

Norsk **Ordbok** 2014

**Ordbok over det norske folkemålet og det
nynorske skriftmålet**

Kvitbok 2012

Norsk Ordbok 2014

Innleiing

Då Stortingsmelding 22 (1999-2000) lanserte Norsk Ordbok som eitt av fire nasjonale oppslagsverk som skulle fullførast til feiringa av Grunnlova i 2014, arbeidde redaksjonen i ordboka ut ei kvitbok for eit slikt prosjekt. Den skulle gje eit samla oversyn over dokument som kasta lys over arbeidet med ordboksverket så langt, og som peika framover. Kvitboka frå 2000 bidrog saman med seminarrapporten *Norsk Ordbok – nynorskens leksikografiske kanon?* frå 1997 med viktig grunnlagsdokumentasjon då planen for prosjektet Norsk Ordbok 2014 blei laga, og kvitboka har også vore eit viktig arbeidsverktøy og sentralt opplæringsdokument i perioden etter at NO 2014-prosjektet blei sett i gang. NO-redaksjonens målsetjing med seminarrapporten og kvitboka var "å leggja fram alle relevante opplysningar, slik at fullføringa av NO under nye vilkår skal verta minst like vellukka og halda minst same faglege nivå som utgjevingane dei siste tiåra" (Kvitbok 2000: 5).

Når det no blir arbeidd ut ei ny kvitbok for prosjekt etter 2014, vil sjølvstilt innhaldet både i seminarrapporten frå 1997 og i kvitbokdokumentet frå 2000 utgjera viktige tilleggsdokument. Det same gjeld forprosjektrapporten for NO 2014-prosjektet frå 2001, utarbeidd av Steinulf Tungesvik på oppdrag frå dåverande Institutt for nordistikk og litteraturvitskap ved UiO. Målet med 2012-kvitboka må vera å dokumentera så nøyaktig som mogleg kva kompetanse, metodar, verktøy og tilfang NO 2014 rår over, og å identifisera ny prosjektverksemd desse ressursane peikar fram mot. For å identifisera ny prosjektverksemd er det viktig å ha med ein bolc med offentlege og andre dokument som legg føringar for vilkåra og rammene den vitskaplege leksikografien vil ha etter 2014 (sjå del B av kvitboka).

Prosjektet Norsk Ordbok 2014 har ved utgangen av 2011 til punkt og prikke halde fristane som den ambisiøse og stramme tidsplanen sette for utgjevingsarbeidet. Redaksjonen leverte i mars 2012 band 10 av ordboksverket, og då står to band att før verket er ferdigstilt. Prosjektet kviler tungt på den solide planlegginga og dei presise omfangsutrekningane som la grunnlaget for den bindande produksjons- og bemanningsplanen for prosjektet. Berre eit fullstendig oversyn over mengda av materiale som skal analyserast, saman med klåre rammor for omfanget arbeidet skal ha, gjer det mogleg å planleggja produksjon og måla resultatane så nøye som det er naudsynt å gjera. Berre på den måten kan både redaksjon og oppdragsgjevarar vera sikre på at verksemda kvalitativt og kvantitativt når dei oppsette måla.

Ein annan vesentleg føresetnad for at prosjektet så langt har vore mogleg å gjennomføra etter planen, er den omfattande digitaliseringa av grunnlagsmaterialet som blei gjort på 1990-talet og den påfølgjande indekseringa av materialet i NO 2014s elektroniske *Metaordbok* (jf artiklane Grønvik 2005, Ore & Tvedt 2005 og Runde m.fl. 2005 i del A av kvitboka). Ikkje mindre viktig er den fullelektroniske redigeringsplattforma som prosjektet no har nytta i nærare 10 år, og som redaksjonen har vore med på å utforma i nært samarbeid med Eining for digital dokumentasjon (EDD). Det er EDD som har levert infrastrukturløysingane til prosjektet, og den kontinuerlege utviklinga av dei digitale ressursane har skjedd i nært samarbeid mellom fagleksikografi og vitskapleg

programmering (jf Bakken & Grønvik 2008, Grønvik 2005, Grønvik 2006 og Ore & Tvedt 2005 i del A av kvitboka). Eit omfattande og solid redaksjonelt opplæringsprogram, ein detaljert redigeringsmanual og ein ditto programmanual er også vesentlege for gjennomføringa av prosjektplanen. Både redigeringsmanualen og programmanualen inneheld dokumentasjon som bør ha stor overføringsverdi til nye prosjekt.

Organiseringa av verksemda som eit eige prosjekt og dei omfattande rasjonaliseringstiltaka som er sette i verk, har gitt NO 2014 gode rammer til å halda jamn fart, og dei har gitt oss kontroll med verkemidla me har hatt til rådvelde for å halda produksjonen oppe. Ein annan faktor som har vore vesentleg for prosjektgjennomføringa, er at ordboksmanuskripta er godkjende som vitskapleg produksjon på nivå 1 i det nasjonale teljekantsystemet for publiserte forskingsresultat. Bidraga i Norsk Ordbok er dermed meritterande for den einskilde redaktøren.

Dokumenta i denne kvitboka, supplert med kvitboka frå 2000, forprosjektrapporten frå 2001, redigeringshandboka og programmanualen til NO 2014, skal spela inn til fleire post 2014-prosjekt. Utvalet av dokument i 2012-kvitboka skal dekkja fagmateriale som gjeld NO 2014-prosjektet og som vil ha overføringsverdi i planlegginga av nye prosjekt, og det skal omfatta offentlege dokument og utreiingar om språkstoda i Noreg og om behova på det leksikografiske fagfeltet. Det skal òg dekkja fagmateriale som gjeld konkrete post 2014-prosjekt. Eit utval konkrete prosjekt er nemnde i punkta a-d nedanfor, men desse representerer ikkje ei uttømmande liste:

- a) **Fullføring av NO-databasen.** Tiltaket gjeld revisjon og elektronisk fullintegrering av band 1-4,5 av Norsk Ordbok. Desse første banda av ordboka er produserte før den fullelektroniske redigeringsløyisinga låg føre. Banda treng innhaldsrevisjon, og det må for kvar artikkel gjerast tilpassingar til det moderne, stringente databaseformatet. Slik den elektroniske basen framstår ved utgangen av prosjektperioden, omfattar han berre alfabetbolken frå I til Å. For at den store og viktige elektroniske ordboksressursen skal bli mogleg å bruka til andre føremål som gjeld leksikografi og språkdokumentasjon, må alfabetbolken A-H også vera integrert i basen.
- b) **Arbeidet med dokumentasjon av moderne bokmål.** Institutt for lingvistiske og nordiske studium har under planlegging eit omfattande prosjekt for dokumentasjon av moderne bokmål. St.meld. 35 (2007-2008, kap. 8.3.6.3) legg føringar på dette prosjektet, i form av at departementet gjev uttrykk for ei forventning om at kompetanse og ressursar frå NO 2014 skal takast med inn i bokmålsprosjektet. Føringane gjer at 2012-kvitboka til NO 2014 òg vil vera ein viktig arbeidsreiskap i forprosjekteringa av bokmålsprosjektet.
- c) **Arbeid med andre leksikografiske prosjekt og produkt.** Det Norske Samlaget, som står bak utgjevinga av Norsk Ordbok, signaliserer at dei svært gjerne vil vera med på å utvikla nye nynorskleksikografiske produkt etter fullføringa av NO 2014-prosjektet. Det finst mange gode idéar til konkrete prosjekt som den ferdige NO-basen kan levera materiale til (t.d. norsk natur sett gjennom norske dialektar, ordbok for norsk mat- og måltidskultur osv.). Også for desse føremåla er det viktig at prosjektorganisasjonen dokumenterer kva ressursar og metodar som står til rådvelde.

- d) **Kontinuerleg vedlikehald, oppretting av feil og innarbeiding av tillegg** i den komplette NO-basen er ei permanent institusjonsoppgåve for Universitetet i Oslo. Fagdokumenta i 2012-kvitboka vil vonleg støtta opp under arbeidet med å etablere ein organisasjon som skal ha ansvar for desse arbeidsoppgåvene etter 2014.

Fagartiklane som er tekne inn under del A) i kvitbokdokumentet, dekkjer ulike sider ved leksikografisk dokumentasjonsverksemd. Bakken 2006, Berg-Olsen & Wetås [i kjømda], Dysvik 2011, Knudsen 1948, Runde 2011, Vikør 2003 og Wetås 2011 tematiserer planar for og arbeid med leksikografiske prosjekt og produkt, medan Bakken & Grønvik 2008, Gismerøy & Ore 2008, Grønvik 2005, Grønvik 2006, Ore & Tvedt 2005 og Runde m.fl. 2005 dekkjer sentrale moment i utviklinga av moderne leksikografisk metode og ditigale verktøy.

Dokumenta i del B) gjeld planar og strategiar for dokumentasjon av norsk språk (Brev til KUD frå Nynorsk forum 4.7.11, Ore 2011, BRO-avtalen 2011, St.meld. 35 (2007-08) kap. 8.3, UiO Strategi 2020). To av dokumenta i denne bolken (Ore 2011 og Grønvik 2011) gjeld planar for elektronisk fullintegrering av eldre ikkje-digitalt produsert ordbokstekst.

Dokumenta i A) og B) vil samla og kvar for seg vera viktige for vidare arbeid med dei nemnde prosjektidéane ovanfor (a-d).

Me håper at artikkelsamlinga i denne kvitboka vil stå som dokumentasjon både av den sterke viljen til å sjå framover, den faglege iveren til å prestera innanfor definerte og strenge rammer og den velfungerande og effektive organisasjonsstrukturen i NO 2014-prosjektet. I tillegg ønskjer me at kvitboka skal dokumentera den vitskapelege metoden, materialbehandlinga, dimensjoneringsmodellane og dei høgteknologiske verktøya som så langt har gjort det mogleg for oss å gjennomføra det store vitskapelege løftet Norsk Ordbok 2014 i tråd med planen.

Åse Wetås

prosjektleder NO 2014

03.02.12

INNHALDSOVERSYN

A) FAGARTIKLAR **9**

Bakken, Kristin & Oddrun Grønvik. 2008. Materialsortering på digital plattform. Eit steg mot høgare dataintegritet i den vitenskaplege leksikografien. I <i>Nordiske studier i leksikografi 9</i> : 31–42	11
Bakken, Kristin. 2006. Relevans, legitimitet og resultatkrav – eit komparativt blikk på utfordringane for dei store nasjonale ordboksverka i Norden i dag. I <i>Nordiske studier i leksikografi 8</i> : 11–24	23
Berg-Olsen, Sturla & Åse Wetås. <i>Norsk Ordbok</i> på nett: utfordringar og løysingar. Kjem i <i>Nordiska studier i leksikografi 11</i>	37
Dysvik, Sylvi. 2011. Nynorsk fagleksikografi. I <i>DKNVS. Skrifter 1/2011</i> : 39–64	51
Ekker, Mette Gismerøy & Christian-Emil Ore. 2008. Norrøn metaordbok. Rapport fra en øvelse. I <i>Nordiske studier i leksikografi 9</i> : 89–104	77
Grønvik, Oddrun. 2005. Redigeringsprogrammet for Norsk Ordbok 2014. I skjæringspunktet mellom menneskeleg skjønn og automatisering. I <i>Nordiske studiar i leksikografi 7</i> : 156–65	93
Grønvik, Oddrun. 2006. Verknader av digitalisering på materialvurdering, redaksjonell metode og opplæring. I <i>Nordiske studier i leksikografi 8</i> : 129–42	103
Knudsen, Trygve. 1948. Det norske litterære ordboksverk i 25 år. <i>Maal og minne</i> : 1-11	117
Ore, Christian-Emil & Lars Jørgen Tvedt. 2006. Digital sats eller digital satsing? I <i>Nordiske studier i leksikografi 8</i> : 315–22	129
Runde, Ålov m.fl. 2005. Metaordboka som reiskap for lemmaseleksjon og ordboksproporsjonering. I <i>Nordiske studiar i leksikografi 7</i> : 326–30	137
Runde, Ålov. 2011. Norsk, særleg nynorsk, dokumentasjonsleksikografi i framtida. I <i>DKNVS. Skrifter 1/2011</i> : 201–10	143
Vikør, Lars. 2003. Norsk Ordbok ved ein krossveg. Rapport frå eit omstillingsprosjekt. I <i>Nordiske studier i leksikografi 6</i> : 391–6	153
Wetås, Åse. 2011. Norsk Ordbok og 2014-prosjektet. i <i>DKNVS. Skrifter 1/2011</i> : 65–84	159

B) ANDRE RELEVANTE DOKUMENT **179**

Brev frå Nynorsk Forum til Kulturdepartementet 04.07.2011	181
Grønvik, Oddrun. 2011. Digitalisering og tilrettelegging for innlasting i ordboksbasen av Norsk Ordbok band 1-5 (a-h). Upublisert rapport, NO 2014	185
Ore, Christian-Emil. 2011. Notat om leksikografi ved ILN etter 2015	191
Samarbeidsavtale for prosjektet Bokmålets og riksmålets ordbase (BRO). 2011	201
Stortingsmelding 35 (2007-08). Kapittel 8.3. Språkleg dokumentasjon	203
Universitetet i Oslo. Strategi 2020. Et grensesprengende universitet [om infrastruktur]	212

Del A

Fagartiklar

Materialsortering på digital plattform

Eit steg mot høgare dataintegritet i den vitskaplege leksikografien

For Norsk Ordbok 2014 linguistic documentation comes in the forms of scanned paper slips, electronic slips and corpus quotations, organised through the index system called the Meta Dictionary. Each entry in Norsk Ordbok is directly linked to the corresponding materials in the Meta Dictionary. However, with up to several thousand instances to sort into a complex entry structure, editors needed a sorting tool for classifying instances and storing the sorting itself. This would both improve editorial stringency, allow for variation in approach method and save time. This article, entitled *Classifying instances of linguistic documentation on a digital platform – one step further towards data integrity within academic lexicography*, describes the tool developed for Norsk Ordbok to meet this need, and sums up user experiences through the first two years as positive but still with space for improvement.

1. Innleiing

I 2002 blei redaksjonen ved Norsk Ordbok omorganisert. Omlegginga innebar nyordning både når det galdt administrasjon og struktur, finansiering, faglege løysingar og ikkje minst digitalisering og utvikling av nye verktøy. Ëi viktig erfaring vi har gjort desse fem første åra, er at dei faglege omleggingane og alle tiltaka som har med digitalisering å gjere heng uløyseleg saman. Vi vil her sjå nærare på det verktøyet som vi i redaksjonen av Norsk Ordbok kallar *Sorteraren*. Dette er kanskje det beste dømet på korleis utviklinga av tekniske løysingar hos oss både heng tett saman med faglege prioriteringar og får faglege konsekvensar.

Eit krav redaksjonen hadde på seg i 2002 var å dra nytte av at kortarkivet vårt alt var digitalisert. Under eit slikt krav låg sjølvstapt primært ønsket om effektivisering. Dei første åra vart det då utvikla ein redigeringsapplikasjon som hadde direkte kontakt med det digitaliserte arkivet via den digitale indeksen for språksamlingane, kalla Metaordboka (for nærare omtale sjå Ore 2000). Ordutval og dimensjonering blir no gjord delvis automatisk med utgangspunkt i det digitaliserte arkivet, og underlagsmaterialet er tilgjengeleg på skjerm. I teorien treng altså ikkje lenger redaktørane forhalde seg til skuffar med arkivmateriale. Men eitt problem stod att: korleis sortere digitale arkivkort? Mange av redaktørane

såg seg framleis nøydde til å sortere underlagsmaterialet på skrivebordet i små bunkar med arkivkort. Dette problemet blei akutt for den delen av redaksjonen som arbeider ved universitetet i Trondheim, for det hadde ikkje tilgjenge til det fysiske arkivet, og var såleis nøydde til å bruke dei digitaliserte arkivkorta, som dei berre fekk sjå eitt for eitt på skjermen.

Mot denne bakgrunnen gjekk vi i dialog med utviklarane ved Eining for digital dokumentasjon. Vi ville ha ein digital sorterar som løysing på eit praktisk problem. Sorteraren skulle vere eit tenleg verktøy i materialanalysen, og helst skulle ein spare tid i denne arbeidskrevjande fasen av det redaksjonelle arbeidet. Vi ønskte også å hente kvalitative gevinstar, dvs. å sikre at kvar og ein gjorde så godt empirisk grunnarbeid som mogleg. Metodisk såg vi her ei moglegheit for at Norsk Ordbok i større grad kunne bli empirisk etterprøvbar, med andre ord at dataintegriteten kunne bli høgare med ein digital sorterar på plass. Til sist var det svært viktig for oss at verktøyet vi etablerte ikkje innebar ei einsretting av arbeidsformene. I forkant hadde vi nemleg kartlagt korleis redaktørane faktisk arbeider med materialet. Vi hadde sett at dei arbeider forbausande ulikt. Og her vil vi understreke at dette dreier seg om heilt grunnleggjande tilnæringsmåtar som heng saman med personlegdomen, og at det er umogleg å seie at den eine tilnæringsmåten er "betre" enn den andre. Dette kjem vi tilbake til, i denne omgangen held det med å understreke at vi ønskte å ta vare på denne metodiske diversiteten. Vi ville at det nye sorterarverktøyet skulle bli eit fleksibelt verktøy som kunne tilpassast ulike, individuelle arbeidsformer. Det måtte for all del ikkje bli ei metodisk tvangstrøye. Dette var då ønsket vårt: Vi ville ha eit fleksibelt verktøy som opnar for metodisk diversitet, som sparar tid, hevar kvaliteten og aukar dataintegriteten.

2. Løysing: Eit funksjonsrikt system

2.1 Plassering i NO-systemet

IKT-løysinga for Norsk Ordbok 2014 er eit system av databasar og IT-verkty til ulike føremål. Norsk Ordbok blir redigert i ein applikasjon som matar data rett inn i ein database. Det er lagringsstaden for ordboksartikkelen, som skal vere resultatet av ei mogleg, men ikkje ei obligatorisk sortering. Tilfanget kjem frå Metaordboka og blir hekta rett på artikkelhovudet, og på denne måten er det direkte samband mellom Metaordbok og ordboksdatabase.

Sorteraren er eit tilbod, eit slags siderom på dette sambandet, der ein kan ordne og annotere tilfanget. Sorteraren ligg derfor utanfor redigeringsapplikasjonen og ordboksdatabase. Sorteraren er tilgjengeleg berre frå applikasjonen, via den einskilde artikkelen.

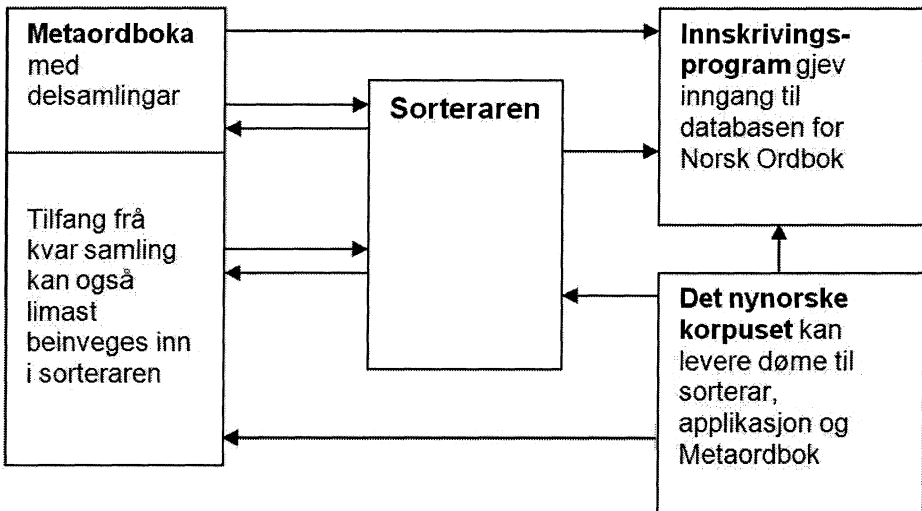
2.2 Utforming

Først eit ord om ein tanke vi forkasta. Det hadde vore mogleg å lage eit simultert "skrivebord" på skjermen, der ein brukte musa til å dra belegg over til rett "bunke". Dette ville vere visuelt attkjenneleg. Men setelbunkane har aldri vore noka ideell løysing. Dei har den veikskapen at ein berre ser øvste setel, og er avhengig i postit-lappar på bunkane for å minnast sorteringskriterium. Det var også eit poeng å unngå musbruk og halde seg mest mogleg til tastaturet.

Vi valde i staden ei løysing som tek vare på den mentale sida ved sorteringsprosessen, nemleg sjølv kategoriseringa. Poenget er då å kunne (1) merkje eller annotere kvar setel, (2) sjølv velje sorteringskriterium, (3) kunne merkje opp setlane etter så mange kriterium som ein vil, (4) kunne sortere framlengs og baklengs etter eitt eller fleire kriterium samstundes, (5) lagre sorteringa som grunnlag for framtidig arbeid.

Sorteraren ser ut som eit rekneark, der venstre kolonne listar opp alle belegg linje for linje. Kvart belegg frå Metaordboka er representert med eit symbol. Når ein klikkar på symbolet kjem setelen opp til venstre på skjermen. Ein kan leggje til kolonnar (for sorteringskriterium) og gje dei tittel etter eige skjøn i høve til kvart ord.

Ein stor føremon ved denne arbeidsmåten er at når tilfanget er ferdig annotert og sortert ser ein straks kva ein har i kvar "bunke", og ein ser heilskapen i sitt eige sorteringssystem. Sorteraren gjer dermed tilfanget langt meir oversiktleg enn ein nokon gong før har hatt det.



Figur 1. NO 2014: Kopling av tilfang til og frå sorteraren. Pilene syner kva veg tilfang kan koplast.

2.3 Innhald, deling og flytting

Når ein opprettar ein sorterar, kan ein hente inn alle belegg som finst i Metaordboka under det aktuelle oppslaget. Men ein kan også hente inn tilfang seinare, frå korpuset eller nytt tilfang som kjem inn i samlingane under Metaordboka. Dette er nyttig både under redigering og manusrevisjon.

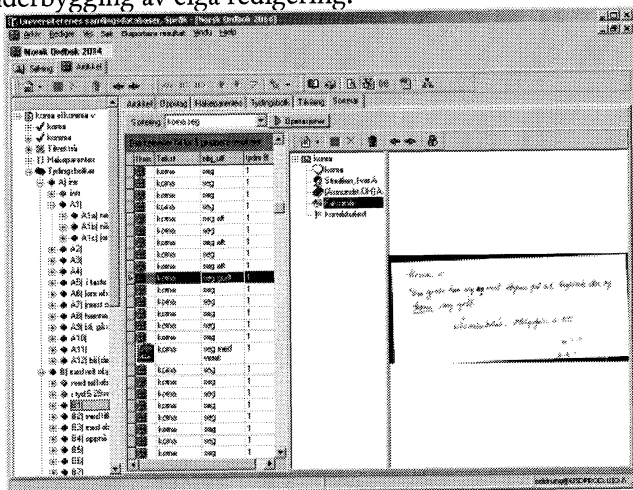
Det kan vere mange grunnar til at éin sorterar bør delast i fleire, frå homograf-separering til å halde styr på til dømes tydingsinndeling i store materialmengder. Dette er lett å få til gjennom den innebygde funksjonsmenyen. En kan også flytte ein sorterar til ein annan artikkel, eller slette ein mislykka sorterar.

Frå sorteraren kan ein kopiere belegg direkte over på alle nodar i artikkeltreet. Dette er særleg kjekt når ein skal grunngje ei uventa form eller ei sjeldsynt tyding frå eit ellers stort artikkeltilfang.

2.4 Skrivning og innsyn

Berre redaktøren kan skrive direkte i sorteraren. Det er såleis hans eller hennar private notatblokk. Men Norsk Ordbok er eit ope prosjekt der alt heng saman med alt. Andre redaktørar kan ha bruk for å sjå korleis ein kollega arbeider med eit visst ord; systemet for kvalitetskontroll av manus kan ha bruk for å sjekke både heilskap i materialbruk og ein detalj her og der. Derfor kan alle redaktørar sjå på alle sorterarane attom kvar artikkel, om dei vil.

Når Norsk Ordbok ein dag blir ei elektronisk ordbok, vil det vere mogleg å gjere sorterarane tilgjengelege for innsyn om ein så vil. Det som i alle fall blir tilgjengeleg, er den dokumentasjonen som er knytt direkte til nodar i artikkeltreet. Førebels avgjer kvar redaktør kor mykje dokumentasjon han eller ho vil leggje fram som underbygging av eiga redigering.



Figur 2. NO 2014: Sorterar til verbet *koma* som dekkjer sambandet "koma seg" (Tyding nr B1). Til venstre ser ein artikkeltreet, med tyding B1 markert, til høgre ser ein det belegget som er markert i sorteraren.

3. Resultat: erfaringar og evalueringar

Intensjonen ein har med ein applikasjon blir ikkje alltid realisert av brukarane slik ein hadde tenkt. Etter knappe to år med sorteraren på plass var det då svært interessant for oss å kartleggje bruken av han. 25 redaktørar i to ulike byar har fått eit tilbod om å bruke sorteraren. Dei har ikkje fått noko *påbod* om det, for i utgangspunktet skulle sorteraren dekkje eit uttalt behov hos nokre, men det var jo slett ikkje alle som hadde etterlyst ein digital sorterar. Etter knappe to år er det akkumulert 1446 artiklar som har sorterarar knytte til seg. Det er opplagt at redaktørane i ulik grad og på ulik måte har tatt dette verktøyet i bruk.

Artiklar med sorterar	færre enn 5	10–60	60–100	fleire enn 100
redaktørar	4	13	4	4

Figur 3. Redaktørane ordna i grupper etter kor mange artiklar dei har oppretta ein sorterar for.

Ved å gå databasen etter i saumane kan ein finne ut ganske mykje om bruken av sorteraren. I tillegg har vi intervjuet nokre av redaktørane i kvar gruppe for å få svar på desse spørsmåla:

- Når opprettar ein redaktør ein sorterar?
- Korleis blir sorteraren nytta?
- Sparar ein tid?
- Er kvaliteten på arbeidet blitt høgare?
- Er dataintegriteten blitt betre?

Når ein analyserer svara på desse spørsmåla, ser ein om intensjonane bak utviklinga av sorteraren er blitt realiserte.

3.1 Når opprettar ein redaktør ein sorterar?

Som venta er det slik at ein sorterar blir oppretta når ein redaktør skal redigere eit "stort" ord, dvs. eit ord med stort materiale bakom seg i arkivet. Dette hadde vi rekna med, for ein når jo ei smertegrense for kor mykje informasjon ein kan greie å ha oversyn over i hovudet. På eit eller anna tidspunkt må ein lage seg "hugselappar". Men det vi ikkje hadde rekna med, er kor store forskjellar det er på kva redaktørane reknar som "store" ord, med andre ord kor mykje materiale den einskilde tykkjer det er greitt å halde oversyn over utan hjelpemiddel. Nokre definerte eit stort ord som eit ord med meir enn 20 belegg attom seg i arkivet, medan andre definerte eit stort ord som eit ord med fleire enn 250 belegg attom seg.

Ein tilleggsgrunn for fleire var forventninga om kompleks tydingsstruktur. Om ein raskt ser at ordet kjem til å bli redigert med mange tydingsnummer fordi tydingsstrukturen er kompleks, opprettar redaktøren ein sorterar sjølv om ordet ikkje har svært mange belegg attom seg i arkivet. Det same gjeld for ord som må

homografseparerast, dvs. at belegga under eitt oppslag i arkivet faktisk skal fordelast på to forskjellige homografar og dermed to forskjellige oppslagsord.

Til sist nemnde fleire at typiske dialektord, dvs. ord med stor målførevariasjon, ofte får knytt ein sorterar til seg. Norsk Ordbok skal jo omfatte både den normerte skriftlege nynorsken og dei munnlege dialektane. Det gjer ein ved å ordne og føre opp dialektvariasjon under kvart oppslag. I mange tilfelle er formvariasjonen svært stor, og i slike tilfelle blir sorteraren nytta for å kartleggje denne typen av variasjonsmønster.

Ikkje uventa ser det altså ut som om ein sorterar blir oppretta for å behalde oversyn over komplekst materiale. Fleire seier eksplisitt at dei alltid vurderer tidsbruken ved å opprette ein sorterar. Der dei trur dei greier å halde oversynet utan sorterar, lét dei vere å opprette ein, for dei trur altså at dei sparar tid på denne måten. Det må også seiast at dei som relativt rutinemessig opprettar ein sorterar for alle ord som har meir enn 20-30 belegg attom seg, har integrert sorterarbruken i den vanlege arbeidsrutinen sin som redaktør. Vi kjem tilbake til erfaringane desse redaktørane har gjort seg.

3.2 Korleis blir sorteraren nytta?

Svaret på dette spørsmålet var det ikkje så lett å lese ut av databasen, men etter å ha intervjuet ulike redaktørar, teiknar det seg eit bilete av to ganske forskjellige måtar å bruke sorteraren på. Interessant nok er det ikkje eintydig om desse to ulike måtane fordeler seg på anten storbrukarane eller dei som brukar dette verktøyet minst.

Vi har valt å karakterisere dei to bruksmåtane for ei induktiv tilnærming og ei deduktiv tilnærming. Den reinaste induktive metoden blir av ein av brukarne også kalla "notatblokkmetoden". Redaktøren går her til materialet utan førehandsoppfatningar om struktur, og opprettar ein sorterar "nedanfrå" så å seie. Det vil blant anna seie at kategoriane som vert oppretta får ganske innhaldstunge overskrifter, og at samstundes med at materialet blir gjennomgått, blir kategoriseringssystemet raffinert og justert ved at kategoriar blir delte eller at nye kjem til. Redaktøren har ikkje primært fokus på eit ryddig sluttprodukt, og det blir typisk nok ikkje eit ein-til-ein-forhold mellom artikkelstrukturen og strukturen i sorteraren. Om redaktøren skulle ønskje å kople relevant grunnlagsmateriale til kvart tydingsnummer i artikkeltreet, krev det ein eigen ryddeprosess til slutt. Dei redaktørane som arbeider slik, viser til fordelane ved å gå til materialet utan førehandsoppfatningar, for slike oppfatningar kan kanskje vere galne og stå i vegen for den analysen som høver beste med det materialet faktisk viser. Dei meiner nok også at metoden yter materialet størst rettferd.

Den reinaste deduktive metoden går ut på at ein i forkant har laga seg eit grovt kart over forventede relevant semantisk og dialektbasert variasjon, og at ein opprettar ein sorterar med kategoriar som speglar desse forventede variablane. Slike hypotesar om struktur blir etablerte på grunnlag av grunnmanuskriptet til Norsk

Ordbok (ein kompilasjon av alle dei eksisterande ordbøkene over nynorsk pr. 1940), og på bakgrunn av andre ordboksverk, grammatikkar og dialektoversyn.

Med denne arbeidsforma blir sorteraren nytta til å teste og justere hypotesar som er formulerte på førehand, og ein relaterer sin eigen materialgjennomgang meir eksplisitt til tidlegare arbeid. Metoden gjer det lettare å etablere eit ryddig sluttprodukt, og den gjev typisk nok eit større samsvar mellom sorterarstruktur og ordboksstruktur. I neste omgang blir det lettare å kople materialet i sorteraren til nodane i artikkelstrukturen.

Vi vil likevel understreke at i praksis er det sjølvstøtt oftast slik at metodane i nokon mon glid over i kvarandre. Dei som primært arbeider deduktivt får plutsleg auge på eit særtrekk i materialet som gjer at dei må justere den opphavlege strukturen. Typisk nok er det kanskje dei faste ordsambanda, dvs. dei konvensjonaliserte kollokasjonane som slik "materialiserer seg" undervegs. På den andre sida er ikkje dei som arbeider mest induktivt heller utan føresetnader, og det er sjølvstøtt slik at dei konsulterer tidlegare arbeid før dei set i gang. Dei lét likevel tidlegare arbeid i mindre grad styre eigen analyse.

Det som er heilt klart er at intensjonane våre med å lage eit fleksibelt verktøy som opnar for metodisk mangfald, er oppfylde. Reint konkret er det slik at nokre nummerer tydingane i sorteraren med éin gong, t.d. "1", "2", "2a", "2b" osv. Andre gjev tydingane titlar i form av meir innhaldstunge stikkord, t.d. "lokativ", "temporal", "durativ" osv. Nokre plukkar berre ut målføresetlane og markerer dei t.d. med "Målf.", medan andre annoterer dei ulike dialektformene for seg. Somme gonger blir tekstdøme merkte med å føre opp kollokat, og såleis blir det lagt til rette for ei strengt kontekstbasert sortering. Nokre tilpassar sorteraren slik at dei kan føre inn alle heimfestingane, og alle belegg frå eldre kjelder, medan andre igjen markerer alle idiom. Nokre hakar av dei belegga dei tykkjer er gode kandidatar for å bli tatt med i den endelege ordartikkelen som sitat. Nokre brukar mange kolonner eller fleire sorterarar for å halde orden på alle informasjonstypene, medan andre heller lagar ulike innhaldstunge annoteringar innanfor same kolonne i sorteraren. Felles for alle systema er prinsippet om at all konsekvent annotering gjer det mogleg å søke fram dei aktuelle delmengdene av materialet etterpå.

Dette inneber ei raffinering av tilfanget som også aukar eigenverdien av ord-samlingane for seinare forskning. Om eit stort underlagstilfang for eit einskildord er merkt opp i sorteraren med kollokat, til dømes, blir det enkelt å plukke ut og bruke dei setlane til andre føremål seinare.

Den store fleksibiliteten systemet opnar for, er slik vi ser det ein stor føremon. Det gjer det mogleg for den einskilde redaktøren å vidareføre si eiga arbeidsform. På den andre sida kan det tenkjast at dette opne systemet kan framstå som krevjande for dei av redaktørane som ikkje, eller i svært liten grad, har tatt sorteraren i bruk.

3.3 Sparar ein tid?

Det er viktig å nemne at fire av redaktørane så å seie ikkje har tatt sorteraren i bruk. Dette er fire av dei eldste redaktørane (over 60 år), og sams for dei er at dei viser til at dei for lenge sidan har etablert gode rutinar som fungerer for dei. Dei har halde fram med å sortere papirsetlar på skrivebordet, og meiner det ville føre til unødig tidsbruk å forsere den tekniske terskelen som sorteraren ber med seg. Dei forventar altså ikkje at bruk av sorteraren skal gje nokon tidsgevinst. I samband med denne undersøkinga har vi sett at vi bør bruke litt meir tid på å gje innføring i bruk av sorteraren for dei som argumenterer med tidkrevjande teknisk terskel.

Bortsett frå desse fire redaktørane meiner alle dei andre at sorteraren gjev tidsgevinstar. Som alt nemnt gjeld dette for mange med visse atterhald. Sorteraren gjev tidsgevinstar for visse (store) ord, men ikkje generelt, og redaktørane gjer slike vurderingar av moglege tidsgevinstar frå ord til ord.

Tre tidssparingspunkt blir nemnde. For det første ligg det potensiell tidssparing i det å ikkje miste oversynet over ei stor materialmengd. Utan merking er faren der alltid for at ein må gjere operasjonar fleire gonger fordi ein ikkje finn ting att, eller fordi ein gløymde noko i førre runden. Sorteraren hjelper slik til med at ein raskt kan orientere seg i materialet og utan omvegar finne fram til relevante døme.

For det andre er det svært viktig for mange at sorteraren gjer det mogleg å annotere og kategorisere etter fleire kriterium samstundes. Når ein sorterer arkivbelegg i haugar på eit skrivebord, blir det eit problem om det finst fleire måtar å kategorisere materialet på som ikkje er synkroniserte. Typisk nok er dette tilfellet med dialektbasert, kjeldebasert eller heimfestingsbasert kategorisering i møte med semantiske kategoriar. Papirløysinga er å sortere i fleire omgangar. Ulempen med denne metoden er at sorteringane blir borte etter kvart. Ein får ikkje lagra og slik rekapitulert sorteringane om ein har bruk for det. Med den digitale sorteraren kan ein sortere etter like mange kriterium som ein i utgangspunktet har brukt i annoteringa. Og ein kan samsortere etter fleire kriterium, t.d. etter heimfestingsstad og dialektform, eller etter dialektform og tydingsnummer. Ved dei store orda gjev dette vesentlege tidsgevinstar.

Det faktumet at desse sorteringane blir elektronisk lagra saman med artikkelen dei utgjer det empiriske grunnlaget for, blir no utnytta i opprettings- og kontrollfasen. Dette er det tredje punktet der mange redaktørar viser til tidssparing. Når ein kontrollør skal vurdere materialanalysen, eller ein i opprettingsfasen ikkje lenger hugsar kva grunnlaget var for vala ein har gjort, er det ein føremon at materialgjennomgangen faktisk er lagra. Både i funksjon av å vere redaktør som skal rette opp eigne manuskript eller i funksjonen som gjennomlesar som skal kvalitetssikre manusa til andre, peikar redaktørane på tidsgevinstane ein sorterar fører med seg.

Vi meiner då at det er riktig å konkludere med at bruk av sorteraren gjer at ein sparar tid. Men det føreset altså at ein brukar sorteraren slik som no, dvs. etter ei ganske gjerrig vurdering av tidsbruken når ein vel å opprette ein sorterar i utgangspunktet. Eit par av dei som i størst grad har integrert bruk av sorteraren i arbeidsrutinen sin, ser tidsgevinstane i opprettingsfasen som så viktige at dei også vel å opprette sorterarar til etter måten små ord.

3.4 Er kvaliteten på arbeidet blitt høgare?

Alle dei som brukar sorteraren meiner kvaliteten på arbeidet er blitt høgare. Dei viser til litt ulike element i kvalitetsomgrepet. Éin viktig grunn for mange er at dei kjenner at det empiriske grunnlaget for den analysen ordboksartikkelen er eit uttrykk for, er meir solid. Dei er tryggare på at dei har tatt omsyn til alle nyansane i materialet, og sikrare på at dei ikkje har gått glipp av vesentlege døme eller kollokasjonar. Éin redaktør viste til at sorteraren hjelper til med å etablere ei naturleg rekkjefølgd på tydingsnumra. Ein får lettare oversyn over kor prominente dei ulike tydingane er i materialet, og reint visuelt får ein indikasjonar på kva for tydingar som er primære og sekundære.

Dei som arbeider mest induktivt, held det fram som ein kvalitativ føremon at ein kan møte materialet utan hypotesar og slik yte det betre rettferd. Det kan kanskje lesast ut som at den induktive arbeidsforma gjer redaktørarbeidet til eit meir sjølvstendig vitskapleg forskararbeid, med resultat som er etablerte med uavhengig evidens i forhold til tidlegare ordboksarbeid. Vi vil understreke at det same gjeld for ei moderat deduktiv tilnærming. Om ein er open for å justere og endre dei hypotesane om struktur ein hadde då ein byrja å sortere materialet, vil ein ved hjelp av sorteraren gjere empirisk grunnlagsarbeid uansett.

Eit svært viktig punkt i høve til dataintegritet er det som har med kvalitetskontroll å gjere. Det er heilt opplagt at kontroll og oppretting av eigne manuskript er blitt betre fundert. Ein slepp å byrje på nytt med materialet når ein kvalitetskontrollerer manuskripta sine. I staden kan ein lett ta ein titt i sorteraren om det er noko ein lurar på. Det same gjeld for dei som kvalitetsikkrar andre sine manuskript. Spesielt viktig er kanskje dette for redaktørar som er i ein opplæringsfase. Ein svært viktig del av opplæringa er jo å lære seg å analysere materiale rett. No kan dei som har ansvar for manuskripta til nybyrjarane, på ein lettvint måte også kvalitetssikre materialanalysen deira. Dette er blitt eit viktig punkt i redaksjonen vår, for så langt har vi i desse fem åra hatt nye redaktørar under opplæring kontinuerleg. Sjølv om vi no nærast har full bemanning, ser vi at det i ein stor redaksjon vil vere utskifting jamleg, og då gjeld det å gjere opplæringa av nye redaktørar mest mogleg effektiv.

Vi ser slik at det som i utgangspunktet var tenkt som ei løysing på eit praktisk problem, faktisk også har gitt kvalitetsgevinstar vi ikkje hadde sett så klart for oss då vi byrja med arbeidet. Vi hadde sett for oss at det ville bli lettare å skaffe seg oversyn og verre å gå glipp av relevant materiale, men vi hadde ikkje så klart sett

kva for gevinstar dei lagra sorteringane ville ha i kontroll- og kvalitetssikringsarbeidet.

3.5 Er dataintegriteten blitt betre?

Norsk Ordbok er ei vitskapleg ordbok. Grunnen til det er at ein tek sikte på å kartleggje og dokumentere og analysere ord som anten ikkje er kartlagde før i det heile, eller som ikkje er kartlagde på denne måten før. Dokumentasjon er her eit nøkkelord. I ei dokumentasjonsordbok skal ein ved hjelp av materiale eller kjelder underbygge dei påstandane om ordtyding, utbreiing og formell variasjonsbreidd ein kjem med i ordartiklane. Eit krav til god vitskap er at den er empirisk etterprøvbar. Det vil seie at dei påstandane eller konklusjonane ein legg fram, skal kunne prøvast mot grunnlagstilfanget, og avkreftast om tilfanget er brukt urett. Det skal vere mogleg å finne og leggje fram motprov mot den konklusjonen eller den påstanden ein kjem med.

I eit vitskapleg dokumentasjonsordbok er påstandane eller konklusjonane som ein legg fram om språk, implisitt formulert. I den ferdige ordartikkelen vil tydingshierarkiet som ein har ordna tydingane til ordet i typisk nok vere ein slik konklusjon. Ein annan konklusjon vil vere det oversynet over distribusjonen av dialektvariantar som blir presentert. Ein tredje vil vere det sambandet som er etablert mellom oppslagsform, språkhistorisk opphavsform og dialektvariantar i artikkelen, det vil seie sjølve avgrensinga av lemmaet. I mange tilfelle vil dette sambandet bli etablert for første gang i ei ordbok av vår type som handterer både målføre, normert tale og skrift.

Spørsmålet blir då om sorteraren har gjort Norsk Ordbok meir *vitskapleg* enn ho var før. Ein annan måte å stille dette spørsmålet på er å spørje seg om dataintegriteten er blitt betre. Med dataintegritet meiner vi her at det er eit godt samsvar mellom konklusjonane om materialet i ordboka og materialet sjølv. Og vi vil leggje til — dette samsvaret skal også kunne etterprøvast.

Det viktigaste einskildelementet i den nye Norsk Ordbok-redaksjonen som betrar dataintegriteten, er faktisk ikkje sorteraren, men Metaordboka. At det digitaliserte arkivet gjennom Metaordboka er direkte kopla til den elektroniske ordboksdatatabasen, gjer at ein kan gå fram og tilbake mellom ordartikkel og arkiv, og slik *etterprøve* ordartiklane. Dette er ein revolusjonerande ny måte å tenkje dokumentasjonsordbok på, og det er ikkje før Norsk Ordbok er tilgjengeleg for alle som elektronisk ordbok, at dette potensialet blir fullt ut realisert. For oss i redaksjonen er det likevel allereie ein realitet, vi har ei direkte kopling mellom empiri og ordartikkel, og vi utnyttar dette til etterprøving i det daglege kvalitets-sikringsarbeidet.

Sorteraren representerer i denne samanhangen ei raffinering av denne koplinga mellom materiale og ordbok. Koplinga mellom materialet i Metaordboka og artiklane i ordboksdatatabasen er jo ikkje nærare spesifisert og fortolka. Den som skal gå analysen etter i saumane må i og for seg gjere heile det leksikografiske

grunnlagsarbeidet om igjen. Sorteraren bøter på dette, for i sorterarane er det jo nettopp spesifiserte og fortolka analysar av tilfanget som blir lagra saman med artikkelen. Ved å kople sorterarane til ordboka, ja gjerne til dei einskilte nodane i artikkelstrukturen, er det etablert ei moglegheit for etterprøving på meir detaljert nivå. I dei tilfella der ein sorterar er kople direkte mot kvar node i artikkelstrukturen, er kvar node empirisk fortolka mot eit større materiale enn det illustrerande dømet som før fekk ivareta denne funksjonen.

Før vi konkluderer på dette punktet, må vi understreke at undersøkinga vår har vist at denne eksplisitte oppkoplinga mellom nodane i ordartiklane og sorteraren til vanleg ikkje blir gjort. Med eitt unntak er det ingen redaktørar som tek den ekstrajobben det er å kople sorteraren opp mot ordartikkelen, node for node. Hovudgrunnen har med tidsbruk å gjere. Slik oppkopling tek tid som ikkje vil gje gevinstar for produksjonen av papirutgåva. Omsynet til ei framtidig elektronisk utgåve er ein luksus redaktørane ikkje unner seg med dei krava til framdrift som dei har å forhalde seg til. Den andre grunnen er at dei som i størst grad brukar sorteraren som ei nedanfrå-og-opp notatblokk, ikkje har etablert ein så tydeleg struktur i sorteraren at ein utan vidare kan kople den opp mot ordboksdatabase. Her må det verkeleg eit reelt tidkrevjande ryddearbeid til før slik kopling kan gjerast.

Idealet om auka dataintegritet er først og fremst realisert på den digitale plattformen i og med Metaordboka. Opprettinga av sorteringar i samband med artiklane i databasen inneber at etterprøvbarheita er realisert på eit meir spesifisert nivå. I og med oppkoplingsmoglegheita som ligg i systemet, kunne ein tenkje seg ei endå tettare og meir fortolka kopling mellom materiale og ordartikkel, men av grunnar som har med tidsbruk å gjere, er det berre éin redaktør som gjer slik oppkopling. Det er viktig å understreke at dette idealet om dataintegritet berre lèt seg realisere på den digitale ordboksplattformen.

4. Konklusjon

For oss er det opplagt at ei digital kopling mellom materiale og ordbokslemma via Metaordboka gjev auka dataintegritet. Det er blitt mogleg å etterprøve det redaksjonelle arbeidet ved å gå fram og tilbake mellom ordboksdatabase og det digitalisert arkivet. I tillegg meiner vi at dei lagra sorterarane som no i mange tilfelle er hekta på ordboksartikkelen i databasen, gjer dei konklusjonane som ordboksartikkelen er eit uttrykk for, etterprøvbare på eit endå meir spesifisert nivå. I og med sorteringane er materialet fortolka, og desse fortolkningane er lagra. Det høgaste nivået av integrering mellom materiale og ordbok får ein ved å kople sorteringane til artikkel-treet node for node. Dette er ei moglegheit som redaksjonen så langt i liten grad nyttar seg av.

Det er slik eit stort arbeid som ligg framfor oss om den moglegheita vi no har fått til å integrere kjelder og ordboksdatabase skulle realiserast gjennomgåande.

I ein framtidig digitalisert versjon av ordboka kunne vi eksplisitt kople sorteringsresultatet til nodane i artikkel-treet slik at materialanalysen også blei tilgjengeleg for brukarane. Men dette set store krav til sorteringane. Dei må då vere ryddige og ferdiganalyserte med ein struktur som er i samsvar med den ferdige artikkelen. Dette inneber heilt klart eit stort meirarbeid. Om ein verkeleg skulle ha som ambisjon at brukarane av ordboka ved alle ord skulle få tilgjenge til oppkoplale materiale, ligg det dessutan ei kjempestor utfordring på metodesida. Etter vår meining måtte ein då uniformere sorterarbruken i større grad enn no, og dette ville gripe radikalt inn i arbeidsformene til den einskilde. Vi ser altså at dei ideelle krava til dataintegritet er i konflikt for det første med ønsket om effektiv tidsbruk og for det andre med ønsket om individuelt tilpassa verktøy og metodisk mangfald.

Vi vil likevel konkludere med at dei intensjonane vi hadde når vi sette i gang arbeidet med å utvikle ein sorterar, langt på veg er realiserte. Vi har fått eit fleksibelt metodisk hjelpemiddel, verktøyet har gitt tidsgevinstar, kvaliteten på ordboksproduktet er blitt høgare og til sist, dataintegriteten er blitt betre. Vi vil halde fram med å oppmuntre redaksjonen til å realisere det potensialet for betre metode som ligg i sorteraren, men vi vil av omsyn til metodisk fleksibilitet og effektiv tidsbruk førebels ikkje fullt ut realisere den optimale etterprøvbarheita for eksterne brukarar.

Litteratur

Ore, Christian-Emil Smith 2000: Metaordboka. I: *Ord om Ord* 6, 30–32.

Kristin Bakken
prosjektdirektør, f. 1959
kristin.bakken@iln.uio.no

Oddrun Grønvik
hovudredaktør, f. 1947
oddrun.gronvik@iln.uio.no

Norsk Ordbok 2014
Universitetet i Oslo
postboks 1021 Blindern
NO-0315 Oslo

Relevans, legitimitet og resultatkrav – eit komparativt blikk på utfordringane for dei store nasjonale ordboksverka i Norden i dag

Relevance, Legitimacy and Demands of Completion – a comparison of the ways the national dictionaries in Scandinavia are challenged today. Five of the national dictionaries in Denmark, Sweden and Norway are presented in order to compare them with regard to the challenges they meet in contemporary society. Several of them were conceptualised several decades or even more than a century ago, meaning that they were motivated both by different cultural contexts and by different scientific ideals. At the same time these projects are extremely costly, thus demanding public funding. The paper focuses on four challenges pertaining to this situation: the challenge of securing the financial foundations, the challenge of giving the projects legitimate motivation, the challenge of enhancing their social and scientific relevance, and the challenge of dealing with demands of completion.

1. Innleiing

Emnet for føredraget blei valt fordi eg har ansvar for eitt av dei nordiske nasjonale ordboksprosjekta, og strevar dagleg med å sameine dei faglege måla for arbeidet med dei rammvilkåra vi har å halde oss til. Her ligg det mange utfordringar, for når eg brukar ord som ”streve”, impliserer det at dei faglege måla ikkje utan vidare lèt seg harmonisere med dei rammene som er sette for arbeidet vårt. I eit prosjekt som Norsk Ordbok 2014, ville ein kanskje som fagentusiast kunne ønskje seg så store og så oppdaterte og lett tilgjengelege samlingar som mogleg, uendeleg mykje papirplass å boltre seg på, og dessutan romslege økonomiske rammer som gjev rom for ”tilstrekkeleg” mange redaktørstillingar og støttepersonell. Dessutan ville ein jo gjerne ha optimale IT-verktøy, høvelege lokale, og ein ville helst sleppe stressande tidspress og ytre krav om ferdigstilling. På den andre sida er det klart at vi ideelt sett gjerne vil ha eit stort publikum, for vi vil jo også at ordboksproduktet skal vere etterspurt.

For Norsk Ordbok 2014 like lite som for dei andre nordiske ordboksverka, er situasjonen slik. I staden står vi overfor utfordringar på dei fleste av dei felte eg alt no har nemnt. Det gjeld utfordringar som har med grunnlagsmaterialet og arkivsituasjonen å gjere, plassproblem ved at line- og sidetalet ein har til disposisjon i dei trykte utgåvene er avgrensa, det er tids- og kostnadskrevjande å utvikle høvelege IT-verktøy, å sikre finansiering av prosjekta er ei konstant utfordring, og når det gjeld tilhøva til samfunnet, har vi relevans- og legitimitetsproblem samstundes med at vi har utfordringar med å formidle verksemda og produkta våre til potensielle brukarar.

Mot denne bakgrunnen tykkjer eg at det vil vere både interessant og lærerikt å samanlikne vilkåra for nokre av dei store ordboksverka i Norden i dag. Målet med ei slik komparativ tilnærming er for min eigen og ideelt sett også for andre sin del, å få idear til korleis ein best kan møte dei utfordringane det her er snakk om.

2. Ordbøkene

Dei fem ordbøkene eg vil samanlikne er forutan *Norsk Ordbok* (NO): *Norsk riksmålsordbok* (NRO) i Noreg, *Svenska Akademiens ordbok* (SAOB) i Sverige, og *Ordbog over det danske Sprog* (ODS) og *Den Danske Ordbog* (DDO) i Danmark. Eg ville ideelt sett gjerne ha inkludert ordboksprosjekt frå Finland og Island i samanlikninga, men det hadde gjort oppgåva for stor gjeve dei tidsrammene eg har hatt til rådvelde. Andre ordboksverk kunne også ha vore trekte inn, t.d. *Nusvensk ordbok* (1917-72), men medan arbeidet med denne ordboka er avslutta, er dei fem ordbøkene eg kjem til å drøfte, alle i meir eller mindre grad å rekne som igangverande prosjekt. Før eg held fram med sjølve samanlikninga, vil eg takke redaksjonane i Lund og København og NRO-redaksjonen i Oslo for all velviljug hjelp i samband med førebuinga av føredraget.

Statusen til dei fem aktuelle ordbøkene ser i utgangspunktet litt forskjellig ut. ODS, NRO og DDO er "fullførde" ordbøker, ODS og NRO endåtil med supplementsband, SAOB er eit gammalt, vakse prosjekt med 34 band utgjevne (september 2005) og med bokstaven *t* under redigering. NO er eit yngre og mindre omfattande prosjekt enn SAOB; første band av NO kom alt i 1966, men pr. i dag (2005) er vi ikkje komne lenger enn til at fem av tolv planlagde band er utgjevne, og bokstaven *k* er under redigering. Samanlikninga av rammevilkår vil difor til ein viss grad bli basert på historiske tilhøve. På den andre sida er no arbeidet i gang med å vidareføre både NRO, ODS og DDO, og vilkåra for denne vidareføringa vil bli trekte inn i drøftinga nedanfor. Også NO og SAOB bør alt no planleggje revisjonsfasen av arbeidet. Tabell 1 gjev ytre karakteristikkar av dei fem ordbøkene.

Presentasjonen viser for det første at ordboksprosjekta har ganske ulik status. SAOB og NO er framleis under redigering, og dei har 2017 og 2014 som fullføringsår for denne primærredigeringa. Men dei er svært ulike med omsyn til ambisjonsnivå og omfang. ODS er SAOB sin næraste ordboksslektning i Skandinavia. Med supplementsbanda fyller ODS 33 band. Primærredigeringa av ODS blei fullført i 1956, men då var revisjonsarbeidet allereie i gang, og no hausten 2005 blir siste hand lagd på det femte og siste supplementsbandet. Samstundes har ein starta arbeidet med å lage ein elektronisk versjon av ordboka der supplementa skal vere fletta inn i hovudteksten. Den primære redigeringa av NRO var fullført alt i 1957, medan revisjonen var ferdig i 1995. Ein har no sett i gang arbeidet med å digitalisere verket, for deretter å førebu ny revisjon i den elektroniske versjonen. DDO, som er den nyaste av desse ordbøkene, er fullført i seks band no i 2005, og her er ein i gang med

å lage ein elektronisk versjon under overbygningen *ordnet.dk*, som tek mål av seg å integrere DDO, *Korpus 2000* og ODS i ein felles nettressurs (jf. Trap-Jensen 2004). Sidan SAOB alt er tilgjengeleg over internett, og sidan NO frå og med 2004 blir redigert som ein søkbar database (jf. Bakken & Ore 2005), kan ein generalisere og seie at alle desse store dokumentasjonsordbøkene er på veg mot det elektroniske mediet, men står på litt ulike steg i denne prosessen.

	SAOB	NRO	NO	ODS	DDO
Status	under red.	digit./revisjon	under red.	revisjon/digit.	digit.
Produksjonsperiode 1	1893-2017	1930-1957	1950-2014	1918-1956	1991-2005
Produksjonsperiode 2	?	1957-1995	?	1950-2005	
Materialgrunnlag	8,5 mill. setlar	1 mill. + 350.000 setlar	3,2 mill. setlar, korpus 22 mill.	2,5 mill. + 1,5 mill. setlar	korpus 40 mill.
Lemma	540.000	200.000	300.000	250.000	100.000
Band	34 (39)	4 + 2	5 (12)	28 + 5	6
Opplag	3600 >1000	?	2000 >1000	5000 + opptrykk	6400

Tabell 1. *Fakta om ordbøkene* (Kursiv viser til prognosetal)

Mot denne bakgrunnen er det interessant å kommentere opplagstala til papirutgåvene av dei ulike ordbøkene. Bengt Sigurd (1986) dokumenterer korleis SAOB blei lansert under stor entusiasme i 1892, og talet på abonnentar steig raskt til det nådde eit høgdepunkt i 1894 med 3600 (Sigurd 1986:146f.). Men ettersom utgjevingstakta dabba av, fall abonnentane frå, og opplaget er i dag på 1000 eksemplar. NO kom i byrjinga ut i eit opplag på 2000, men er også no nede i 1000. Det samla opplaget av NRO har det ikkje vore mogleg å finne fram til, det opprinnelege opplaget er ikkje kommentert i førsteutgåva og forlaget har ikkje lenger dei relevante arkiva, dessutan er ordboka seinare utgjeven i fotografiske opptrykk både i 1983 og 1991, og i opptrykk av einskildband innimellom og etter det. Det må til saman vere snakk om eit relativt stort samla opplag (10.000-20.000?). ODS sitt førsteopplag var imponerende nok 5000, og ordboka har dessutan blitt gjeve ut tre gonger seinare i uendra opptrykk (1966-70, 1975-77, 1981-83). DDO blir gjeven ut i eit opplag på 6400.

Ein ser klart ein samanheng mellom fullføring og opplagsstorleik. Det er ei trusprøve å abonnere på evigheitsprosjekt som SAOB og NO, medan salet av NRO og ODS, viser at fullførde ordboksverk er salbare, jamvel om det er snakk om ei

kostbar investering i 28 ordboksband slik tilfellet var med ODS. (Salet av ODS blei truleg hjelpt av det populære danske spørjeprogrammet *Fup eller Fakta* som blei sendt på TV i fleire rundar på 1960- og 1970-talet. Programmet brukte ODS som "fasit".) Vilkåra for sal av store kostbare papirordbøker er likevel i endring, og det store spørsmålet er å vite korleis dei fullførde utgåvene av NO og SAOB vil bli brukte og selde når det finst elektroniske parallellutgåver. Når ein no har valt å gje ut DDO i eit opplag på 6400, blir det spanande å sjå om dette er eit rett estimat i høve til marknaden, og korleis den framtidige elektroniske utgåva eventuelt vil påverke salet.

3. Institusjonell tilknytning

Dei fem ordbøkene har ulik institusjonell tilknytning. NO er organisert som eit forskingsprosjekt under Universitetet i Oslo, og er slik det einaste ordboksprosjektet med universitetstilknytning. Dei fire andre ordbøkene har det til felles at dei er organiserte og styrde av vitskapsakademi, men desse akademia er rett nok svært forskjellige. Sjå tabell 2.

SAOB	Svenska akademien	Alm. akad. 18 medl.
NRO	Det norske Akademi for Sprog og Litteratur	Språkpol. alm. akad. 10-21 medl.
NO	Universitetet i Oslo	Universitet
ODS	Det Danske Sprog- og Litteraturselskab	Humvit. akad. 75 medl.
DDO	Det Danske Sprog- og Litteraturselskab	do.

Tabell 2. *Institusjonell tilknytning*

Svenska akademien har som kjent 18 medlemmer, og desse medlemene er for tida anten forfattarar eller professorar frå humaniora eller samfunnsfag. To av dei atten er språkvitskapsfolk. Redaksjonen av SAOB lever eit ganske uavhengig liv frå akademiet, men rapporterer jamnleg om framdrift og økonomi. Det Danske Sprog- og Litteraturselskab står som utgjevar av dei to danske ordbøkene. Dette er på den eine sida eit smalare samansett akademi, jf. at medlemene berre er vitskapsfolk frå språkvitskap, litteraturvitskap eller historie, på den andre sida har det breiare representasjon med så mange som 75 medlemmer. Den einskilde redaksjonen har eit fagleg tilsyn oppnemnt av DSL, og i tillegg eit pengetilsyn som dei rapporterer til om pengebruk, tidsbruk og framdrift. Det Norske Akademi for Sprog og Litteratur er til samanlikning like breitt samansett som Svenska Akademien, dvs. at det i tillegg

til språk- og litteraturforskarar også har medlemmer frå ånds- og kulturlivet meir allment. Det er på den andre sida ganske fåtalig, jf. at det til eikvar tid har mellom 10 og 21 medlemmer. Men den viktigaste forskjellen mellom det norske akademiet og dei tilsvarende nordiske, er at dette akademiet har eit eksplisitt språkideologisk siktemål, dvs. at det har teke side i den norske språkideologiske debatten, og er slik eit interesseorgan for riksmålsrørsla. Arbeidet med supplementsbanda gjorde redaktøren, Harald Noreng, for ein stor del aleine, utan at Akademiet la seg mykje opp i det, korkje i innhald eller ytre rammer. Felles for dei tre academia er at dei er lukka selskap som regenererer seg sjølve.

NO har ei litt anna historie å vise til. Sjølv om initiativet til ordboka kom frå det store nynorskforlaget, Det Norske Samlaget, fekk prosjektet tidleg statsstøtte, anten rett over statsbudsjettet eller via forskingsrådet. I 1972 blei prosjektet lagt inn under Universitetet i Oslo, og i perioden fram til 2002 låg prosjektet inne i den vanlege instituttstrukturen. I 2002 blei redaksjonen omorganisert, og redaksjonen har no status som eit sjølvstendig prosjekt på sida av instituttstrukturen, men stadig med Universitetet i Oslo som overbygnad.

4. Utfordringar

Hovudfokuset her skal vere nokre av dei utfordringane dei store ordboksverka står ovafor, og eg har her valt å løfte fram fire utfordringar som utgangspunkt for samanlikning:

1. Finansiering
2. Legitimering
3. Relevans
4. Resultatkrav

4.1. Finansiering

Å sikre finansieringa for desse svært store, svært langvarige prosjekta er verkeleg ei utfordring, ikkje berre for ordboksredaksjonane sjølve, men for samfunnet som heilskap. Tabell 3 gjev eit lite oversyn over finansieringa av dei fem ordbøkene i den fasen vi no er inne i.

SAOB har sidan 1833 blitt finansiert ved å få utbyttet av den statlege publikasjonen *Post- och Inrikes Tidningar* (tilsvarende det norske *Lysingsbladet*). Det er altså snakk om ei fullfinansiering frå staten si side. Denne ordninga er no under reforhandling i samband med at *Post- og Inrikes Tidningar* skal bli ei elektronisk netteneste. Men det er grunn til å tru at Bolagsverket på ein eller annan måte vil kompensere for dei tapte inntektene og vidareføre SAOB som statsfinansiert prosjekt. Både ODS og DDO i Danmark er fullfinansierte, men ikkje einsidig frå staten. Carlsbergfondet har bidrege på like fot med Kulturministeriet i finansieringa

av dei to prosjekta. "Privat" finansiering kan kanskje ikkje Carlsbergfondet sine bidrag kallast; poenget her er at midlane ikkje kjem over statsbudsjettet. NO er til samanlikning fullfinansiert for statlege midlar: Noregs forskingsråd kostar trykkinga, Universitetet i Oslo dekkjer direkte og indirekte driftsutgifter og delar av lønsbudsjettet, medan størstedelen av lønsbudsjettet blir løyvd direkte over statsbudsjettet. NRO har dårlegare vilkår. Pr. i dag får prosjektet noko støtte direkte over statsbudsjettet, men ikkje slik at det er fullfinansiert i den noverande fasen. I tillegg støt NRO seg på fondsmidlar frå mindre fond som Fritt Ord eller Nansenfondet, og ordboka har tradisjonelt fått bidrag frå privatpersonar, dvs. privat finansiering i eigenleg meining. Forskingsrådet har også vore bidragsytar til NRO, og indirekte har Universiteta i Oslo og Bergen vore med på finansieringa av ordboka, sidan setelarkiva som NRO byggjer på, for ein del har vore bygde opp av tilsette ved universiteta. Men det er viktig å understreke at NRO pr. i dag ikkje er sikra full finansiering for å gjennomføre digitalisering eller revisjon.

SAOB	Inntektene frå <i>Post- och Inrikes Tidningar</i> . Reforhandling: Kompensasjon frå Bolagsverket?	Statleg. Fullfinansiert.
NRO	Kulturdepartementet, private fond, forlag, donasjonar	Priv./statleg. Ikkje fullfinansiert.
NO	Kulturdepartementet, Universitetet i Oslo, Noregs forskingsråd	Statleg. Fullfinansiert.
ODS	Kulturministeriet, Carlsbergfondet	Priv./statleg. Fullfinansiert.
DDO	Kulturministeriet, Carlsbergfondet	Priv./statleg. Fullfinansiert.

Tabell 3. *Finansiering*

Ordboksverka det her er snakk om, er ekstremt kostbare prosjekt, som ofte går over mange år, og som heilt sprenger dei økonomiske rammene ein vanlegvis set for forskings- eller bokprosjekt. Difor er ordboksverka heilt avhengige av statleg finansiering for å bli gjennomførde. Eit fellestrekk ved ODS, DDO og NO i den noverande fasen er at dei kviler på ei avtale om kostnadsdeling. Lis Jacobsen (1956) har gjort greie for bakgrunnen for at ODS fekk fullfinansiering og eit lønsnivå for redaktørane på line med andre statstenestemenn. Ho seier her rett ut at "et forhåndstilsagn om bevilling fra fondet [Carlsbergfondet] var den løftestang, der gjorde det mulig at opnå statsbevillingen" (Jacobsen 1956:407). Same modell med kostnadsdeling mellom Carlsbergfondet og Kulturministeriet blei lagd til grunn for DDO. Diverre har ikkje svenskane eller vi i Noreg noko som liknar Carlsbergfondet, men modellen om kostnadsdeling blei likevel overteken då ein etablerte NO

som nytt prosjekt med ny finansiering. Rektor ved Universitetet i Oslo møtte kulturstatsråden, og Universitetet i Oslo forplikta seg til å dekkje ein del av utgiftene. Vilkåret var at Kulturdepartementet tok seg av auken i lønsbudsjettet. Dette har vist seg som ei gunstig ordning, for den verkar gjensidig forpliktande; det er vanskeleg for den eine parten å bryte sin del av avtala. Prosjektet har då også ved fleire høve eksplisitt appellert til partane ved å vise til den gjensidige bindinga.

4.2. Legitimering

Når ein er avhengig av statsløyvingar, slik som i praksis alle dei fem ordboksprosjekta er, blir spørsmålet om politisk legitimitet heilt avgjerande. Denne legitimiteten er ikkje ein permanent storleik, for i så langvarige prosjekt det her er snakk om, vil det truleg bli naudsynt å legitimere prosjekta fleire gonger og på forskjellige måtar, ettersom den politiske og kulturelle konteksten endrar seg.

Den norske situasjonen gjev ein god illustrasjon av dette poenget. Både NO og NRO får løyvingane sine rett over statsbudsjettet, dvs. løyvingane må vedtakast av Stortinget kvar haust. Det må argumenterast for løyvingssummane overfor regjeringa og i kulturkomiteen, og sidan det alltid er fleire ønskemål enn ressursar, må politikarane prioritere mellom dei kulturiltaka dei har framfor seg. Korleis skal ein så greie å argumentere for at det å bruke 100 millionar kroner eller meir på ei ordbok som kanskje færre enn 1000 nordmenn er interesserte i å kjøpe, er rett prioritering av offentlege midlar?

Den norske stoda er interessant, fordi den avvik litt frå den danske og svenske, og fordi ordbøkene meir eksplisitt er legitimerede som kulturpolitiske tiltak. NO flyt på sin status som nynorskiltak. Sidan nynorsken er eit statleg ansvar, og sidan målforma er under konstant trugsmål om marginalisering og på sikt utradering, er det politisk korrekt å stimulere og støtte tiltak som kan betre nynorsken sin status i samfunnet. Dette kjem NO til gode. I tillegg har NO eit demokratisk eller distriktspolitisk aspekt, som gjer det kulturpolitisk omsetjeleg. For NO dokumenterer ikkje berre det nynorske skriftmålet, men også det ”norske folkemålet”, dvs. dialektane. I Noreg, der dialektane og distrikta står så sterkt, blir difor NO politisk korrekt også på dette punktet. Her kan heile folket og alle samfunnslag finne sitt eige kvardagsspråk dokumentert. Om ikkje dette er nok, kan vi appellere til museumsaspektet av NO, dei gamle dialektorda i NO er eit aspekt av den folkelege og nynorskklitterære kulturarven som for ein del er på veg til å forsvinne. Mange ser behovet for å dokumentere denne språklege kulturarven før han er tapt for alltid.

Prosjektet NO 2014 skal i tillegg legitimere seg innanfor Universitetet i Oslo, for universitetet må også ville ønskje å vidareføre tilskotet til ordboksarbeidet innanfor sine pressa budsjett. I denne samanhengen har nok NO 2014 som vitskapleg digitaliseringsprosjekt mest appell og legitimeringstyngd. I og med digitaliseringsarbeidet som har gått føre seg dei siste åra, blir nemleg språkarkiva våre no lagde til

rette for anna utnytting enn den leksikografiske bruken. Setelarkivet og korpuset er tilgjengelege for ålmenta via internett, og ordboka er no ein elektronisk database som opnar opp for ei mengd nye utnyttingsmåtar som tener universitetsverksemda.

I møte med dei løyvande myndigheitene kan ein sjølvsagt også vise til kor mykje arbeid som faktisk alt er lagt ned i ordboka, og at det er umogleg å stanse ei ordbok i bokstaven *k*. Men røymsler frå grannelanda viser at dette kontinuitetsargumentet ikkje har avgjerande vekt. (Jf. OSD som no i praksis er terminert, og ONP i Danmark som etter band 3 (*de-em*) ikkje vil kome ut som papirbok.) Det er verdt å merkje seg at funksjonelle argument, slik som samfunnsnytte eller praktiske konsekvensar for den skrivande og lesande ålmenta i Noreg, er underordna i argumentasjonen for NO.

NRO utgjer eit interessant samanlikningspunkt. Ordboka skal jo legitimerast i den same kulturpolitiske konteksten som NO, men det har vist seg å vere vanskelegare å få gehør for dette prosjektet (jf. Guttu 1996:57f.). Dette heng truleg for ein del saman med statusen til riksmålet. I motsetnad til nynorsken har riksmålet ikkje offisiell status, og er slik ikkje ein konkurrent til nynorsken, men til bokmålet. Rett nok er dei reelle skilnadene mellom det offisielle bokmålet og riksmålet i ferd med å viskast ut, men i denne samanhengen er det underordna. Medan nynorsksaka har gjeve NO legitimitet og eit preg av politisk korrektheit, har riksmålsassosiasjonen truleg vore ei belastning for NRO i møte med dei løyvande myndigheitene. I praksis er då også dette tona ned i den eksplisitte presentasjonen av NRO, jf. Lars Bucher Johannessen og Lars Roar Langslet i forordet til tilleggsbanda som kom ut i 1995 (NRO 5:v). Derimot har ein appellert til forskingsrelevansen, dvs. ved å vise til faglege omsyn slik som behovet for å dokumentere utviklingane i det skriftlege notidsspråket. Denne argumentasjonsvegen er heller ikkje problemfri, for det hadde truleg vore lettare å argumentere for NRO som forskingsprosjekt om ordboka hadde hatt offisiell universitetstilknytning og tettare samheng med den språklege dokumentasjonsverksemda ved universiteta. NRO hadde truleg vore tent med å vere betre organisatorisk forankra som forskingsprosjekt, for slik det er, manglar NRO dei formelle forskingsgevanta det kunne ha vore opportunt å ikle seg.

Det sterkaste kortet NRO har å bruke andsynes dei løyvande myndigheitene, er bruksaspektet. Dette er den einaste dokumentasjonsordboka på bokmålssida (orsak termen!), NRO er fullført, og det har vist seg å vere behov for verket. Bokmålet er jo stetta av ei mykje større brukargruppe enn nynorsken, og dermed er det ein marknad for NRO. Ironisk nok kan kanskje dette også vere problematisk frå eit statleg finansieringssynspunkt, for forlagsinteressa for NRO er større enn for NO, og med god grunn.

Situasjonen i Danmark og Sverige er eg naturleg nok dårlegare kjend med, men i Danmark ser i alle høve ordboksprosjekta ut til å vere meir fagleg og funksjonelt motiverte. ODS blei lansert som eit nasjonalt dokumentasjonsprosjekt, avløyseren

til Molbechs då utdaterte ordbok. I utgangspunktet skulle ordboka dokumentere samtidspråket avgrensa frå 1700 til redigeringstidspunktet, eit tidspunkt som flytta seg ettersom verket skreid fram (jf. Vikør 1999). Men i og med at ein sette kronologisk strek for ekserperinga ved sluttåret for primærutgjevinga (1955), har ordboka gjennom heile supplementsperioden blitt redigert som ei *historisk* dokumentasjonsordbok.

Ein tilskuv til at DDO blei sett i gang, var at Danmark gjekk inn i den europeiske unionen. På 1980-talet, i kjølvatnet av EU-avstemmingane, var det lett å få gehør for tiltak som styrkte det sær danske, og det blei advart mot at det danske språket kanskje ville bli truga innanfor rammene av den nye europeiske fellesskapen. DDO er altså delvis legitimert av behovet for å styrke det danske i møte med det framande. Fagleg sett var det lett å argumentere for at ein trong å dokumentere samtidspråket, ODS inkluderer jo berre språk og språkleg utvikling frå før 1955. DDO blei difor nasjonalt og funksjonelt legitimert i møte med politikarane. Det galdt å dokumentere, men også ideelt sett å styrke det moderne danske språket i møte med EU. DDO måtte difor greie å sameine omsynet til eit vidt spekter av brukarar, for den alminnelege dansken måtte ikkje få eit så komplekst ordboksformat at han ikkje greidde å bruke det, samstundes med at ordboka også måtte kunne vere eit forskingsverktøy for lingvistar.

Når det gjeld DDO, og seinare *Korpus 2000* og *ordnet.dk*, ser det ut som DSL på ein vellukka måte fungerer som fagleg garantist opp mot finansieringsinstansane. DSL-prosjekta framstår slik som i utgangspunktet fagleg vellegitimerte, slik at ein kan bruke kreftene sine på å argumentere kultur- eller forskingspolitisk for desse projekta i møte med finansieringskjeldene. I Noreg manglar vi ein slik samlande fagleg koordinator av ordboksprojekta, noko eg trur er ei ulempe. I alle høve kan det oppstå situasjonar der til dels konkurrerande prosjekt kan kome til å bruke krefter på å argumentere fagleg opp mot finansieringskjeldene, i staden for kultur- eller forskingspolitisk.

SAOB si opprinnelege legitimering var som eit språkleg og kulturelt dannelsesprosjekt, jf. Svenska Akademiens valspråk *Snille och smak*, og Gustaf 3. sitt ønske om å dyrke fram "Svenska Språkets renhet, styrka och höghet" (Loman 1986:5ff.). Den svært tunge tradisjonen som no står opp under SAOB, og det så langt sikre finansieringsgrunnlaget, har nok gjort det mindre presserande å legitimere prosjektet i samfunnet. Det er blitt eit nasjonalt klenodium, og slik er det heva over kritiske eller kjetterske spørsmål. Det er kanskje også ei bakside ved denne situasjonen, for utan legitimeringsbehov mistar ein noko av kontakten med det samfunnet ein skal kommunisere med. Det er slik interessant at mange av dei entusiastiske abonnentane til dei første hefta av SAOB var lærarar og skulefolk, og forlaget lanserte ordboka som ei bok "för hela folket" (Loman 1986:14). Denne situasjonen er endra, og Bengt Sigurd (1986:176) skriv: "För allmänheten – även den s.k. bildade – är nog SAOB

okänd, och i ökande grad." I tillegg ser ein at den tunge tradisjonen også lyfter SAOB over fagleg kritikk. Lundbladh (2003) dokumenterer korleis utgjevingane av SAOB-band og -hefte i svært liten grad har vore emne for faglege meldingar.

Oppsummeringsvis kan ein konkludere med at dei store nordiske ordbøkene er ulikt legitimerede. Ein ser at legitimeringsbehovet er ulikt til ulike tider, og at nasjonsbyggings- eller dannelsingsprosjekt kanskje ikkje er like "salbare" no som for 100 år sidan. Det kan slik sett vere ei utfordring å endre argumentasjonen undervegs.

4.3. Relevans

Så langt har eg drøfta korleis offentleg finansiering kviler på kva slags politisk legitimering ordboksprosjekta får. Men ein får tru at politisk aksept også heng saman med kva slags sosial eller kulturell relevans ordbøkene *faktisk* har. Ein kan jo påstå at dette er eit prosjekt som bringer fram uunnverleg forskning, eller at dette er eit prosjekt som styrkjer den vanlege mannen sitt språklege medvit, men spørsmålet er om ordbøkene faktisk er relevante for dei brukargruppene ein ønskjer å nå, og i høve til den argumentasjonen som blir brukt for å legitimere ordbøkene politisk.

Vi har vel alle saman eit ideal om at ordbøkene vi produserer skal kjennast relevante for brukarane våre. På den andre sida er det eit empirisk faktum at vi har få brukarar. Både den ålmenne kulturelle konteksten og dei forskingsparadigma som var gjeldande då ordbøkene opprinneleg blei initierte, er endra, og dermed må ein spørje seg om ordbøkene våre faktisk og framleis er relevante.

Eit hovudintrykk er at SAOB, ODS, NO og for ein del NRO har låg samfunnsrelevans. DDO har enno ikkje fått tid på seg til å stadfeste relevansen sin. SAOB og NO er lite kjende i samfunnet, og dei sel i små opplag. ODS og NRO som er fullførde, er meir kjende. Presseoppslaga om NO får gjerne ei museal eller kuriøs, eller endåtil humoristisk vinkling. Dette gjeld då særleg framdriftstakta. Og det er klart at det formatet vi arbeider innanfor ser avvikande ut i ei "tabloid" verd. I forsøka på å svare på utfordrande spørsmål frå pressa, kan vi dessutan ikkje lenger bruke ord som *danning*, *nasjonal* eller *kulturarv*, for dei er ikkje utan vidare honnørord i dag. Utfordringa blir difor å gjere ordbøkene meir samfunnsrelevante, og eg vil nemne nokre tiltak på vegen mot eit slikt mål.

Det er viktig å bruke media meir offensivt. Ofte er det slik at vi blir kontakta av pressa for at dei skal få stadfest sine fordømmar om ordboksverksemd av oss. Her er det truleg mykje å hente ved å ta initiativ på eigne premissar. Lis Jacobsen ser ut til slik å ha brukt media svært medvite til beste for ODS i dei åra ho styrde ordboka. Stoffet vi arbeider med eignar seg desutan til mange andre og meir tabloide format enn som ordboksartiklar. Eg tenkjer her på Dagens Ord-spalter, ein redaktør hos oss har fast program på NRK kvar veke der han tek imot spørsmål frå lyttarane, og andre TV- og radiokonsept kunne også ha ordfokus.

Det norske TV-programmet *Typisk norsk* viser dessutan at språk og ord er noko

folk er svært interesserte i. I Noreg er det stor aktivitet rundt om i bygdene når det gjeld å registrere gamle dialektord. Etymologiske ordbøker sel godt både i Noreg og Sverige. Skandinavar er opptekne av litteratur og forfattarskapar. Og svært mange løyser kryssord eller spelar scrabble. Utfordringa vår er å knyte an til denne interessa på ein kreativ måte. Målet må vere å greie å kanalisere språkinteressa hos folk over på ordbøkene via dei meir tabloide formidlingsformata.

I denne samanhengen er digitaliseringsprosessane som alle dei nordiske ordboksprosjekta på ein eller annan måte står midt oppe i, avgjerande viktige. Ordboksverka våre er store, og dyre og plasskrevjande. Elektroniske versjonar av ordbøkene er ingen av delane, og dei inviterer dessutan til meir mangesidig bruk av ordbøkene (jf. Bakken & Ore 2005). I freistinga på å nå fleire brukarar i framtida vil det elektroniske mediet truleg bli heilt avgjerande.

Så langt har eg litt upresist snakka om samfunnsrelevansen. Eg vil avslutningsvis avgrense omgrepet litt, og sjå nærare på forskingsrelevansen av ordbøkene våre. Dei er trass alt vitskaplege dokumentasjonsordbøker, og forskarsamfunnet er truleg den primære brukargruppa for fleire av dei. Utgangspunktet for meg er ein påstand om at koplinga mellom den generelle språkvitskapen og den vitskaplege leksikografien blei gradvis svekt frå 1950 og i tiåra framover. Grunnane kan vere mange, men éin grunn er at nye fokus kom til innanfor språkvitskapen, og desse fokusa prioriterte syntaks og teori heller enn empiri og leksikon. Ein annan grunn kan vere at det blei færre formelle koplingar mellom ordboksverka og forskarsamfunna, det var ikkje lenger så vanleg å dele tida si mellom universitet og ordboksredaksjon, jf. det tette personalhopehavet mellom Københavns universitet og den opprinnelege redaksjonen i ODS. Dessutan konsoliderte leksikografien seg som eige fagfelt, slik at det blei skipa til eigne leksikografitidsskrift, eigne undervisningstilbod, eigne konferansar og avhandlingar med reint leksikografisk fokus. Under éin synsstad er dette sjølvsagt eit gode, men det er grunn til å tru at det ikkje har fremja kontaktane mellom den genrelle språkvitskapen og leksikografien. Sjå for øvrig Christensen 2002 for ei interessant samanlikning av korleis disiplinen semantikk blir praktisert innanfor SAOB og innanfor universitetet.

Mitt poeng med å peike på desse faktorane som tidlegare har skapt avstand og slik sett undergrave relevansen til ordbøkene for universitetssektoren, er å vise til at desse vilkåra på viktige punkt er endra, og at ein difor no har eit godt utgangspunkt for å knyte nye band mellom språkvitskapen og leksikografien.

For det første rår no nye språkvitskaplege forskingsparadigme med ny interesse for faktisk språkbruk og breie empiriske undersøkingar. Her er korpuslingvistikken truleg det beste dømet. Dessutan er det større interesse for leksikonet og for morfologi enn 1970- og 80-talet var prega av, og det innanfor ulike teoretiske overbygningar. Grensene mellom datalingvistikken og den databaserte leksikografien er i ferd med å viskast ut, noko som har vist seg å opne for nye spanande perspektiv.

Innanfor leksikografien har dessutan dei nye databaserte metodane gjeve oss eit heilt anna utgangspunkt for å møte andre språkforskarar på halvvegen. Viktigast her er det at datateknologien har brote ned grensene mellom ordbøker, databasar og kjelder. Innanfor prosjektet NO er t.d. no setelarkiva våre elektronisk tilgjengelege og søkbare, vi har etablert eit tekstkorpus som på same måte er tilgjengeleg og søkbart, og sjølve ordboka er ein database med direkte koplingar til kjeldegrunnlaget. Innanfor overbygningen *ordnet.dk* vil ein etter kvart finne koordinerte foredla språkressursar som innbyr til å utnytte desse ressurane på nye måtar. Ikkje minst viktig vil det vere at dei nye elektroniske formata opnar opp for systematisk og ikkje berre partikulær bruk av ordbøkene, noko som absolutt bør auke relevansen av dei for andre språkforskarar.

Eg trur altså digitalisering vil vere heilt avgjerande for meir forskingsrelevant bruk av ordbøkene. Men for å nå dette målet, er det i alle fall fire forhold som det trengst å arbeidast vidare med. Vi må arbeide med å bryte ned fordommane som finst mot ordboksgenren i forskarsamfunna, vi må gjere dei nye ordboksprosjekta synlege i forskarsamfunna, og vi må demonstrere for andre forskarar korleis ein ny vitskapleg ordboksbruk kan sjå ut. Målet må vere at auka forskingsrelevans og utnytting i neste omgang kan brukast som legitimeringsargument opp mot forskingssektoren, ikkje berre mot kultursektoren.

4.4. Resultatkrav

Til slutt vil eg stutt kommentere den utfordringa som gjeld resultatkrav. Det er no slik at alle dei fire fullfinansierte ordbøkene ha, eller har hatt resultatkrav stilte til seg. NO fekk eit slikt krav i og med omorganiseringa i 2002, medan SAOB fekk eit sluttår for arbeidet på 1990-talet. Med unnatak av NRO, som altså ikkje er fullfinansiert, er resultatkravet formulert av finansieringskjeldene, som også fungerer som rapporteringsinstans. Når resultatkrav no er formulerte, er oppdragsaspektet av verksemda tydeleggjort. Det ser elles ut til at det er litt ulikt kor tett oppdragsgjeverane følgjer oss opp.

Det har ofte blitt fokusert på to negative sider ved det å redigere ordbok med stramme tidsfristar. Det kan for det første vere stressande å arbeide med strenge tidsfristar, og for det andre oppstår det ofte ein konflikt mellom progresjon og kvalitet. Begge problema vil kjennast mest akutte når tidsfristane blir urimeleg strenge. NO 2014 har likevel insistert på å få rapportere om framdrifta direkte til Kulturdepartementet og ikkje berre til vårt eige styre. Vi ønskjer å fokusere på det faktumet at vi har fått eit oppdrag frå samfunnet, og slik brukar vi sluttåret som utgangspunkt for å argumentere eksternt for finansiering. Internt er sluttåret eit absolutt orienteringspunkt når vi må gjere val som gjeld framdrift, dimensjonering og avgrensing.

Innanfor rimelege grenser trur eg faktisk det er bra å arbeide med absolutte

tidsfristar. For det første må all vitskap avgrensast; det gjeld å svare på oppgåva ein har sett seg på adekvat måte. Dessutan er det ofte slik at når eit prosjekt manglar tidsavgrensing, vil gjerne dei kvalitative og kvantitative normene endre seg undervegs, jf. den naturlege utviklinga av SAOB, ODS og NO opp gjennom åra. Dette viser også at kvalitet er ein variabel storleik som ikkje kan vurderast uavhengig av rammene for oppgåva. Eg trur dessutan at det å greie å innfri eit samfunnsoppdrag er heilt avgjerande for å gjere arbeidet vårt relevant, noko ikkje minst salet av dei fullførde ordbøkene NRO og ODS demonstrerer. Vi må demonstrere ein vilje til å møte samfunnet på dei premissane andre legg for oss. Eg må likevel presisere at det gjeld å finne fram til ei tidsavgrensing som står i rett forhold til omfanget av arbeidet.

5. Konklusjon

Denne samanliknande drøftinga av utfordringar som er sams for dei fem nordiske ordboksverka eg her har sett nærare på, hadde som primært mål å gje meg innsikter som kan kome NO til gode. Eg vil difor konkludere med dei personlege læringsresultata eg sjølv sit igjen med. Mi von er at dette også kan gagne andre.

På vegner av alle ordboksprosjekta trur eg for det første at vi må legitimere prosjekta på ein meir medviten måte enn tradisjonelt har vore gjort. For det andre er det viktig å betre ordbøkene sin samfunnsrelevans og forskingsrelevans. For det tredje må vi vere meir utoverretta om vi ønskjer å synleggjere og halde på det samfunnsoppdraget vi faktisk har fått, og for det fjerde må vi ta resultatkrava på alvor. For NO sin del vil eg leggje til at vi har mykje å gå på når det gjeld fagleg legitimering, medan vi ligg betre an når det gjeld kulturpolitisk legitimering. Fagleg legitimering må i siste instans forankrast i fagleg kvalitet, og vi bør difor halde fram med å styrkje den faglege kvaliteten internt i redaksjonen. Og til sist vil eg halde fram at digitaliseringsarbeidet er avgjerande i høve til mange av dei utfordringane vi no står overfor. Digitaliseringsarbeid synleggjer, skaffar nye brukarar, lyfter det interne medvitsnivået, inviterer til ny og meir systematisk bruk av ordbøkene, gjer ordbøkene meir vitskaplege og fremjar framdrift. Det er difor gledeleg at alle dei fem ordboksverka eg her har sett nærare på, no vidareutviklar ordbøkene sine langs dette sporet.

Litteratur

Ordbøker:

DDO 2003-2005 = E. Hjorth/K. Kristensen (red.): *Den Danske Ordbog*. 6 band. Udgivet af

Det Danske Sprog- og Litteraturselskab. København: Gyldendal.

NO 1966- = *Norsk Ordbok*. 5 av 12 band. Oslo: Det Norske Samlaget.

NRO 1937-1957, 1995 = T. Knudsen/A. Sommerfeldt/H. Noreng (red.): *Norsk Riksmålsord-*

- bok*. 6 band. Utgitt av Det norske Akademi for Sprog og Litteratur. Oslo: Aschehoug/Gyldendal/Kunnskapsforlaget.
- Nusvensk ordbok* 1917-1972. 10 band. Stockholm: Wahlström & Widstrand.
- ODS 1918-1956 = *Ordbog over det danske Sprog*. 28 band. Udgivet af Det Danske Sprog- og Litteraturselskab. København: Gyldendal.
- ODS-S 1992-2005 = *Ordbog over det danske Sprog. Supplement*. 5 band. Udgivet af Det Danske Sprog- og Litteraturselskab. København: Gyldendal.
- ONP 1989-2004 = *Ordbog over det norrøne prosasprog. Register*, 3 bind. Udgivet af Den arnamagnæanske kommission. København.
- OSD 1991 = *Ordbok över Sveriges Dialekter*. 1 band. Uppsala: Arkivet för Ordbok över Sveriges dialekter.
- SAOB 1898- = *Ordbok över Svenska Språket*. 34 av 39 band. Utgiven av Svenska Akademien. Lund: Gleerup.

Annan litteratur:

- Allén, Sture, Loman, Bengt & Sigurd, Bengt 1986: *Svenska Akademien och svenska språket. Tre studier. Svenska Akademien 200 år*. (Svenska Akademiens handlingar från år 1986. Fjärde delen 1986.) Stockholm: Norstedts.
- Bakken, Kristin & Ore, Christian-Emil 2005: Norsk Ordbok – også ei elektronisk ordbok? I: *LexicoNordica* 12, 7-18.
- Christensen, Lisa 2002: Universitetssemantik och ordbokssemantik. I: Mattisson, Anki mfl. (red.): *Alla ord är lika roliga. Festskrift til Lars Svensson 28 februari 2002*. Svenska Akademien. Distr. Nordstedts: Stockholm, 24-39.
- Guttu, Tor 1996: Norsk Riksmålsordbok V-VI. I: *Mål og Makt* 4, 53-58.
- Jacobsen, Lis 1956: Da den danske ordbog blev en institution. I: Aakjær, Svend, Hald, Kr. & Hansen, Aage (red.): *Festskrift til Peter Skautrup 21. januar 1956*. Århus: Universitetsforlaget, 403-410.
- Loman, Bengt 1986: En Inrättning, ägnad endast til Språkets förbättring. I: Allén, Loman & Sigurd 1986, 1-142.
- Lundbladh, Carl-Erik 2003: Kritiken av SAOB. I: *LexicoNordica* 10, 99-118.
- Sigurd, Bengt 1986: Ordboken, ordlistan och några andra av Svenska Akademiens språkliga insatser under 1900-talet. I: Allén, Loman & Sigurd 1986, 145-232.
- Trap-Jensen, Lars 2004: Et net af ord – ordnet.dk. I: *Mål & Mæle* 4, 24-30.
- Vikør, Lars Sigurdsson 1999: Fleirgenerasjonsordbøker og tida. I: *Nordiska studier i lexicografi 4. Rapport från Konferensen om lexicografi i Norden, Esbo 21-24 maj 1997*. Helsingfors: Nordiska föreningen för lexicografi, Helsingfors, 395-405.

Norsk Ordbok på nett: utfordringar og løysingar

Sturla Berg-Olsen og Åse Wetås

Norsk Ordbok 2014, Universitetet i Oslo

Summary

Norsk Ordbok (abbreviated NO) is a scientific dictionary covering Norwegian in its written Nynorsk form and all spoken dialects. The dictionary is under production and will be finished in 12 volumes in 2014. The work on NO began in 1930, and the first volume was published in 1966. From 2002 the project organisation NO 2014 has been working towards the completion of the dictionary, all the time on the premise that the final product will be a printed dictionary. Over the last few years a growing number of dictionaries have been published in an electronic format on the World Wide Web – either as a supplement to or an alternative to a printed edition. We see a development where more and more lexicographical products are exclusively published in an electronic form. This format renders possible a more user-friendly search functionality and more flexible ways of displaying dictionary entries. In addition, the format makes it fairly easy to link lexicographical resources with other existing, relevant electronic resources (cartographic data, source texts, encyclopaedic resources etc.). We at NO 2014 would also like to publish our dictionary electronically, but we have found ourselves faced with certain lexicographical challenges that have to be looked into in more detail. In this article we discuss some of the problems encountered in the work on the online version of NO, with particular emphasis on search mechanisms.

Nøkkelord

e-leksikografi, eittspråklege ordbøker, søkjemekanismar, dialektordbøker

1. Innleiing

Norsk Ordbok (NO) er ei vitenskapleg dokumentasjonsordbok som dekkjer det nynorske skriftspråket og det norske folkemålet. Ordboka er under produksjon og skal vere ferdig i 12 band til 200-årsjubileet for den norske Grunnlova i 2014. Arbeidet med NO blei starta i 1930,

og det første bandet av ordboksverket kom ut i 1966. Frå 2002 har prosjektorganisasjonen Norsk Ordbok 2014 (NO 2014) arbeidd med ferdigstillinga av ordboksverket, og det har heile tida vore føresetnaden at verket skulle gjevast ut på papir. Dei seinaste åra har fleire og fleire ordbøker blitt publiserte elektronisk på nettet, anten som supplement til ei prenta utgåve eller som eit alternativ til prenting. Utviklinga går i retning av at stadig fleire leksikografiske arbeid utelukkande blir publiserte elektronisk. Dette opnar for meir brukarvenleg søkjefunksjonalitet og fleksibel artikkelvisning, og det elektroniske publiseringsformatet legg også til rette for ei samankopling av dei leksikografiske ressursane med andre tilgjengelege og relevante elektroniske ressursar (kartdata, grunnlagskjelder, ensyklopediske ressursar osv.). I NO 2014-organisasjonen ønskjer vi òg å publisere ordboksverket på nett, men dette publiseringsmediet utløyser ein del leksikografiske utfordringar som må diskuterast meir i detalj. I denne artikkelen kjem vi nærare inn på nokre av problemstillingane vi har drøfta i arbeidet med ei nettutgåve av ordboka, og då spesielt spørsmål som gjeld søkjemekanisar.

2. Bakgrunn

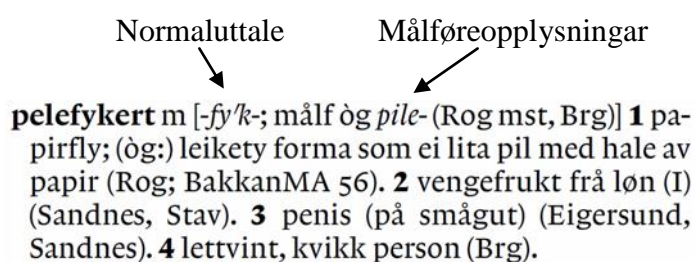
NO dekkjer både det norske folkemålet (dvs. alle dei norske dialektane) og nynorsk skriftspråk frå 1860-talet og fram til i dag. Dette gjer at vi trygt kan ta utgangspunkt i at ordboksverket har eit samansett publikum. Mange potensielle brukarar vil primært vere interesserte i å hente ut informasjon om ord i eigen dialekt, medan andre kan vere på leit etter ulike nemningar for eit gjeve fenomen, eller etter eldre skriftspråklege belegg av eit gjeve ord. Eit mål vi har sett oss i arbeidet med ein nettversjon av ordboka, er at brukarane både skal få dei rette forventningane til kva dei kan finne i NO, og at dei samstundes får god rettleiing til å finne det dei leiter etter.

Artikkelmaterialet i NO bruker 1938-normalen for nynorsk som redaksjonell målform, med nokre mindre modifikasjonar som ikkje er så viktige i denne samanhengen. Dei formene frå 1938-normalen som er valde ut, ligg stort sett ikkje så langt frå ein moderne midtlinenynorsk, men normalen skil seg frå dagens nynorsk på nokre punkt. T.d. har substantivet *jente* oppslagsforma *gjente* i NO.¹ Ordboka tek utgangspunkt i at nynorsk skriftspråk og norske talemål er to sider av ei og same sak, eller siamesiske tvillingar, om ein vil. Det inneber at den redaksjonelle handsaminga av ord som berre er dokumentert nytta dialektalt, også kan få ei

¹ Mellom dei sentrale systemdraga i NO-nynorsken er *a*-infinitiv, supinum på *-i* (*skriv*, *lest*) og adjektivendinga *-ut* (*strip*, *rugl*).

normert nynorsk form. Den nynorske forma slike ord får i NO, følgjer 1938-normalen. Døme på dette er belegga bak substantivet *myseprim*, der mange har form /²møsbrøm/, eller substantivet *matgjerd*, som finst i mange dialektar som /²majær/ eller /²majær/.

For at tilhøvet mellom normert nynorsk oppslagsform og talemålsvariantane som ligg under skal komme til uttrykk, blir målføreformene gjevne att i ein eigen del av artikkelen. Her er det berre avvik ut over det som er heilt regelbunde, som blir gjevne att. Det inneber at vanlege, utbreidde former ofte ikkje blir gjorde greie for i artiklane. Til grunn for det ligg det ein føresetnad om at brukarane kjenner til regelbundne målføredrag (slik som lenisering av *p*, *t*, *k* til *b*, *d*, *g* i Sør-Rogaland og Agder, jf. figur 1).



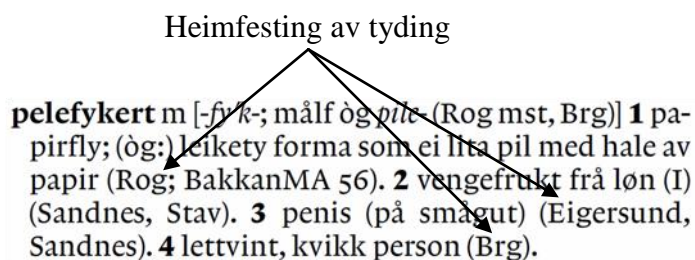
Figur 1. NO-artikkel (i papirversjon) med normaluttale og målføreopplysningar.

I artikkelen **pelefykert** kjem opplysningar om normaluttale av ordet først i den innleiande parentesen med skarpe klammer («hakeparentesen»). Grunnen til at det er ført den typen opplysningar i *pelefykert*-artikkelen, er at trykkplasseringa er avvikande frå den som er den vanlege i norsk. Etter opplysningane om normaluttale følgjer registrert målførevariasjon. Belegga frå Rogaland gjer greie for uttalen /pilə¹fy:gərt/. Dette området har lenisering, men i og med at dét målføredraget blir rekna for å vere sjølvgjevne, blir det ikkje gjort greie for i NOs framstilling.

Derimot representerer ikkje *e > i* i førsteleddet av dette ordet eit regelbunde målføredrag, og målføreopplysninga *pile-* blir difor førd. Målføreformer som ikkje er sjølvgjevne, blir førde inn i hakeparentesen i ordartiklane med geografisk plassering på heradsnivå. Heradssystemet reflekterer kommuneinndelinga i Noreg slik ho var før dei store kommunesamanslåingane på 1960-talet.

Vidare fører NO heimfesting av ordtydingar i tydingsdelen av ordartiklane. I den elektroniske redigeringsapplikasjonen blir heimfestingane førde i eit eige felt knytt til den aktuelle definisjonsteksten. I prenta artiklar kjem dette inn som opplysningar i parentes etter definisjonane (jf. figur 2 nedanfor). Dette er aktuelt i alle dei tilfella der ordtydinga ikkje er

vanleg over heile landet. For ordtydingar som er ålment utbreidde, blir det ikkje ført heimfestingar.



Figur 2. Pilene viser heimfestingar av ordtyding.

NO i prenta versjon er tilgjengeleg for tingarar og gjennom forskings- og folkebibliotek som har verket i boksamlinga si. Når vi nå presenterer ordartiklane i ei nettbasert ordbok, er det rimeleg å vente at gruppa av potensielle brukarar av verket blir mykje større. Fordi ordboka i nettversjon møter ei stor og samansett brukargruppe, er det viktig å gjere artikkelpresentasjonen på nett så brukarvennleg som mogleg. Det aller viktigaste å sikre då, er at brukarane så langt som råd får treff som svarer til det dei faktisk ønskte å søkje på. For ei ordbok som presenterer analysar av både talemålsmateriale og materiale frå skriftspråklege kjelder, er utfordringa endå større enn for dokumentasjonsordbøker som berre dekkjer den eine av desse kjeldetypane.

Det vi ønskjer for nettversjonen av NO, er at brukaren ikkje berre skal kunne søkje på oppslagsord, men også søkje i fritekst eller på geografisk stad ned til heradsnivå. For å sikre mest mogleg relevante treff, ønskjer vi vidare at brukaren også skal kunne søkje seg inn i ferdige ordartiklar gjennom dei koplingane som finst mellom ordartiklar og grunnlagsmateriale. Der er koplingspunktet NOs Metaordbok.

Metaordboka er ein elektronisk indeks der materiale frå ulike materialdatabasar er samla. Dette omfattar mellom anna NOs setelarkiv, Trønderarkivet, grunnmanuskriptet til NO, *Nynorskordboka* (NOB), *Bokmålsordboka* (BOB), Matias Skards *Nynorsk ordbok* (1954) og ei rekkje kjelder som ligg inne i Ordbokshotellet.² Belegg frå desse kjeldene er i Metaordboka samla under normerte oppslagsord og grovlemmatiserte.³

² Ordbokshotellet inneheld lokale ordsamlingar i elektronisk form, saman med ordartiklar frå ymse leksikografiske kjeldeverk, slik som *Nynorsk etymologisk ordbok* (Torp 1919) og *Norsk landbruksordbok* (Rommetveit 1979). For meir informasjon om Ordbokshotellet sjå Tvedt, Lien og Eide (2007).

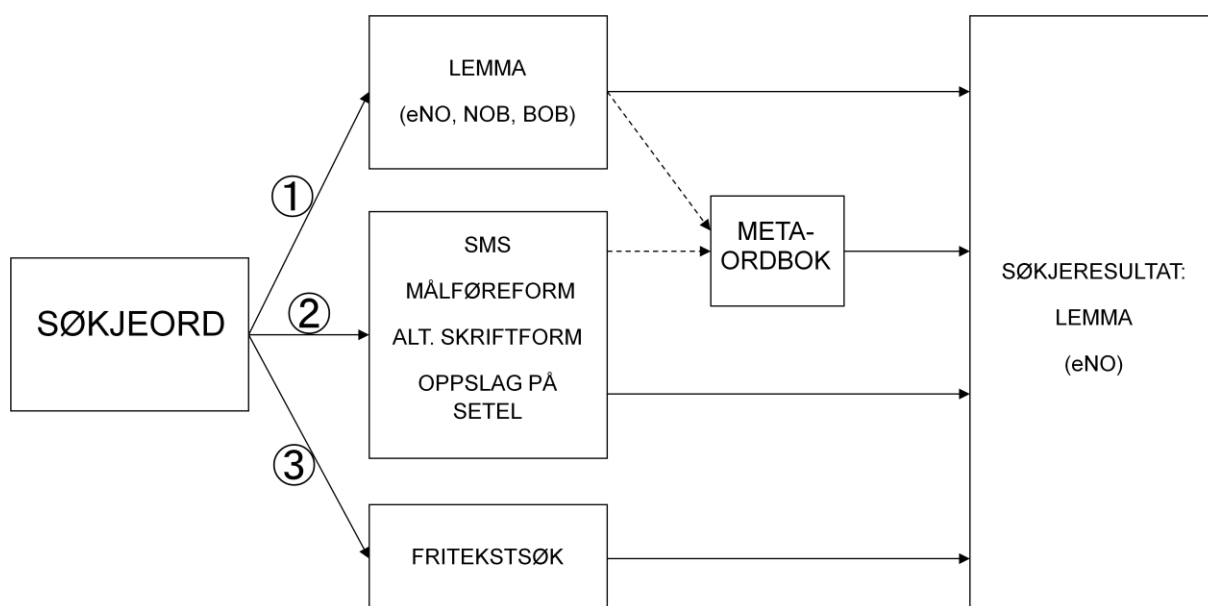
³ Metaordboka skil ikkje mellom homografe lemma som høyrer til same ordklasse, eller same genus for substantiv. Såleis ligg til dømes belegg på homografane **I singel** «småkorna stein, grov grus» og **II singel** «1. singelplate; 2. spel med éin deltakar på kvar side» i same metaordboksartikkel, medan adjektivet **V singel** har eigen artikkel i Metaordboka. For meir om Metaordboka sjå Svoldal (2003) og Ore og Tvedt (2006).

Det er Eining for digital dokumentasjon (EDD) ved Universitetet i Oslo som står bak NOs redigeringsapplikasjon, og EDD står også for utviklingsarbeidet bak nettutgåva av ordboka. Til første versjon av NO på nett har vi valt å prioritere funksjonalitet høgare enn design. Planen er å arbeide vidare med utsjånaden på sidene etter at første versjon av eNO er lagt ut og testa. Etter testing vil vi kunne gjere endringar i måten vi presenterer og legg til rette for søk, for å gjere treffprosenten endå høgare og brukarane endå meir nøgde.

3. Søkjevegar ved enkelt søk i eNO

Dei fleste brukarane av eNO, eller i alle fall dei som har minst kjennskap til ordboka frå før, vil truleg bruke funksjonen «enkelt søk». Denne funksjonen vil ha eit søkjefelt utan valmoglegheiter. Vi har som mål å lage ein søkjemekanisme for enkle søk som gjev flest mogleg brukarar treff på det dei er ute etter, men dette lèt seg ikkje nødvendigvis gjere så lett. Dels kjem det av faktorar som alle ordbøker må ta høgd for – som at brukargruppa er heterogen og at nokre brukarar truleg vil vere ute etter informasjon som nett denne boka ikkje er designa for å gje (t.d. gjev ikkje eNO opplysningar om gjeldande offisiell rettskriving). Men vi støyter også på ein del utfordringar knytte til særdrag ved verket NO, som skissert ovanfor. Det gjeld til dømes diskrepansar på ein del punkt mellom oppslagsformene i NO (som følgjer normalen frå 1938) og dei formene som høyrer til dagens norm. Endå meir krevjande er tilhøvet mellom oppslagsformene i standardisert nynorsk, og målføreformene som ein gjeven artikkel representerer, og som ofte skil seg heller mykje frå den normaliserte forma. Etersom NO – og dermed eNO – også er ei dialektordbok, vil ein tru at mange brukarar vil vere opptekne av å finne ord og uttrykk frå talemålet ulike stader i Noreg. Korleis kan vi utforme eit søkjesystem som fangar opp flest mogleg søk på målføreformar, leier brukaren til rett normalisert lemma og samstundes viser fram tilhøvet mellom den målføreforma det blei søkt på og lemmaet?

Ein stor føremon for søkjesystemet er det faktumet at alle innhaldsartiklar i NO er kopla til ein artikkel i Metaordboka. Slik kan Metaordboka tene som nav i søkjesystemet, fordi mange formalt ulike inngangar til eit lemma er samla her. Korleis dette vil fungere i praksis, er illustrert i nokre konkrete søkjescenario i avsnitt 7. Figur 3 gjev eit oversyn over korleis søkjesystemet for enkle søk i eNO er tenkt å operere.



Figur 3. Søkjevegar for enkle søk i eNO.

Nummera i figuren illustrerer rekkjefølgda søka opererer i. Når brukaren tastar inn eit søkjeord i rubrikken for enkelt søk, skjer følgjande:

1) Systemet søkjer etter ordet blant oppslagsformene i eNO, og samstundes også i oppslagsformene i *Nynorskordboka* (NOB) og *Bokmålsordboka* (BOB), som begge er kopla saman med eNO gjennom Metaordboka. Alle oppslagsord som svarer til søkjeordet i desse tre basane, gjev treff og blir viste i søkjeresultatet. På denne måten vil ein brukar som søkjer på substantivet *jente*, få treff på ordartikkelen *gjente* i eNO (jf. det som blei sagt om denne forma i pkt. 2 ovanfor).

2) Om ein ikkje har fått treff, blir det søkt fleire andre stader i eNO-artiklane og i metaordboksartiklane. Det gjeld: a) Samansetjingsrekkjer inne i eNO-artiklar (forkorta SMS i figur 3). Desse er rekkjer med samansette ord som ikkje er definerte i NO, men berre lista opp under (ei tyding av) det grunnordet som utgjer førsteleddet i samansetjinga. Om søkjeordet er ei slik samansetjing, skal ein få opp den aktuelle grunnordsartikkelen der samansetjinga er førd opp. b) Målføreformer som er førde i eNO-artiklar. c) Alternative skriftformer som er førde i eNO-artiklar. Dette er former som ikkje fell saman med oppslagsforma, men som er eller har vore utbreidde i skrift. d) Oppslag på setlar i det elektroniske setelarkivet, som er kopla til dei grovlemmatiserte artiklane i Metaordboka. Oppslaget (eller overskrifta) på ein setel er nemleg ikkje alltid identisk med oppslagsforma i eNO (eller NOB eller BOB), og som vi skal sjå nedanfor, gjev det i nokre tilfelle høve til å fange opp søk som ikkje blir fanga opp på annan måte.

3) Om ein framleis ikkje har fått treff, blir det søkt i all tekst i eNO. Fritekstsøk vil til dømes kunne gje treff i unormaliserte sitat eller i oppslagsformer i eldre leksikografiske kjelder der ordet ligg føre.

4. Nokre moglege søkjescenario, problem og løysingar

I det følgjande skal vi gå gjennom nokre tenkte søkjescenario. Det er verdt å understreke at desse scenarioa er baserte på kva vi tenkjer oss at folk kan komme til å vilje søkje etter. Dei representerer med andre ord ikkje reelle søk. Når eNO er lansert for publikum, vil vi analysere korleis folk bruker søkjefunksjonen og om naudsynt justere søkjefunksjonen for å gje fleire relevante treff. Vi ser også for oss at informasjon frå søkjeloggar bør supplerast med andre undersøkingar som kan gje oss ein peikepinn ikkje berre om kva folk søker på og kva treff dei får, men også om kva dei faktisk var ute etter å finne då dei søkte.

4.1 Søkjescenario 1: «sjangle»

Brukaren tastar inn søkjeordet «sjangle» med føremålet å finne artikkelen for verbet som ein kan skrive slik i infinitiv på nynorsk og bokmål. Søkjeveg 1 i figur 3 fører for det første til treff på oppslagsforma *sjangle* i eNO, som er eit substantiv i hokjønn. Søket gjev også treff på verbet *sjangle* i NOB og BOB, som i Metaordboka er kopla til verbartikkelen **sjangla** i eNO (som nemnt i avsnitt 2 bruker NO *a*-infinitiv av verb, i tråd med 1938-normalen).

sjangle substantiv, hokjønn

innslag av striper i ty (Holla)

sjangla verb -a

gå ustøtt, rava (ll.3a), skjaga (1a), vingla; (òg:) halda på å falla; vakla

han sjangla mest sanselaus ut på sleen då han skulde køyre heim att (Mehl.Hall 231)

han (fekk) eit drag over nakken so han sjangla burtetter golvet (NFL11Ød 41)

Figur 4. Søkjetreff ved søk på «sjangle»: artikkane **sjangle** og **sjangla**.

Søkjer resultatet (figur 4) inneheld to artiklar, mellom dei verbartikkelen som brukaren ønskte å finne. Vi ser altså at sjølv om infinitiven i eNO-artikkelen endar på *-a*, kan vi ved hjelp av Metaordboka og koplingane til NOB og BOB lett fange opp søk der *e*-infinitiv blir nytta. Dette gjeld berre i dei tilfella der ordet det blir søkt på står i NOB og/eller BOB. Om det ikkje gjer det, må vi stole på andre metodar.

4.2 Søkjesenario 2: «skjangle»

I dette tilfellet søker brukaren etter verbet *sjangla* med ei ikkje-standardisert skriftform (*skj* i staden for *sj*) og med *e*-infinitiv. Søkjeveg 1 i figur 3 gjev ingen treff, ettersom *skjangle* ikkje er oppslagsord korkje i eNO, NOB eller BOB. Rett nok har eNO ein tilvisingsartikkel **skjangla** som viser til **sjangla**, men *a*-infinitiven i oppslagsordet hindrar søket frå å bli fanga opp på dette stadiet. Søkjeveg 2 fører heller ikkje til treff, men søkjeveg 3 (fritekstsøket) gjev treff i feltet for spesifisering av eldre kjelder:

The screenshot shows a search interface for the verb "sjangla" with the suffix "-a". At the top right, there are several icons: a speech bubble, a magnifying glass, a leaf, a red arrow, a red cross, and a red 'X'. Below the search bar, there are three main sections:

- Eldre kjelder**: A box containing the text "H u sjake, A1 633 skjangle, A2 skjangla og u skjaga, R skjangla, R3".
- gå ustøtt, rava (ll.3a), skjaga (1a), vingla; (òg:) halda på å falla; vakla**: A box containing two example sentences: "han sjangla mest sanselaus ut på sleen då han skulde køyre heim att (Mehl.Hall 231)" and "han (fekk) eit drag over nakken so han sjangla burtetter golvet (NFL11Ød 41)".
- np**: A box containing the sentence "ein trøyt natranglar kom skjanglande etter gata (Forsb.DNR 71)".

Figur 5. Søkjetreff ved søk på «skjangle»: artikkelen **sjangla**.

I feltet «Eldre kjelder» i artikkelen **sjangla** er det ført opp at kjelda A¹ (dvs. Aasen 1850) fører ordet på side 633 med forma *skjangle*. Via denne forma får brukaren opp nett denne artikkelen (og berre denne). Vi ser for øvrig at skriveforma med *skj*- også er attestert i eit av døma i artikkelen. Ein kan sjølvsagt hevde – med rette – at vi i dette tilfellet er heldige, ettersom det var langt frå gjeve at forma *skjangle* skulle dukke opp i artikkelen i det heile. Dette er likevel ein illustrasjon av at det gjeld å nyttiggjere seg informasjon i mange delar av artiklane ved søk i eNO.

4.3 Søkjescenario 3: «sjaundeferd»

Her søker brukaren på det samansette ordet *sjaundeferd*. Søkjeordet fell saman med den normaliserte forma av ordet. I og med at det ikkje finst ein artikkel med dette som oppslagsord korkje i eNO, NOB eller BOB, gjev ikkje søkjeveg 1 noko resultat. Derimot er ordet nemnt saman med andre samansetjingar i artikkelen **sjaunde-**, som er ein føreleddsartikkel som gjev opplysningar om i prinsippet alle ord som har dette føreleddet. Dette blir plukka opp av søkjeveg 2, og søkerresultatet må i dette tilfellet bli føreleddsartikkelen **sjaunde-**:



Figur 6. Søkjetreff ved søk på «sjaundeferd»: føreleddsartikkelen **sjaunde-**.

I tilfelle som dette, der søkjeordet står inni ein artikkel, bør treffet uthevast eller markerast med ein særskild farge, slik at brukaren lett ser nett det ho var ute etter å finne.

4.4 Søkjescenario 4: «sjondarbåt»

I dette scenarioet søker brukaren på «sjondarbåt», som er ei målføreform (attestert frå Hålandsdal i Hordaland) av det lemmaet som i NO er normert **sjaundebåt**. I mange tilfelle vil det truleg vere naturleg for brukarane å søkje på dialektord i den forma dei er vande med å høyre og bruke, fordi dette ofte er ord som ein til vanleg ikkje ser i skrift. Her gjev søkjeveg 1 ingen resultat, men søkjeveg 2 vil gje treff på oppslaget på den eine setelen frå Hålandsdal der forma *sjondarbåt* er brukt. På papirsetelen er *sjondarbåt* førd som tentativ normalisert form av innskrivaren, og denne forma er førd vidare som oppslag då papirsetelen blei digitalisert, slik vi ser i figur 7, som gjev att eit utsnitt frå Metaordboka der denne eine setelen er markert:

Oppslag	Grammatikk	K-status	Antal
☐ sjaund* <i>e</i> *båt	m	ok	2
☐ sjon <i>dar</i> båt	subst mask ubf sg appell	ok	
☐ sjundar-båt	m	ok	

Figur 7. Metaordboksartikkelen **sjaundebåt**; setelen med oppslaget *sjondarbåt* er markert.

Via Metaordboka kjem ein då fram til artikkelen **sjaundebåt**, som representerer det lemnaet søkjaren var ute etter:

sjaunde-båt substantiv, hankjønn



båt brukt til likferd, sjaund (SHI R, Hålandsdal)

Figur 8. Søkjetreff ved søk på «sjondarbåt»: artikkelen **sjaundebåt**.

4.5 Søkjesenario 5: «sennagress»

Dette scenariolet liknar mykje på det førre. Igjen søker brukaren på ei målføreform, «sennagress», som er attestert frå Vardø i Finnmark. Lemnaet det er snakk om, er i NO normalisert **sine-** eller **senne-gras**. NOB har oppslaget **sennegras**, medan BOB har **senne-gras** eller **-gress**. Som vi ser, fangar ingen av desse oppslaga opp den forma brukaren søker på. Men med søkjeveg 2 blir det igjen treff, her i form av eit oppslag i Ordbokshotellet, som er éin av ressursane som er samla i Metaordboka (sjå avsnitt 3). I figur 9 er det aktuelle oppslaget det fjerde nedanfrå. Det dreier seg om eit oppslag i Støme (1994), ei ordsamling frå Vardø.

Oppslag	Grammatikk	K-stat...	Antal
☐ sine*gras	n	ok	31
📖 senagras	subst nøyt ubf sg appell	ok	
📖 senegras	subst nøyt ubf sg appell	ok	
📖 senegras	subst nøyt ubf sg appell	ok	
📖 sen-gras	subst nøyt ubf sg appell	ok	
📖 sennegras	Ukjent	ok	
📖 sennelgras el. ~gress	Ukjent	ok	
📖 sinegras	n	ok	
📖 sinagras		ok	
📖 sennægras		ok	
📖 sennagress	n.	ok	
📖 senngræs	n.	ok	
📖 sennegras	subst nøyt ubf sg appell	ok	
📖 sinegras	n	ok	

Figur 9. Utdrag av metaordboksartikkelen **sinegras**; oppslaget *sennagress* er markert.

Også i dette tilfellet får altså brukaren treff på den artikkelen ho var ute etter, nemleg denne:

sine- eller **senne-gras** substantiv, inkjekjønn 

1. BOTANIKK, plante(art) i storfamilien; *Carex vesicaria* (VSO; Snåsa, Velfjord, Rødøy, Skånland, Nordreisa; J.LidElvenNF 1020) 

gjetorde fór ut yver verda som eld i sennegras (KohtTB 13)

2. (myr)gras el stort som blir turka og brukt som isolasjon i kommagar og anna skoty (Forsand, Nesna, Korgen, Fauske, Gratangen) 

inni kommagan bruka di tørt, mjukt senngræs, og då e di godt beskytta på føten (Håløygm 1962,76)

3.

a. eit slag gras som veks tett; (òg:) **SINE** (I,1a) (Sande på Snm, Ulstein, No A2, Beiarn, Vardø)

b. **tettegras** (Hetland)

c. namn på plante(arte)n kornstorr (No R4)

Figur 10. Søkjetreff ved søk på «sennagress»: artikkelen **sine-** eller **sennegras**.

4.6 Søkjesenario 6: «sennagras»

Til slutt vil vi sjå på eit meir problematisk, men ikkje urealistisk, søkjescenario. Brukaren søkjer på «sennagras». Dette er ikkje oppslagsform i eNO, NOB eller BOB, og ein finn det heller ikkje brukt som oppslag på setlar eller andre ressursar kopla til den aktuelle artikkelen i

Metaordboka. Heller ikkje eit fritekstsøk hjelper oss i dette tilfellet, og søket gjev dermed ingen treff. Det hjelper lite at forma *senna-* er førd opp som målførereform i føreleddsartikkelen **II sine-**, der nettopp opplysningar om samansetjingar av typen *sennagrass* er samla:

II sine-



Målføreformer (målførevariasjon annan enn lik oppslagsord eller sjølvgjeven)

- sina-*
Ulstein
- sene-*
No R4
- sena-*
Forsand, Sauda
- sen-*
No A2, Korgen, Beiarn
- senn-*
Fauske, Senja
- senna-*
Nesna, Gratangen, Vardø

til **II sine**

Figur 11. Føreleddsartikkelen **II sine-** med målføreformer.

Det er svært viktig kva informasjon brukaren får i dei tilfella der søket hennar ikkje gjev noko treff. Ho bør få informasjon om normeringsprinsippa som ligg bak forma på oppslagsorda i eNO (vi kan ikkje forvente at alle brukarane av eNO har eit reflektert forhold til tilhøvet mellom målføreformer og normaliserte standardformer). Samstundes bør ho få tips om meir avanserte søkje strategiar som kanskje kan hjelpe henne å finne rett artikkel. I tilfellet med *sennagrass* vil til dømes eit trunkert søk på «senn%gras» føre fram.

8. Konklusjon

NO er ei ordbok der artiklane har ein kompleks og informasjonstett struktur. I tillegg føreset den prenta utgåva at brukarane har ein viss kompetanse i norsk målførevitskap og ein viss kjennskap til normutviklinga i nynorsk (jf. til dømes at ordboka ikkje fører sjølvgjevne målføreformer, og at verket er redigert etter 1938-normalen). Formidlinga av dette ordboksmaterialet til eit breitt publikum er ei utfordring, og vi må gjere ein god del tilpassingar av informasjonsstrukturen og presentasjonen av artiklane i eNO. Søkjesenarioa vi har sett opp her, viser at den fullelektroniske ordboksbasen og dei elektroniske koplingane

mellom artiklar og underliggjande grunnlagsmateriale likevel gjev oss eit godt utgangspunkt for å tilby brukarane god søkjefunksjonalitet og relevante treff. Redigeringsapplikasjonen er utvikla av informasjonsteknologar ved EDD, som også står for utviklinga av nettversjonen eNO. Fordi redaksjonen og EDD står i eit fortløpande samarbeids- og utviklingsforhold, kjenner både leksikografane og informatikarane form, innhald og søkjefunksjonalitet i dei mange samankopla basane og i ordboksapplikasjonen. Dette gjer utviklingsarbeidet lettare.

eNO blei lansert i ein testversjon med alfabetbolken *i-skodda* 15. mars 2012. Neste steg i arbeidet med eNO vil vere å teste denne førsteversjonen på eit større publikum og bruke responsen vi får attende til vidare arbeid med utsjånad og funksjonalitet i eNO.

9. Litteratur

Aasen, Ivar 1850: Ordbog over det norske Folkesprog. Kristiania.

BOB = Bokmålsordboka <<http://www.nob-ordbok.uio.no>> (september 2011)

eNO = Nettutgåve av Norsk Ordbok. Testversjon: *i-skodda*. <no2014.uio.no> (april 2012)

Tvedt, Lars Jørgen, Elisabeth Lien og Øyvind Eide 2007: Ordbokshotellet – varig lagring og formidling av norske ordsamlingar. I: Torben Arboe (red.) Nordisk dialektologi og sociolingvistik: foredrag på 8. Nordiske Dialektologkonferanse, Århus 15.–18. august 2006. S. 379–388. <www.edd.uio.no/artiklar/leksikografi/Ordbokshotellet.pdf> (september 2011)

NO = Norsk Ordbok. Ordbok over det norske folkemålet og det nynorske skriftmålet. 1966–. Oslo: Det Norske Samlaget.

NOB = Nynorskordboka <<http://www.nob-ordbok.uio.no>> (september 2011)

Ore, Christian-Emil Smith 2001: Metaordboken – et elektronisk rammeverk for Norsk Ordbok? I: Martin Gellerstam et al. (red.): Nordiska studier i leksikografi 5. Rapport frå Konferens om Lexikografi i Norden, Göteborg 26.–29. maj 1999. S. 202–216.

Ore, Christian-Emil og Lars Jørgen Tvedt 2006: Digital sats eller digital satsing? I: Henrik Lorentzen og Lars Trap-Jensen (red.): Nordiske studier i leksikografi 8. Rapport fra Conference om Leksikografi i Norden Sønderborg 24.–28. maj 2005. S. 315–322.

Rommetveit, Magne (red.) 1979: Norsk landbruksordbok. Band 1–2. Oslo.

Skard, Matias 1954: Nynorsk ordbok for rettskriving og litteraturlesnad. Femte utgåva. Oslo.

Støme, Oddvar 1994: Vår eigen stemme. Tartu.

Svardal, Terje 2003: Norsk Ordbok – om normering av Metaordboka. I: Zakaris Svabo Hansen og Anfinnur Johansen (red.): Nordiske Studier i Leksikografi 6. Rapport fra Konferanse om leksikografi i Norden Tórshavn 21.–25. august 2001. S. 351–356.

Torp, Alf 1919: Nynorsk etymologisk ordbok. Kristiania.

Sturla Berg-Olsen
dr.art., redaktør/forskar
Norsk Ordbok 2014
Postboks 1021 Blindern
NO-0315 Oslo
sturla.berg-olsen@iln.uio.no

Åse Wetås
dr.art., prosjektdirektør
Norsk Ordbok 2014
Postboks 1021 Blindern
NO-0315 Oslo
ase.wetas@iln.uio.no

Nynorsk fagleksikografi

Sylvi Dysvik

Innleiing

Denne artikkelen byggjer i stor grad på ei masteroppgåve i nynorsk skriftkultur ved Høgskulen i Volda. Emnet for oppgåva var nynorsk fagspråk og terminologi i lys av nynorsk fagleksikografi, og som ein del av oppgåva prøvde eg å samle det som er vorte utgjeve av nynorsk fagleksikografi i tabellform. I denne artikkelen vil eg prøve å gje eit mest mogleg samla oversyn over den nynorske fagleksikografien, og samstundes ta opp noko av det som er særmerkt ved nynorsk fagspråk og terminologi. Oversynet gjeld berre trykte ordbøker, ord-lister og oppslagsverk. I tillegg har eg med noko om det som særmerkjer nynorsk fagspråk og terminologi, og litt om viktige aktørar innanfor den nynorske fagleksikografien.

Den nynorske skriftkulturen har gjerne vore knytt til andre sjangrar enn det som høyrer inn under sakprosa og fagspråk, og ein av dei folkelege mytane om nynorsk er at denne målforma skal høve særleg godt til skjønnlitteratur og anna diktning. Her kan det då liggje implisitt ein motsett myte om at nynorsk skal høve dårlegare enn bokmål til andre språklege ytringar, t.d. innanfor fagspråket. Dersom det skulle vere slik, vil dette seie at nynorsk ikkje kan reknast som eit fullverdig språk som er like godt brukande til alle føremål. Talet på fagleksiko-grafiske utgjevingar på nynorsk skulle tyde på at så ikkje er tilfelle.

Allmennleksikografi og fagleksikografi: avklaring av omgrep

Mange tenkjer gjerne på ordbøker som reint språklege hjelpemiddel, t.d. ved innlæring av eit framandspråk eller til kontroll av rettskriving. Men dei kan like gjerne vere eit hjelpemiddel til å auke forståinga innanfor eit særskilt fagområde, og til å utarbeide nye termar på grunnlag av etablert terminologi.

Allmennordbøker er ordbøker som ikkje er spesialiserte når det gjeld språkvarietetar, tekstar, ordboksstil, brukargrupper o.l. Her finn vi rettskrivings-ordlister og definisjonsordbøker med ordforklaringar

og definisjonar, gramma-tiske opplysningar og praktiske døme på bruken. Lars S. Vikør gjer nærmare greie for dei nynorske allmennordbøkene på s. 7-20.

Spesialordbøker er ei samlenemning for ordbøker som er spesialiserte når det gjeld språkvarietetar, tekstar, ordboksstil og brukargrupper, t.d. høyrer dei etymologiske ordbøkene og framandordbøkene heime her og likeins rimord-bøker, namneordbøker, kryssordbøker o.l. Den viktigaste gruppa i denne saman-hengen er likevel fagordbøkene eller dei terminologiske ordbøkene, som tek for seg terminologien på eitt eller fleire fagområde med sikte på dokumentasjon og standardisering, og kan rette seg både mot lekfolk og fagfolk. Ofte er slike bøker eit resultat av eit samarbeid mellom fleire fagfolk og/eller terminologar. I fag-ordbøkene vert det gjerne lagt mindre vekt på språkinformasjon og meir på definisjonar og forklaringar. Terminologiske ordbøker kan ofte ha ein norme-rande funksjon, fordi det her vert teke sikte på å standardisere terminologi som ikkje har vore fastlagt innanfor eit fag. Men dei kan òg vere deskriptive, dvs. først og fremst registrere førekomsten og bruken av eit ord eller ein term.

Nynorsk fagspråk og terminologi: særtrekk og problem

Mange av dei første fagordbøkene på nynorsk var fornorskingsordbøker med framlegg til nynorske avløysarord til dansknorsken. I den norske ordlagings-tradisjonen har omgrepslånet gjerne vore sett på som den beste ordlagingsmåten, med freistnader på å finne eit «heilnorsk» ord i staden for direkte omsetjing. Avløysarorda, eller fornorska ord, har såleis alltid hatt ein høg status i nynorsk. Normering av ordtilfanget har alltid stått sentralt i utviklinga av nynorsken, og puristiske haldningar når det gjeld det nynorske ordtilfanget har gjerne gjort at nynorsk har vorte sett på som eit forbodsspråk. Krava til språkleg purisme, bruk av avløysarord og termdanning med avleiing skapar gjerne problem på nynorsk i og med at eit fagspråkleg ordtilfang – eller ein terminologi – ofte inneber bruk av affiks som i ein puristisk tradisjon gjerne vert rekna som «dårleg nynorsk».

I fagspråket trengst det meir presise tydingar og definisjonar enn det som ofte krevst i allmennspråket. Praktisk terminologiarbeid har oftast ein tverr-vitskapleg karakter, ettersom både språkkunnskapar og fagkunnskapar trengst for å kome fram til eit best mogleg resultat.

Fagtermar må då kunne reknast både som ein del av språket og ein del av faget. Krava til fagspråkleg presisjon er høge og bør innebere ein høg grad av standardisering. I fagspråket er ofte import av ord og termar frå andre språk vanleg, anten som framandord eller lånord. Her er det framandorda som avvik mest, både i uttale, stavemåte og bøyingsmønster. Etter ei tid med aktiv bruk vil svært mange av dei importerte orda verte tilpassa norske språkreglar.

Nynorsk som målform har gjerne hatt ein lågare status i mange fagmiljø, eller kanskje i like stor grad: den spesielle språksituasjonen i Noreg med to jamstelte målformer har ikkje vore sett på som relevant eller interessant når ny terminologi skal fastsetjast. Då er termene berre vortne omsette utan at fagfolk har vore med på å fastsetje dei, og difor har termene gjerne små sjansar til å verte nytta i praktisk arbeid. Oftast er det bokmålsmiljøet som har lagt premissane, fordi bruken av bokmål vert rekna som det «normale», og utarbeidinga eller valet av ein term på nynorsk vil i stor grad vere prega av det som alt finst på bokmål. Dersom bokmålstermen har ei form og eit leksikalsk innhald som kan godtakast i begge målformer, så risikerer ein ikkje noka endring av innhaldsnyansar og fare for mistyding eller ulik tolking, og heller ikkje at termen går ut over særpreget eller formverket. Men ofte vil ein bokmålsterm vere såpass avvikande frå det nynorske formverket at det må utarbeidast ein eigen nynorskterm. Alternativet kan vere å nytte eit framandord med gresk/latinsk opphav, for framandordet er like «framandt» i bokmål som i nynorsk. Elles kan utanlandske ord og uttrykk ofte tilmåstast utan større problem dersom dei kan bøyast greitt, stavemåten tilpassast og elles kan gli nokolunde lett inn i det norske språket.

I nynorsk har det gjerne vore eit mål å unngå nominalisering og skrive om med verbalfomer. Dette er ei løysing som ofte er å finne i fornorskingsordbøker eller synonymordbøker mellom bokmål og nynorsk, og som gjerne fungerer betre i allmennspråkelege tekstar enn i fagspråket. Terminologiar er svært dominerte av substantiv og substantivsamansetningar, noko som ofte skapar tilleggsproblem på nynorsk. Omskrivingar kan ofte fungere dårleg, t.d. ved indeksering og bruk av oppslagsord i faglitteratur og fagordbøker.

Den nynorske fagleksikografien

Historia til den nynorske fagleksikografien kan først attende til Ivar Aasen. Han var den første nynorske fagleksikografen, og såg på

fagspråket som ein viktig del av utviklinga av eit nytt fullverdig språk som skulle kunne nyttast på alle område i samfunnet. Dei to store ordbøkene hans danna grunnlaget for det nynorske ordtilfanget. Mykje av arbeidet med omsetjing og uarbeiding av ordbøker i den tidlegaste nynorskperioden var prega av idealisme, og vart utført av fagfolk som ikkje først og fremst var filologar, men som var inspirerte av arbeidet til Ivar Aasen og opptekne av at nynorsk skulle kunne nyttast til alle føremål.

I tabellane nedanfor er dei ulike ordbokstypene førde opp i eigne kronologiske oversikter med kommentarar etter kvar tabell. Det er ikkje skilt mellom ordbøker og ordlister.

1 Oppslag og forklaringar på nynorsk (ev. supplert med bokmål og andre språk)

<i>Forfattar</i>	<i>År</i>	<i>Tittel</i>	<i>Utgjevar og stad</i>	<i>Merknader</i>
Medisinar-mållaget «Eir»	1939	Norsk medisinsk ordliste	Olaf Norlis Forlag, Oslo	Utgjeven av medisinar-mållaget «Eir», Oslo.
Magne Rommetveit (red)	1978	Norsk landbruks-ordbok I-II	Det Norske Samlaget, Oslo	Fleirspråkleg ordbok i to band. Utarbeidd av Nemnda for Norsk landbruksordbok / Studentmållaget i Ås.
Jan Engh og Lars Anders Kulbrandstad	1978	Språkvitenskapleg ordliste	Universitetsforlaget, Oslo	I Even Hovdhaugen (red.): Språkvitenskap: en elementær innføring s. 172-185

Nynorsk fagleksikografi

<i>Forfatter</i>	<i>År</i>	<i>Tittel</i>	<i>Utgjevar og stad</i>	<i>Merknader</i>
Tom Schmidt	1981	Fagordliste for norsk namnegransking	Universitetet i Oslo, Institutt for namnegransking	Oppslagsord parallelt på nynorsk og bokmål, forklaringar på nynorsk
Rådet for Norsk teknisk ordbok	1981	Norsk teknisk ordbok	Det Norske Samlaget, Oslo	Oppslag på både bokmål og nynorsk. Ny utgåve 1984 med ein del nye oppslagsord og definisjonar.
Audun Øyri	1988	Norsk medisinsk ordbok	Det Norske Samlaget, Oslo	Åtte utgåver, den siste kom i 2007 og har med engelske termar. Finst òg i elektronisk utgåve.
Audun Øyri	2002	Medisinsk lommeordbok	Det Norske Samlaget, Oslo	Boka inneheld engelske oppslag i parentes.
Hallstein Grønstøl	2004	Veterinærmedisinsk ordbok	Det Norske Samlaget, Oslo	Boka inneheld engelske oppslag i parentes.

I medisinfaget har mykje av lærestoffet vore på framandspråk. Latinske fagtermar er sjeldan noko problem. Her nærmar vi oss eit grunnleggjande problem for nynorsk terminologi som har med purisme, godtaking av framand-ord og bruk av avløysarord å gjere. I tillegg til ordboka frå 1939 gav medisinar-mållaget «Eir» i Oslo ut heftet Rettleiing i journalskriving i 1956, med omset-jingar av journalmønster som nynorskskrivinge lækjarar kunne nytte.

Medisinararen Idar Handagard gjorde eit viktig terminologiarbeid, sjølv om han ikkje utarbeidde reine ordbøker. Han skreiv m.a. bøkene Vanlege sjukdomar (1925) og Mannalikamen og helsa

(1927). Begge desse er opplysningsbøker som inneheld mykje oppnorska terminologi.

Den store medisinske ordboka til Audun Øyri kom første gongen i 1988, og var den første større medisinske ordboka på norsk. Ho kunne difor få innpass òg blant bokmålsbrukarar. Ordboka har gode etymologiske opplysningar, latinske nemningar og omfattande forklaringar til mange av oppslagsorda. I dei åtte utgåvene som er komne ut, har sidetalet auka mykje, og den siste utgåva inneheld ei ordliste med engelske oppslagsord, slik at ordboka no kan reknast som fleispråkleg. Etter at Audun Øyri døyddde i 2002 har kona hans, Bjørg Øyri, ført arbeidet vidare og er redaktør for den siste utgåva. Medisinsk lommeordbok inneheld eit utval på 7000 oppslagsord med korte forklaringar frå den store medisinske ordboka. Hallstein Grønstøl si Veterinærmedisinsk ordbok, som er utarbeidd etter same mønster som frå den medisinske lommeordboka, bør også nemnast her. Denne boka er den einaste på sitt fagområde, og har ein del av tilfanget sitt frå bøkene til Øyri. Ho inneheld dessutan eit vedlegg om diagnosar og definisjonar til bruk i veterinærfaget.

Landbruksfaga har vore eit sterkt domene for nynorsk språk. Lærebøkene på området har òg vore mange og har hatt stor utbreiing. Til landbrukshøgskulen på Ås kom det studentar frå heile landet, og mange av desse hadde nynorsk som hovudmål. Mange av studentane ønskte seg eit fagspråk som var mindre prega av dansk og tysk, og etterlyste ein norsk fagterminologi. Studentmållaget i Ås samla inn både midlar og ordtilfang til ei ordbok, og Det Norske Samlaget kom med i dette arbeidet frå 1952. Likevel kom ikkje Norsk landbruksordbok ut før i 1978, etter nærmare 40 års arbeid med innsamling og systematisering av fagtermar. Både NAVF og Landbruksdepartementet løyvde midlar, og Magne Rommetveit vart tilsett som redaktør i 1954. Ordboka har ei spesiell utforming i og med at ho i tillegg til eit hovudband med oppslagsord på norsk (bokmål og nynorsk), grammatiske opplysningar og forklaringar, òg har eit registerband med synonymordlister på dansk, svensk, engelsk, tysk, finsk og islandsk, med omsetjingar til norsk. I tillegg finst det vitskapelege nemningar og ein del samiske synonym frå relevante landbruksfaglege område. Ordboka er no overført til elektronisk format, og ordtilfanget ajourført og utvida. Det finst ingen tilsvarande ordbøker på bokmål, noko som truleg har gjeve denne ordboka ei større utbreiing enn om ho «berre» hadde måtte tevle med ei tilsvarande ordbok på bokmål.

Arbeidet med den store Norsk teknisk ordbok var ei viktig bragd som Studentmållaget i Nidaros hadde teke initiativ til. Ordboka var eit resultat av innsamling av fagtermar gjennom mange år, og mykje av arbeidet vart gjort som frivillig innsats der ordtilfanget vart samla inn av målinteresserte studentar og sivilingeniørar og jamført med definisjonar i ordbøkene frå Rådet for teknisk terminologi (RTT-ordbøkene) og i norske og internasjonale standardar. Boka har ein heller enkel bunad og er utan illustrasjonar. Ein slik utgjevingsmåte var vald for at det skulle bli rimelegare å gje ut nye utgåver (sjå føreordet), men boka kom berre i éi ny utgåve, i 1984.

2 Frå bokmål til nynorsk

<i>Forfattar</i>	<i>År</i>	<i>Tittel</i>	<i>Utgjevar og stad</i>	<i>Merknader</i>
Knut Knutson Fiane	1928	Norsk (landsmål). Ei lita ordliste til bruk i Telegrafstyret og andre telegrafkontor	Oslo	Kjenner ikkje til utgjevar. Er òg usikker på om boka er fleispråkleg, men det dreier seg truleg om ei fornorskingsordliste.
Nikolaus Gjelsvik	1929	Juridisk ordliste: med ei utgreiding om den juridiske stilen	Prenta hjå Nikolai Olsen, Oslo	Oppslagsord på bokmål.
Eigil Lehmann m.fl.	1931	Teologisk ordliste	Noregs Mållag, Oslo	Utarbeidd av ei nemnd av teologistudentar
Einar Sagen	1933	Militær ordliste med nynorsk tydning	Forsvarsdepartementet, Oslo	

<i>Forfattar</i>	År	Tittel	Utgjevar og stad	Merknader
Gustav Kullerud	1935	Nynorsk i apotek	Øens Prenteverk, Førde	Inneheld òg «bruksrett-leidingar» og brevdøme.
Edvard Os	1941	Militær ordliste	Det Norske Samlaget, Oslo	Nynorsk-bokmål og bokmål-nynorsk. Ny og utvida utgåve 1949.
Torstein Torsrud	1950	Teknisk ordliste	Studentmållaget i Nidaros	Bokmål-nynorsk.
Ingeniør-mållaget	1956	Teknisk ordliste. Bokmål-nynorsk	Noregs Boklag, Oslo	
Asgeir Olden	1978	Juridisk ordliste	Skuleboknemnda åt Studentmållaget i Oslo	Bokmål-nynorsk. Fleire opplag.
Karl Arne Utgård	2002	Juridisk og administrativ ordliste	Det Norske Samlaget, Oslo	
Språkrådet	2008	Kort administrativ ordliste bokmål-nynorsk	Språkrådet, Oslo	

Innanfor det juridiske fagspråket har nynorsk hatt lange tradisjonar. Den som truleg har gjort mest for det nynorske lovmålet, er professor Nikolaus Gjelsvik, som utarbeidde ei juridisk ordliste i 1929 og skreiv den språkpolitiske boka *Von og veg* i 1938. Den juridiske ordlista hans har ei lang innleiing der han gjer greie for kva som særmerkjer den juridiske stilen. Han åtvarar mot ord-for-ord-omsetjingar, og legg vekt på at ein må frigjere seg frå opphavsteksten. Gjelsvik omsette mange lover til nynorsk, og gav m.a. ut *Lover i utval på nynorsk* i 1936. Han hadde såleis stor innverknad på

utforminga av det nynorske lovmålet. Det juridiske språket har vore svært prega av tysk-danske vendingar, både i terminologi og syntaks, og det har vore eit stort arbeid å finne fullgode ekvivalentar på nynorsk, særleg fordi nynorsk er rikare på verb og verbale seiemåtar enn det bokmålet er.

Mindre kjend er kanskje den juridiske oppslagsboka med artiklar frå Norsk Allkunnebok, samla av Ingvald Svandal og utgjeven i 1969. Boka er meint både for studentar, skrivar- og lensmannskontor og andre som treng vite tydinga av juridiske uttrykk på nynorsk. Ettersom artiklane er henta direkte frå eit leksikon, er dei meir grundige og omfattande enn det som elles plar vere vanleg i ordbøker.

Det finst fleire gode ordbøker på det juridiske området, m.a. ei av høgste-rettsdommar Karl Arne Utgård. Den juridiske og administrative ordlista hans inneheld eit nyttig føreord med råd til dei som skal skrive nynorsk i juridiske og administrative samanhengar.

I denne ordboksgruppa finst det fleire lister som skulle danne grunnlaget for større ordbøker seinare, t.d. Nynorsk i apotek med «bruksretteleidingar» og brevdøme, utgjeven av Gustav Kullerud i 1935, og fleire tekniske ordlister utgjevne av Studentmållaget i Nidaros (1950) og Ingeniørmållaget (1956).

Militærstell er kanskje ikkje å rekne som noko eige fagområde, men eg har valt å ta dei med på grunn av den militære ordlista som vart utarbeidd av Edvard Os, som var offiser, organisasjonsmann og forfattar, og hadde ei rekkje verv innanfor målrørsla. I 1949-utgåva har han ei lang «Føremæling» der han gjer greie for prinsipp for utarbeiding av terminologi og fagspråk, og peikar på noko av det som er mest sentralt ved utarbeidinga av nytt fagspråk.

Ein annan militær målmann var kaptein, seinare major Einar Sagen. I 1910 gav han ut den vesle boka Norsk her-mål. Størstedelen av boka har parallelltekst på dansk og nynorsk, med prøver på omsetjingar av lover og føresegner, og døme på brev, som på same måte som brevbøkene til Koht og Os skulle vere til hjelp for dei som ville nytte den nye målforma. I nokre av breva har han teke med alternative seiemåtar, ordforklaringar og opplysningar om etymologi. Først i 1933 kom ordlista som han nemner i føreordet til boka frå 1910, Militær ordliste med nynorsk tydning, utgjeven av Forsvarsdepartementet.

3 Nynorsk til/frå framandspråk

<i>Forfatter</i>	<i>År</i>	<i>Tittel</i>	<i>Utgjevar og stad</i>	<i>Merknader</i>
J.O. Hennem	1886	Anatomiske Termini fra det Norske Landsmål	Det Steenske Bogtrykkeri, Kristiania	Utgjeve som «Tillægshäfte til Norsk Magazin for Lægevidenskaben».
Halvdan Koht	1898	Ordbok. Norsk umsetjing utav danske, franske, latinske og andre framande handelsord	Bertrand Jensen, Kristiania	I Norsk brevbok (s. 41-74). Ordliste med meir enn 1 800 handelsord.
Edvard Os	1927	Norsk brevbok	Det Norske Samlaget, Oslo	I alt fem utgåver, den siste kom i 1958. Oppslagsordliste på 15 sider med framlegg til nynorske termar.
Einar Austlid	1971	Norsk-engelsk ordliste for fiskarar	Bjarne H. Reenskaug forlagsaksjeselskap, Oslo	Hovudoppslag og ordforklaringar på nynorsk, med synonym på bokmål og engelsk.
Johan Ottesen	1971	Engelsk-norsk sosiologisk ordliste	Skulebokne mnda åt Studentmållaget i Oslo	Liste utan ordforklaringar.
Johan Ottesen	1973	Engelsk-norsk ordliste i EDB	Skulebokne mnda åt Studentmållaget i Oslo	Liste utan ordforklaringar.
Johan Ottesen	1974	Engelsk-norsk ordliste i politikk	Skulebokne mnda åt Studentmållaget i Oslo	Liste utan ordforklaringar.

<i>Forfattar</i>	<i>År</i>	<i>Tittel</i>	<i>Utgjevar og stad</i>	<i>Merknader</i>
Johan Ottesen	1974	Engelsk-norsk ordliste i sosiologi	Skulebokne mnda åt Studentmål-laget i Oslo	Liste utan ordforklaringar. Ny utgåve av ordlista frå 1971.

Alt i 1886 kom dr.med. J.O. Hennem med «Anatomiske termini fra det norske landsmål» som tilleggshefte til «Norsk Magazin for lægevidenskabem». Her har han samla norske nemningar for anatomiske fagord og -uttrykk. Han har med forklaringar, og i mange tilfelle finst det opplysningar om kvar i landet ordet er henta frå. Ofte er òg svenske nemningar tekne med. Lista er ordna etter «legemets inndeling» og ikkje alfabetisk, men det finst eit alfabetisk register med sidetilvising.

Brevboka til Koht var den første boka med nynorsk handelsterminologi. Slike brevbøker var freistnader på å lage høvelege uttrykksmåtar til bruk i handelsbrevveksling på nynorsk. Boka inneheld ei lang ordliste med norsk om-setjing av framande handelsord, og gjev elles døme både på «Ervebrev», «Erve-pakt millom ægtefolk», «Hyrebrev» og «Varetilbod». Koht held seg mest til dei offisielle brevdøma, medan Edvard Os i si brevbok frå 1927 òg har med malar for skrivning av både takkekort, søknader, fullmakter, møtebøker o.l. på nynorsk. Brevboka til Os har òg ei rettleiing om rettskriving og målbruk og ei for-norskingsordliste på 15 sider. Ho kom i fem utgåver i åra 1927-58. Begge desse brevbøkene har truleg vore til stor hjelp for folk som ville nytte nynorsk i meir offentlege samanhengar. Bortsett frå dei gamle brevbøkene til Os og Koht finst det svært få ordbøker på området økonomi, handel og næringsliv.

Den første fagordboka med begge målformer er Austlids ordliste for fiskarar, der hovudoppslaga og ordforklaringane er på nynorsk, men der det òg finst ei kolonne med synonym på bokmål.

Dei tre fagordlistene i edb, sosiologi og politikk er reine ordlister utan forklaringar og definisjonar, der noko av føremålet skulle vere å finne gode avlørsarord til engelske faguttrykk. Ottesen seier i føreordet til 2. utgåva av sosiologiordlista at motstanden mot ei slik fornorsking kan ventast å verte stor i fagmiljøet, der mange av framandorda er vortne innarbeidde utan for mange motførestellingar. Desse ordlistene har ikkje kome i nye utgåver.

4 Språkblanda bokmål og nynorsk, til/frå framandspråk

Forfatter	År	Tittel	Utgjevar og stad	Merknader
Gösta Bergman (red.)	1970	Turistord i Norden	Norsk språk-nemnd. Skrifter nr. 4, Borgå	Dansk, finsk, færøysk, islandsk, svensk, bokmål og nynorsk
Kolbjørn Heggstad (red.)	1976	Norsk dataordbok	Universitetsforlaget, Oslo	Tospråkleg bok. Nynorsk er oppført der det er avvik frå bokmål.
Rådet for teknisk terminologi, RTT	1979	Ordbok for tele-teknikk (RTT 40)	Universitetsforlaget, Oslo	Fleispråkleg ordbok. Nynorsk er oppført der det er avvik frå bokmål.
Rådet for teknisk terminologi, RTT	1980	Ordbok for varslings- og alarmteknikk (RTT 41)	Universitetsforlaget, Oslo	Fleispråkleg ordbok. Nynorsk er oppført der det er avvik frå bokmål.
Rådet for teknisk terminologi, RTT	1980	RTT 42: Ordbok for VVS (RTT 42)	Universitetsforlaget, Oslo	Fleispråkleg ordbok. Nynorsk er oppført der det er avvik frå bokmål.
Rådet for teknisk terminologi, RTT	1980	Ordbok for petroleums-virksomhet (RTT 43)	Universitetsforlaget, Oslo	Norsk-engelsk / engelsk-norsk. Nynorsk er oppført der det er avvik frå bokmål.
Rådet for teknisk terminologi, RTT	1980	Ordbok for bygningsglass (RTT 44)	Universitetsforlaget, Oslo	Fleispråkleg ordbok. Nynorsk er oppført der det er avvik frå bokmål.

Nynorsk fagleksikografi

<i>Forfatter</i>	År	Tittel	Utgjevar og stad	Merknader
Norsk språkråds komité for markedsføringsterminologi	1981	Norsk markedsføringsordbok	Universitetsforlaget, Bergen	Språkblanda bok. Nynorsk er oppført der det er avvik frå bokmål.
Rådet for teknisk terminologi, RTT	1981	Ordbok for vannturbiner (RTT 45)	Universitetsforlaget, Oslo	Fleispråkleg ordbok. Nynorsk er oppført der det er avvik frå bokmål.
Norsk språkråd	1982	Oljeordliste	Universitetsforlaget, Oslo	Engelsk og norsk. Nynorsk er oppført der det er avvik frå bokmål. Samarbeid mellom Språkrådet og RTT.
Rådet for teknisk terminologi, RTT	1983	Ordbok for varmekraft. Del 2: Kjel- og turbinanlegg (RTT 47)	Universitetsforlaget, Oslo	Fleispråkleg ordbok. Nynorsk er oppført der det er avvik frå bokmål.
Håvard Hjulstad og Bjarne Norevik (red.)	1984	Norsk teknisk fagordbok	Universitetsforlaget, Bergen	Språkblanda bok med engelske ekvivalentar, storparten er ei samanstøyping av stoff frå andre fagordbøker. Innslaga frå Norsk teknisk ordbok er på nynorsk.

Nynorsk fagleksikografi

<i>Forfatter</i>	<i>År</i>	<i>Tittel</i>	<i>Utgjevar og stad</i>	<i>Merknader</i>
Patrick Chaffey (red.)	1984	ADNOM. Norsk-engelsk glossar. Navn og termer fra offentlig og privat virksomhet	Universitetsforlaget, Oslo	Samarbeid mellom UiO, Norsk termbank, UD og Univ. of Surrey. Språk-blanda bok. Nynorsk er oppført der det er avvik frå bokmål.
Rådet for teknisk terminologi, RTT	1984	Ordbok for varmekraft. Del 1: Reaktor-anlegg (RTT 46)	Universitetsforlaget, Oslo	Fleirspråkleg ordbok. Nynorsk er oppført der det er avvik frå bokmål.
Rådet for teknisk terminologi, RTT	1985	Ordbok for arbeidervern (RTT 49)	Universitetsforlaget, Oslo	Fleirspråkleg ordbok. Nynorsk er oppført der det er avvik frå bokmål.
Rådet for teknisk terminologi, RTT	1986	Ordbok for energiteknikk (RTT 51)	Universitetsforlaget, Oslo	Fleirspråkleg ordbok. Nynorsk er oppført der det er avvik frå bokmål.
Rådet for teknisk terminologi, RTT	1986	Ordbok for maling-, lakk- og trykkfargeteknikk (RTT 52)	Universitetsforlaget, Oslo	Fleirspråkleg ordbok. Nynorsk er oppført der det er avvik frå bokmål.
Rådet for teknisk terminologi, RTT	1986	Ordbok for korrosjonsteknikk (RTT 53)	Universitetsforlaget, Oslo	Fleirspråkleg ordbok. Nynorsk er oppført der det er avvik frå bokmål.

Nynorsk fagleksikografi

Forfatter	År	Tittel	Utgjevar og stad	Merknader
Rådet for teknisk terminologi, RTT	1987	Ordbok for varslings- og alarmteknikk (RTT 54)	Universitetsforlaget, Oslo	Fleirspråkleg ordbok. Nynorsk er oppført der det er avvik frå bokmål.
Patrick Chaffey (red.)	1988	Norsk-engelsk administrativ ordbok. Navn og termer fra offentlig virksomhet	Universitetsforlaget, Oslo	Ny utg. av ADNOM. Samarbeid mellom UiO, Norsk termbank, DU, Univ. of Surrey og RTT. Språkblanda bok. Nynorsk er oppført der det er avvik frå bokmål.
Ingrid Dahlø og Aksel Lydersen	1988	Ordbok for kjemiteknikk	Tapir, Trondheim	Norsk-engelsk / engelsk norsk. Nynorsk er oppført der det er avvik frå bokmål.
Rådet for teknisk terminologi, RTT	1988	Ordbok for plast- og gummiteknikk (RTT 55)	Universitetsforlaget, Oslo	Fleirspråkleg ordbok. Nynorsk er oppført der det er avvik frå bokmål.
Rådet for teknisk terminologi, RTT	1988	Ordbok for fluidteknikk (RTT 56)	Tapir, Trondheim	Fleirspråkleg ordbok. Nynorsk er oppført der det er avvik frå bokmål.
Rådet for teknisk terminologi, RTT	1989	Ordbok for kart og oppmåling (RTT 57)	Statens Kartverk	Fleirspråkleg ordbok. Nynorsk er oppført der det er avvik frå bokmål.

<i>Forfatter</i>	<i>År</i>	<i>Tittel</i>	<i>Utgjevar og stad</i>	<i>Merknader</i>
Rådet for teknisk terminologi, RTT	1994	Ordbok for romvirksomhet (RTT 59)	Formatic AS	Fleirspråkleg ordbok. Nynorsk er oppført der det er avvik frå bokmål.
Rådet for teknisk terminologi, RTT	1994	Ordbok for fiberoptikk (RTT 62)	Formatic AS	Fleirspråkleg ordbok. Nynorsk er oppført der det er avvik frå bokmål.
Kjell E. Innli	1995	Gastronomisk ordbok	Yrkeslitteratur	Fleirspråkleg ordbok. Nynorsk er oppført der det er avvik frå bokmål.
Nordiska språk- och informations-centret	1995	Pensjonstermer Finland-Norge-Sverige	Nordiska språk- og informations-centret, Helsingfors	Fleirspråkleg ordbok.
Nordisk språkråd	1996	Nordisk møteordliste	Nordisk språkråd	Fleirspråkleg ordbok. Både nynorsk og bokmål. Ny utgåve i 2003 under tittelen "Nordisk mötesordlista".
Henning Bergenholtz, Ilse Cantell, Ruth Vatvedt Fjeld, Dag Gundersen, Jón Hilmar Jónsson og Bo Svensén	1997	Nordisk leksikografisk ordbok	Universitetsforlaget, Oslo	Fleirspråkleg ordbok. Redigert på bokmål. Nynorske parallell-termar.

Nynorsk fagleksikografi

Forfatter	År	Tittel	Utgjevar og stad	Merknader
Erling Joar Fløttum	1997	Ordlister og definisjoner i økonomisk statistikk. Engelsk-bokmål-nynorsk	Statistisk sentralbyrå, Oslo	Engelsk-norsk / norsk-engelsk ordliste. Eigne bolkar med nynorsk og bokmål.
Rådet for teknisk terminologi, RTT	1997	Ordbok for energi (RTT 60)	Formatic AS	Fleirspråkleg ordbok. Nynorsk er oppført der det er avvik frå bokmål.
Rådet for teknisk terminologi, RTT	1998	Ordbok for elektronikk (RTT 64)	Formatic AS	Fleirspråkleg ordbok. Nynorsk er oppført der det er avvik frå bokmål.
Norsk språkråd	2000	Ordbok for fysikk og kjemi til skole-bruk	Norsk språkråd, Oslo	Norsk språkråds skrifter nr. 7. Nynorsk- og bokmålstermar. Både nynorsk og bokmål er nytta som redigerings-språk.
Arild Torvund Olsen og Ingvild Aursøy Måseide	2002	Medieordlista: nynorske fagomgrep for film- og medievitskap	Studentmål-laget i Nidaros	Bokmål-nynorsk, med ein del engelske oppslagsord. Finst òg i elektronisk utgåve.
Jens Morten Nilsen	2005	Realfagsordlista	Studentmål-laget i Nidaros	Ein del av oppslagsorda er på engelsk. Finst òg i elektronisk utgåve.

<i>Forfatter</i>	<i>År</i>	<i>Tittel</i>	<i>Utgjevar og stad</i>	<i>Merknader</i>
Sigbjørn Hjelmbrekke	2007	Økonomisk ordliste: nynorske fagomgrep for samfunnsøkonomi	Studentmål-laget i Oslo	Todelt liste der hovuddelen er ei engelsk-norsk liste (om lag 600 oppslag) og den andre delen har om lag 200 oppslag på bokmål med nynorsk omsetjing. Finst òg i elektronisk utgåve.

Denne tabellen omfattar berre ordbøker og -lister som er utgjevne etter 1970, og er den mest omfattande lista. Grunnen til dette er nok at det no vart sett fram strengare krav til at slike ordbøker skulle ha nynorskoppføringar.

Norsk teknisk fagordbok inneheld fagterminologi frå Norsk termbank og frå ei rekkje andre tekniske ordbøker, og termoppslaga har tilvisingar til desse kjeldene. Oppslagsorda og forklaringar finst på begge målformer, alt etter kvar termane og forklaringane er tekne frå. Boka har engelske ekvivalentar, og eit engelsk-norsk register på 90 sider. I tillegg finst det nyttige bolkar om terminologi, måleiningar, omrekningstabellar o.l. I føreordet til denne boka òg står det at ho var tenkt utgjeven i nye, ajourførde versjonar. Dette har diverre ikkje skjedd.

Når det gjeld tekniske fag elles, er det RTT som har gjeve ut det meste av nynorsk terminologi, men i og med at denne terminologien oftast har vore reine omsetjingar frå bokmål, veit vi lite om kor innarbeidd han er vorten i fagmiljøa. På desse fagområda skjer òg utviklinga så snøgt at det er vanskeleg å utarbeide oppdaterte ordbøker i det heile.

Frå 1970-talet og framover kom det mange fagordbøker som i hovudsak er redigerte på bokmål, men som har med nynorskformer der det er avvik mellom målformene. Ordbøkene frå RTT er typiske døme på dette. Den første av ordbøkene med denne typen redigering kom i 1979.

Nynorsk fagleksikografi

I denne perioden gav Nordisk språksekretariat ut fleire felles-nordiske ordbøker, m.a. ei turistordbok (1970), ei liste med pensjonstermar (1995) og ei møteordliste (1996).

Føremålet med dei to ADNOM-bøkene med nomenklatur over administra-tive namn og nemningar var å samle og standardisere namna på norske yrkes-nemningar, institusjonar, organisasjonar o.l. på engelsk. Prosjektet vart nedlagt på 1990-talet.

Dei siste tre ordlistene i tabellen har oppslag både på engelsk og norsk. Alle listene er utarbeidde på frivillig basis av målmedvitne studentar, som såleis fører vidare ein lang tradisjon innanfor den nynorske fagleksikografien ved å gjere ein innsats for å gje nynorsk ein plass i faget. Studentmållaget i Nidaros har dess-utan laga ei organisasjonsordliste med avløysarord, som berre finst tilgjengeleg i nettversjon.

Samla kan vi seie at den særlege støtteordninga for nynorsk litteratur gjennom Norsk kulturfond, som vart innført frå 1966, kan ha ført til at det vart satsa meir på nynorsk òg innanfor faglitteraturen. Men like viktig var kanskje auken i studenttalet og at nynorskbrukande studentar engasjerte seg stadig sterkare i kampen for å få gjennomføre det lovfesta kravet om at lærebøker skal liggje føre i begge målformer til same tid og til same pris. Studentmållaget i Oslo kom på banen att og gav ut ein serie med fagordlister i tidsrommet 1971-74 og ei juridisk ordliste i 1978.

5 Spesialordbøker

<i>Forfatter</i>	<i>År</i>	<i>Tittel</i>	<i>Utgjevar og stad</i>	<i>Merknader</i>
Ivar Aasen	1876	Norsk Maalbunad	Det Norske Samlaget, Oslo	Utgjeven først i 1925. Elektronisk utgåve 1997.
Severin Eskeland	1919	Framandord-bok	Olaf Norlis Forlag, Oslo	Utgjeven på nytt i fotografisk avprent i 1995.

<i>Forfatter</i>	<i>År</i>	<i>Tittel</i>	<i>Utgjevar og stad</i>	<i>Merknader</i>
Alf Torp	1919	Nynorsk etymologisk ordbok	H. Aschehoug & Co, Kristiania	Ny faksimile-utgåve 1992 ved Bjørn Ringstrøms Antikvariat, Oslo. Ordforklaringane er på dansk.
Universitetet i Bergen, Nordisk institutt	1972	Nynorsk baklengsordliste	Universitetsforlaget, Bergen	Prosjekt for datamaskinell språkbehandling. Norsk språkdata, rapport 2. Tilsvarande rapport finst på bokmål.
Per Vestbøstad (red.)	1989	Nynorsk frekvensordbok. Dei vanlegaste orda i skriftleg nynorsk	Alma Mater Forlag AS	Henta frå aviser, sakprosa og romanar 1978-84.
Kolbjørn Hauge	2000	Stor norsk rimordbok	Det Norske Samlaget, Oslo	Både bokmål og nynorsk. To utgåver.

Norsk Maalbunad er ei synonymksamling der mykje av tilfanget er henta frå prosastykka i allkunneboka Heimsyn, som Aasen gav ut i 1875. Boka var resultatet av ei ordsystematisering som Aasen hadde halde på med heilt sidan han systematiserte dyre- og plantenamn på 1830-talet.

Severin Eskeland gav ut ei framandordbok i 1919. Boka kom ut på «nynorskforlaget» Olaf Norli, men vart òg utgjeven som tillegg til vekebladet For bygd og by. Den sistnemnde utgåva var utvida med eit «Fyrerit» med døme på brev, søknader o.l. som likeins med brevbøkene til Koht og Os inneheld døme på korleis ymse brev, vedtekter og andre dokument kunne skrivast på nynorsk.

Av andre større fagleksikografiske arbeid bør vi nemne den store Nynorsk etymologisk ordbok frå 1919, som finst i ei tilsvarande utgåve «over det norske og danske sprog» av Falk og Torp.

6 Leksikon, nynorsk

Forfattar	År	Tittel	Utgjevar og stad	Merknader
Arnulv Sudmann (hovudred.)	1948	Norsk Allkunnebok	Fonna Forlag, Oslo	Ti band, det siste kom i 1966. Verket omfatta òg eit atlas (1962).
Ingvald Svandal (red.)	1969	Norsk Allkunneboks juridiske oppslagsbok	Fonna Forlag, Oslo	Signerte artiklar om jus og økonomi frå Norsk Allkunnebok.
Rolf Ridar, Ingebrigt Davik, Jan Strand og Olav Åm	1979	Hermes leksikon. For skule og heim	Hermes Forlag, Oslo	Nynorsk leksikon i to band. Finst òg på bokmål.
Ottar Lundstøl	1987	Nynorsk leksikon	Faktum Forlag, Oslo	To band. Finst òg på bokmål som Faktum leksikon

Mykje av institusjonsbygginga innanfor den nynorske skriftkulturen har gått føre seg etter eit prinsipp om parallell eller alternativ organisering, der nynorsk-miljøa har bygd opp sine eigne institusjonar der nynorsk vart nytta som mål-form. Den parallelle organiseringa som har prega nynorskmiljøet, har ført til eit eige leksikon på nynorsk: Norsk Allkunnebok (ti band, 1948-1966). I tillegg kom Hermes leksikon i 1979 og Nynorsk leksikon i 1987, dei einaste leksikona som ligg føre både på bokmål og nynorsk. Norsk Allkunnebok hadde eit opplag på 10 000 eksemplar, og men utgjevingane vart hemma av at forlaget fekk problem med økonomien på slutten av 1950-talet, slik at verket måtte kortast ned. Dei første åtte banda går fram til og med bokstaven n, medan dei siste to banda inneheld resten av alfabetet. I føreordet nemner redaktørane dei særlege problema som er knytte til at det mangla nynorsk terminologi på mange fagområde. Den nynorske leksikontradisjonen vert no førd vidare i det digitale oppslagsverket Allkunne, som vart lansert hausten 2009.

Tabellane syner ein klar auke i utgjevingstalet for fagordbøker, spesialord-bøker og leksikon frå 1970-talet og framover. Når talet på utgjevne fagordbøker held seg høgt, kan grunnen vere at fleire av dei verka som kom ut, hadde vore under arbeid i lengre tid, m.a. dei tekniske ordlistene (utgjevne i 1950 og 1956). Fleire av dei andre utgjevingane var det kjende og aktive personar i nynorsk-rørsla som stod for, t.d. Edvard Os og Nikolaus Gjelsvik.

Viktige aktørar

Blant utgjevarane av nynorskordbøker samla sett er det to viktige aktørar som bør nemnast, nemleg forlaga og studentmållaga. Når det gjeld dei språkblanda bøkene som er oppførde i tabell 4, er dessutan Rådet for teknisk terminologi (RTT) ein viktig aktør.

I den tidlegaste tida til nynorsken var det eit uklart skilje mellom avis- og forlagsdrift, og det var gjerne boktrykkjarane og seinare forleggjarane, som stod for sjølve boksalet. Dei fleste ordbøkene som kom ut før 1900, vart utgjevne av slike forlag, og vart ofte selde i kommisjon. Ofte var det forfatarane sjølve som måtte betale for prentinga av bøkene sine.

Samla sett har Det Norske Samlaget stått for dei fleste ordboks-utgjevingane på nynorsk, og dette forlaget har stått for utgjevinga av sentrale fagordbøker som Norsk medisinsk ordbok og Norsk landbruksordbok. Likevel er det Universitetsforlaget som toppar lista når det gjeld fagordbøkene, men dette forlaget har ingen reine nynorskutgjeingar. Den første utgjevinga kom i 1976 (Norsk data-ordbok). Tradisjonelle nynorskforlag som Fonna Forlag, Norsk bokreidingslag og Noregs Boklag har svært få utgjeingar av fagordbøker.

Nynorskbrukande studentar var svært aktive og engasjerte seg sterkt i målpolitisk arbeid, og mange av dei la ned eit stort frivillig arbeid for å setje om fagbøker og utarbeide ordbøker og fagterminologi på nynorsk. Både i Oslo, Bergen, Trondheim og Ås gjorde studentmållaga ein stor innsats. Det var studentane som i stor grad tok i bruk nynorsk som akademisk språk. Det fanst ikkje lærebøker på nynorsk, og mange av studentane gjorde ein stor innsats ved å setje om lærebøker til nynorsk og utarbeide eit nynorsk fagspråk. Studentmållaget i Oslo har nesten vore som eit eige forlag å rekne, som både har hatt eigne utgjeingar, samarbeidd med andre og gjeve pengestøtte til utgje-vingar på andre forlag. Som i målrørsla

elles stod folkeopplysningstanken sterkt. Laget starta utgjevinga av folkeopplysningsserien Norske Folkeskrifter i sam-arbeid med Noregs Ungdomslag og seinare Noregs Boklag (frå 1939). Desse skriftene var viktige for utviklinga og spreinga av nynorsk fagspråk og terminologi på mange ulike område. Mykje av arbeidet med store og sentrale fagleksikografiske verk tok til etter initiativ frå målstudentane. Skuleboknemnda åt Studentmållaget i Oslo har heilt fram til våre dagar vore ein viktig aktør, men hadde den viktigaste perioden sin frå kring 1910 og fram til andre verdskrigen. Fleire sentrale medlemmer av Skuleboknemnda, t.d. Olav Midttun, Ola Raknes og Johannes Lid, vart seinare rekrutterte til arbeid i eller for Det Norske Sam-laget på grunnlag av røynslene dei hadde frå forlagsarbeidet i Skuleboknemnda.

Rådet for teknisk terminologi (RTT) vart skipa ved NTH alt i 1938 som eit samarbeidstiltak mellom fleire universitet, bransjeforeiningar innanfor tekniske fag og offentlege styresmakter, og vart seinare gjort om til ei stifting som var finansiert dels gjennom statlege midlar og dels gjennom oppdrag. Føremålet var m.a. å utarbeide tekniske ordbøker, og over 60 slike bøker er gjevne ut innanfor ulike tekniske fagområde. Det står ingenting i føremålsparagrafen om ansvar for utarbeiding av terminologi på nynorsk, men i den grad nynorske termar er tekne med, vart dette gjort av Språkrådet ved den språklege gjennomgangen av bøkene, eller av språkfolk i RTT. I dei fleste av desse ordbøkene er nynorsk-termar tekne med berre der det er skilnad mellom dei to målformene. I alt kom det ut 20 RTT-ordbøker med slike nynorskoppføringer, dei fleste av dei vart utgjevne av Universitetsforlaget og forlaget Formatic AS.

Elles bør det nemnast at institusjonane alltid har spela ei stor rolle i nynorskhistoria, og har medverka monaleg for å sikre stillinga til og bruken av nynorsk som eit fullverdig skriftspråk på alle område i samfunnet, slik Ivar Aasen såg føre seg at den nye målforma skulle vere. Politiske vedtak og støtteordningar har gjort det mogleg å fremje og sikre nynorsken ein plass i det offentlege, og dei mange utgjevingane av nynorsk fagleksikografi frå 1970-talet og framover er ei direkte følgje av slike vedtak og ordningar.

Konklusjon

Mykje av utbreiinga av nynorsk språk i det heile skuldast idealistiske einskild-personar som gjennom dugnadsånd og frivillig arbeid gjorde ein stor innsats for nynorsken. Ein slik innsats har òg prega arbeidet med den nynorske fagleksikografien, for slikt arbeid har ofte vore sett i verk på idealistisk grunnlag av sentrale målfolk, studentar, språkinteresserte fagfolk og andre som har sett det som viktig og naudsynt at nynorsk skulle takast i bruk på alle område i sam-funnet. Ei slik dugnadsånd var særleg viktig fram i den tidlege perioden til nynorsken og fram til 1960-talet, då statlege støtteordningar vart innførde for å sikre nynorsken ein plass i ordbøkene. Mange ord-boksprosjekt hadde ikkje vore mogleg å gjennomføre utan slik støtte. Likevel finst det framleis nynorskbrukande og målmedvitne studentar som på idealistisk vis og med støtte frå studentmållaga fører denne tradisjonen vidare.

Ordtilfanget i nynorsk har alltid vore eit stridsemne, og oftast har det vore strid om i kor stor grad det nynorske skriftmålet skal opnast for ord som har vore haldne ute tidlegare. Når det gjeld nyare lånord, har det gått eit viktig skilje mellom meir internasjonale ord (frå latin og gresk) og lånord frå dansk og tysk. Ettersom nynorsk voks fram som ein motsetnad til dansknorsk, har motstanden vore større mot den sistnemnde typen lånord, sjølv om slike ord heile tida har vore å finne i folkemålet. Motsetnaden har såleis stått mellom ei puristisk linje og ei meir liberal linje, og purismen har vore ein viktig språkspesifikk faktor i nynorsken. Purisme og strid om kva som vert rekna som «god nynorsk» har gjort arbeidet med utarbeiding av nynorsk terminologi særleg utfordrande.

Det verkar som om det krevst statleg styring og press for å få inn nynorsk terminologi i elektroniske basar, lister og ordbøker, og at den kommersielle marknaden i svært liten grad spør etter nynorsk. Likevel vitnar omfanget av dei fagleksikografiske utgjevingane om vilje til og ønske om at nynorsk skal ha ein plass i arbeidslivet. Både ordbøker og store leksikonverk har lenge vore til-gjengelege som internettutgåver, noko som gjer tekstmassen lettare å søkje og finne fram i og enklare å ajourføre. I framtida vil nok nye bokutgåver vere sjeldne, og nettutgåvene gjer det mogleg å søkje og finne fram til kunnskap på nye måtar, og alt arbeid som har med ajourføring og feilretting vert òg lettare å gjennomføre. Ei slik utvikling bør vere ein føremon for vidareføringa av den rike nynorske fagleksikografien.

Litteratur

Dysvik, Sylvi. 2009. Frå spongvegjer til høgfartsbanar. Om nynorsk fagspråk og terminologi i lys av nynorsk fagleksikografi. Mastergradsoppgåve i nynorsk skriftkultur ved Høgskulen i Volda.

Haugen, Eva L. (1984): *A Bibliography of Scandinavian Dictionaries*. New York: Kraus International Publications.

Hovdenak, Marit. 2001. Bokmål og nynorsk som kjelde- og målspråk i internordiske ordbøker. *LexicoNordica* 8. Oslo.

For andre kjelder viser eg til Dysvik 2009.

Norrøn metaordbok – en rapport fra en øvelse

The term Old Norse usually denotes the (written) languages used in Iceland until 1540 and in Norway until 1370. Old Norse was a weakly normalised language with a wide variety of forms. The scholarly study of Old Norse has long traditions and several dictionaries have been published. *The Dictionary of Old Norse Prose* (ONP) is a current major initiative but so far covering only letters A-E. J. Fritzner's dictionary, FRI, (1883–1896, 1974) and F. Jónsson's *Lexicon Poeticum*, LP, (1931) still remain the major dictionaries. The dictionary group of the Medieval Nordic Text Archive (Menota) has as one of its main objectives to establish a framework for linking the disparate Old Norse lexicographic resources. The fragmented state corresponds closely to the state of the *New Norwegian* lexicography 10 years ago. Thus it was a natural choice to try to apply the methods and tools to Old Norse developed for the *Norwegian Dictionary 2014* project (NO2014), the so called Meta Dictionary, a tool for systematizing lexicographical material for weakly normalised languages. In its basic form the Meta Dictionary is an electronic form thesaurus linking realizations of the lemmas. In this pilot project we have focused on the letter B and merged the headword lists of ONP, FRI, LP and a word list based on five lemmatised texts and 200 lemmatised charters. The process and the structure of the resulting Meta Dictionary are divided in three levels of normalisation. All the original forms are kept and made searchable.

1. *Normalisation*: The first step towards merging was simply to ensure one common system for normalised spelling.
2. *Citation form*: The wordlists follow different principles in choice of the form of the cited headword, e.g. nom.sg versus nom.pl. Consequently we had to make sure that all the wordlists use the same forms.
3. *Meta level, further merging of related word forms*: The last step is to connect the multitude of related word forms, both between and within lists, in meta lemmas. Such related forms can be nouns differing in gender, or small differences in word formation.

We also briefly discuss the linking of the medieval material to the modern dictionary projects, like NO2014. We conclude that such a linking at least for Norwegian will involve considerable manual processing. The realization of an Old Norse Meta Dictionary should also rely on a close cooperation with the ONP project.

1 Bakgrunn

Norrønt er det språket vi finner i Norge og på Island i middelalderen. Det nedfeller seg i skriftlige kilder fra rundt 1150. I Norge pleier en å regne den norrøne periode til ca. 1370 og på Island til første halvdel av 1500-tallet. Språket var ikke statisk i denne perioden. Der var dialektale forskjeller og det skjedde endringer over tid. Det overleverte materialet viser både ortografisk og språklig variasjon. Norrøn og nordisk filologi har utviklet seg som disiplin gjennom nærmere 200 år. Tekstene i utgaver og ordbøker for nordiske middelalderspråk preges av et stort mangfold med hensyn på normering, antatt leksikalisering og ortografi. Slik sett er det en dobbel varians for norrønt, originaltekstene er svakt normert og metadisiplinen har utviklet seg. Situasjonen er dermed ikke ulik den vi finner for nynorsk skrifttradisjon og leksikografi.

De to store pågående ordbokprosjektene for norrønt og nynorsk, *Ordbog over det norrøne prosasprog* (ONP) og *Norsk Ordbok* (NO), har også mange likhetstrekk. Begge ble grunnlagt på 1930-tallet, og for begge prosjektene var det på slutten av 1990-tallet åpenbart at den daværende redaksjonskapasiteten ville kreve ytterligere flere tiårs redigering før verkene var komplette. For begge prosjektenes vedkommende er det i de senere år blitt foretatt en omfattende digitalisering av kildematerialet. NO har dessuten blitt fullfinansiert med en planlagt ferdigstillelse i 2014.

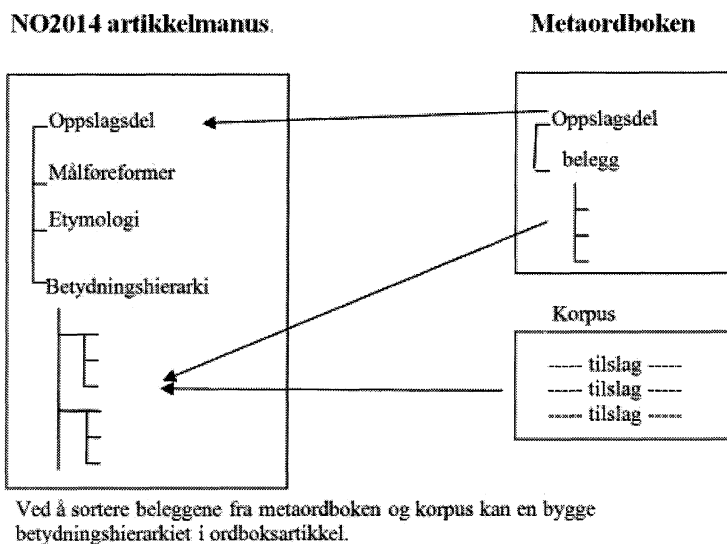
Selv om situasjonen for norrøn leksikografi minner om situasjonen for den nynorske leksikografien på slutten av 1990-tallet, er rammebetingelsene noe annerledes og det organisatoriske atskillig mer komplekst. Norrønt er ikke et levende språk og heller ikke bare knyttet til ett land. Samtidig som norrønt og norrøn tekstutgivelse også er et internasjonalt forskningsfelt, er språket en vesentlig del av Islands og Norges kulturarv.

ONP ble startet på 1930-tallet da Island var underlagt Danmark og må sånn sett sies å være et dansk prosjekt. Norge har siden 1940-tallet hatt et eget Gammelnorsk Ordboksverk (GNO) med det formål å dokumentere alle sider av det gammelnorske språket. GNO har mellom annet utgitt et supplementsbind til J. Fritznors ordbok og arbeider med å få realisert *Mellomnorsk ordbok* som skal beskrive norsk i den perioden ONP ikke dekker, 1370–1580. De leksikografiske hovedverkene for norrønt er fortsatt J. Fritznors *Ordbog over det gamle Norske sprog* (1883–1896, 1972) og F. Jónssons *Lexicon Poeticum* (1931). I tillegg kommer de tre første bindene av ONP. Hele verket vil neppe bli ferdig redigert på flere tiår. Det finnes altså ingen fullstendig og samordnet leksikografisk beskrivelse av norrønt.

I løpet av de siste 15 årene har det likevel skjedd en rekke positive ting. Det har vært en tiltagende produksjon av elektroniske utgaver av nordiske middelalderhåndskrifter, og standardordbøkene er blitt gjort elektronisk tilgjengelig. Ved ONP har en også digitalisert store deler av grunnlagsmaterialet. I 2001 ble det

nordiske nettverkssamarbeidet Medieval Nordic Text Archive (Menota, www.menota.org) etablert. Menota er et nettverk av nordiske forskningsinstitusjoner og -selskaper som har til hensikt å bevare, formidle og utgi middelaldertekster i elektronisk form. Menotahåndboken har gitt oss en utmerket standard for koding av nordiske middelalderhåndskrifter. Samarbeidet om publisering av elektroniske nordiske middelaldertekster har kommet i godt gjenge. Det har kommet på plass et elektronisk tekstarkiv med et tilhørende lemmatiseringsverktøy og et rammeverk for et elektronisk korpus. Menotas gruppe for leksikografi, Nordiske middelalderordbøker, (www.edd.uio.no/menota) ser på mulig samarbeid og deling av leksikalske ressurser mellom de ulike ordboksprosjektene for middelalderspråk i Norden. Arbeidet som er beskrevet i denne artikkelen er gjort i den forbindelse. Vi har i vår øvelse sett på hvordan en kan knytte sammen de leksikalske ressursene i et felles rammeverk for søk i de ulike digitale ordboksressursene og tekstkorpuserne.

2 Hva er norrøn metaordbok?



Figur 1: Norsk Ordboks redigeringsystem, Metaordboken og korpus

Middelalderspråk som norrønt er svakt normerte språk med et stort utvalg av ulike ordformer. For å optimalisere søk og analyser bør en størst mulig del av ordene i de løpende tekstene være lemmatisert og påført bøyninginformasjon. Det finnes ingen utviklet morfologisk analysator (tagger) for norrønt. Som et viktig ledd i å sikre god funksjonalitet trenger korpuset å ha en formtesaurus koblet mot eksisterende ordbøker for norrønt. Denne formtesauren eller metaordboken vil være ryggraden i sammenkoblingen av de leksikalske enkeltdatabasene. Den vil også kunne danne forbindelsen til de moderne ordboksverkene for islandsk og

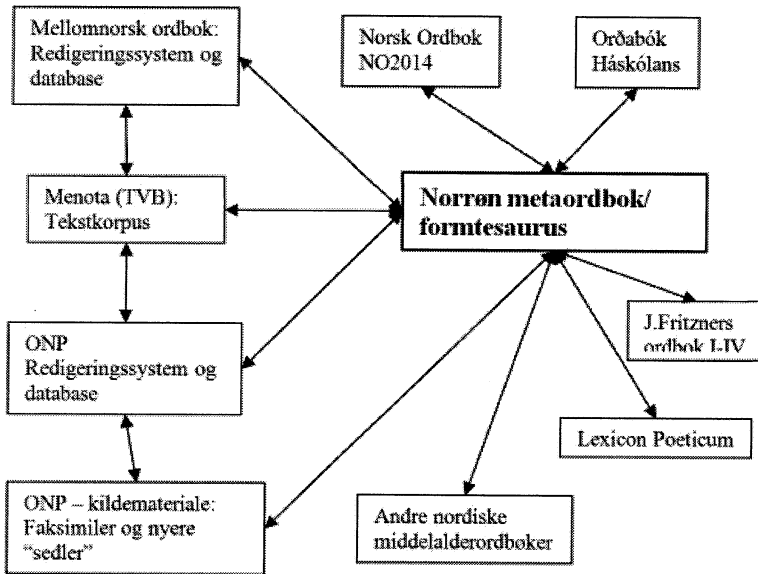
norsk, *Orðabók Háskólans* og *Norsk Ordbok*, som beskriver språkene fra etterreformatorisk tid til i dag. Vi mener at et slikt rammeverk kan bygges over samme lest som Metaordboks-systemet for Norsk Ordbok. Metaordboken er et velprøvd verktøy som har blitt brukt av NO2014 de siste 6 årene. Det er kanskje ikke så velkjent, så det er naturlig å bruke litt plass på å forklare bakgrunnen og hva det går ut på.

Nynorsk leksikografi var på 1990-tallet i samme situasjon som norrøn leksikografi med det unntak at norrøn leksikografi er spredt over flere land og prosjekter. I den nettbaserte virkeligheten spiller det siste en mer organisatorisk enn praktisk rolle. Som et ledd i revitaliseringen av de store norske ordboksprosjektene gjennomførte en på 1990-tallet en storstilt digitalisering av tekstsamlinger, kortarkiv og ordbøker, og det ble lagd systemer for morfologisk merking av løpende tekst. Dette arbeidet har siden fortsatt. Det revitaliserte Norsk Ordbok-prosjektet, NO2014, har nå også et tekstkorpus på mer enn 30 millioner ord. I løpet av siste halvpart av 1990-tallet ble det utviklet en modell for hvordan disse ressursene, både det gamle materialet og det nye elektroniske tekstkorpuset, kunne bindes sammen i et rammeverk for leksikografisk arbeid for å understøtte redigeringsarbeidet, men også for å gjøre materialet tilgjengelig for andre forskere og den interesserte allmennheten. Dette rammeverket ble kalt en *metaordbok* og grunnideene er beskrevet i Ore 1999.

Metaordboken er et verktøy for å samle belegg og normere svakt normerte språk. Hver artikkel i Metaordboken er en samling referanser til kilder som beskriver det aktuelle ordet eller gir brukseksempler. Hver artikkel kan ha et eller flere normerte oppslagsord med tilhørende bøyningsinformasjon. Disse oppslagsordene kan være konstruert i henhold til en gitt rettskrivningsnorm. De brukes som bindeledd til de egentlige ordboksartiklene i Norsk Ordbok. For norrønt vil en kunne ha oppslagsord til meta-artiklene både i henhold til GNOs standard og til ONPs.

Fullt utbygd og korrekt anvendt skal Metaordboken til NO2014 avspeile prosessen med å skape ordboksartikler helt fra ubehandlede samlinger, via omstruktureringer og systematiseringer til en eksplisitt beskrivelse av hva de rå samlingsdataene betyr, se figur 1. I prinsippet skal man ha tilgang til alt materiale som er brukt i redigeringsprosessen, også det som ikke er kommet med i ordboksartikkelen. På den måten vil man komme et steg videre med å gjøre en ordbok vitenskapelig etterprøvable.

Når vi nå har sett på mulighetene av å lage en metaordbok for norrønt, er det uten intensjoner om å lage noen felles nordisk middelalderordbok. Hensikten er å få til et system for å koble sammen alle elektroniske tekster og leksikalske ressurser. Den norrøne metaordboken er ment som et hjelpemiddel både for de norrøne ordboksprosjektene og for tekstdelen av Menota-samarbeidet, se figur 2. Den vil også være en måte for elektroniske versjoner av de to store ordbøkene for nyere tid, *Orðabók Háskólans* og *Norsk Ordbok*, å hekte seg på til det førreformatoriske materialet.



Figur 2: Noen mulige forbindelser gjennom en norrøn metaordbok

3 Grunnlagsmaterialet

Metaordboken kan i sin enkleste form sees som en samindeksering av en rekke kilder og belegg. Vi har foreløpig begrenset oss med visse unntak til denne enkle formen og kun sett på ord som begynner på bokstaven *B* for å finne ut mer om arbeidsmengde og problemer. Kildene vi ønsker å koble er hovedordbøkene for norrønt. Det vil si ONP, J. Fritzners ordbok og F. Jónssons *Lexicon Poeticum*. For å få til en kobling med korpusmaterialet har vi også tatt med en liste over ordforrådet i fem større tekster¹ og 200 norskspråklige diplomer fra før 1308. De elektroniske tekstene er i xml-format og lemmatiseringen er basert på GNOs ordseddelsamlinger. Tekstene svarer til om lag 420 000 løpende ord.

Tabellen i figur 3 gir en oversikt over forholdet mellom løpende ord (tokens), forskjellige ordformer (types på formnivå) og grunnord (types) på lemmanivå i de fem tekstene og diplomene. I snitt utgjør lemma på bokstaven *B* om lag 6,2 prosent av hele lemmautvalget, mens på token-nivå ligger *B*-ordformer nede på 3,3 prosent. Det skyldes antagelig at det ikke finnes så mange høyfrekvente funksjonsord på bokstaven *B*. Men det er verdt å legge merke til at det relative antallet oppslagsord på bokstaven *B* i de tre ordbøkene er i godt samsvar med den faktiske fordelingen i tekstutvalget. For ONP må spesielt totaltallene tas med en klype salt. De er hentet fra nettordlisten. Tallene i raden markert med "hovedord" er

¹ Konungs skuggsjá, AM 243 b α fol; Óláfs saga ins helga, Upps DG 8; Strengleikar, Upps DG 4–7; Þiðriks saga af Bern, Holm perg 4 fol; Thómass saga erkibiskups, Holm perg 17 4to.

de ord som ikke er markert som poetisk, navn, fremmedord, eller yngre islandsk eller norsk. De aller fleste oppslagsordene i *Lexicon Poeticum* er derfor ikke med i ONPs hovedordliste slik den fremstår på nettet. Tilsvarende er ikke navn eller latinske ord i tekstene med der. Yngre norske og islandske ord i Fritzners ordbok er heller ikke med i ONPs hovedordliste.

Tekst	Forskjellige grunnord			Forskjellige ordformer			Løpende ordformer		
	B	alle	%	B	alle	%	B	alle	%
Thóm.	186	3 709	5,0	699	14 284	4,9	2 430	64 380	3,8
Kgs.	280	4 116	6,8	951	14 194	6,7	1 780	56 226	3,2
Þiðr.	235	4 210	5,6	1 305	24 334	5,4	4 139	147 766	2,8
Str.	147	2 597	5,7	487	9 889	4,9	1 100	42 085	2,6
Óláf.	183	3 234	5,7	603	9 738	6,2	1 218	41 060	3,0
Dipl.	312	4 806	6,5	1 235	20 111	6,1	3 050	66 175	4,6
Allc	841	13 598	6,2	4 544	75 567	6,0	13 717	417 692	3,3

Ordbok	B	Hele	%
Fritzner, artikler	2 400	38 500	6,2 %
Lexicon Poeticum, artikler	1 150	1 8000	6,4 %
ONP, ordliste, hovedord	3 900	72 000	5,4 %
ONP, ordliste, alle ord (antall artikler a-em)	5 600	91 700	6,1 %

Figur 3: Lemma- og ordformfrekvenser i de ulike tekstene og i de to hovedordbøkene for norrønt.

4 En tre-trinns prosess for samkjøring av ordlister

Et konkret mål for vår øvelse var å samkjøre de fire ordlistene slik at like lemma i hver av listene ble koblet sammen i en formtesaurus. For å gjennomføre dette satte vi opp følgende tre-trinns prosess:

1. felles normalisering
2. valg av oppslagsform og grammatisk klassifisering
3. metanivå, en videre sammenslåing av relaterte ordformer

4.1 Trinn 1: normalisering

Normalisering innenfor norrønt er konstruert i nyere tid etter mønster av islandsk ortografi. Det finnes flere normaliseringsvarianter. Ingen av de fire ordlistene våre følger nøyaktig samme prinsipper for normalisering. En oversikt over vesentlige forskjeller gis i figur 4.

Det kan være snakk om rent ortografiske ulikheter, som at ONP gjengir lang 'ø' med aksent for markering av vokallengde, mens de andre listene bruker 'oe'-ligatur. Ulik normaliseringspraksis kan også gjenspeile geografisk variasjon, som

når GNOs ordformer viser gammelnorske trekk ved at 'h' ikke står foran 'l', 'n' og 'r' i framlyd.

Fenomen	ONP	GNO	Fritzner	LP
æ	æ	æ	æ	æ
langø	ø	œ	œ	œ
kort ø	ø	ø	ö	ø
o med kvist	o	o	ö	o
lang o med kvist	á	á	á	ó
øy / ey	ey	øy	ey	ey
konsonantisk i	j	i	j	j
pt / ft	ft	pt / (ft)	pt / (ft)	pt
lang K + ny K	ggð, lld	gð, ld	gð, ld	ggð, ld
hl, hn, hr	hl, hn, hr	l, n, r	hl, hn, hr	hl, hn, hr

Figur 4: Normaliseringsforskjeller mellom ordlistene

I tråd med Menotas prinsipper valgte vi å følge ONPs standard for normalisering. Mye av samnormaliseringen kunne utføres ved kontrollerte søk-erstatt rutiner. Men enkelte ulikheter mellom listene måtte lukes ut manuelt. Det siste gjaldt f.eks. for ordlistenes ulike realisering av vokalen 'o med kvist'. Dette var en lav, bakre, rundet vokal som oppstod ved overgangen fra urnordisk til norrønt. Den ble ikke værende i språket. I islandsk falt den etterhvert sammen med kort 'ø' og i norsk hovedsakelig med kort 'o'. ONP og LP bruker et entydig spesialtegn for 'o med kvist'. Fritzner følger nyere islandsk ortografi med bokstaven <ö>, mens GNO bruker bokstaven <o>. Fritzner har dermed sammenfall mellom etymologisk kort 'ø' og kort 'o med kvist', mens GNO har sammenfall mellom etymologisk kort 'o' og kort 'o med kvist'. Denne typen variasjon i listene måtte lukes ut manuelt og føres over til ONPs standard.

Til hjelp i arbeidet med å identifisere avvikende ordformer, foretok vi jevnlig flettinger av ordlistene og kjørte ut *diskrepanslister* over de ordformene i Fritzner, GNO og LP som ikke falt sammen med noen ordform i ONP. Ved fletting fulgte vi følgende prinsipper: GNO-listen er ikke homografseparert, så vi skiller ikke mellom homografer av samme ordklasse. I ordlistene er forledd og etterledd egne innførsler, så vi har valgt å skille mellom oppslagsord markert som selvstendige ord og oppslagsord markert som for- eller etterledd. Former som *bakki* m. og *-bakki* m. ble dermed ikke koblet sammen.

Ved samkjøring av ordformer beholdt vi de ulike listenes opprinnelige oppslagsform. For hvert trinn i prosessen, ble det i hver liste lagt til et ekstra felt for lemma og grammatisk kategori, og alle endringer på trinnet ble foretatt i de nye feltene. Figur 5 viser trinn-1 felt etter samnormalisering av ordet *banahogg*. Listenes opprinnelige form står i feltene til høyre.²

² For feltet 'Gram-1' i tabellen i figur 5 har vi ikke fulgt ONPs notasjon. Vi vil etterhvert følge Menotas grammatiske notasjon som baserer seg på anbefalinger fra Eagles (<http://www.ilc.cnr.it/EAGLES96/annotate/annotate.html>). I tabellen er det brukt en midlertidig arbeidsnotasjon.

Liste_ID	Trinn-1	Gram-1	Oppslagsord	Gram
GNO	banahogg	n	banahogg	n
ONP	banahogg	n	banahogg	sb. n
Fritzner	banahogg	n	banahogg	n

Figur 5: Trinn-1 felt for banahogg med 'o med kvist' ført inn.

Etter fullført samnormalisering, trinn 1, kjørte vi ut nye diskrepanslister og regnet ut avviksprosent for hver av listene i forhold til ONP. Av Fritznerns totalt ca. 2400 ord på B, kom det 180 ord på diskrepanslisten, hvilket tilsvarer avvik på 7,5 %. For LP og GNO var avviket på henholdsvis 5 % og 16 %. Vi kom dermed ganske langt bare ved å samnormalisere listene.

4.2 Trinn 2: valg av oppslagsform og grammatisk klassifisering

4.2.1 Ulike prinsipper for valg av oppslagsform

De ordene vi fikk ut på diskrepanslisten etter trinn 1, falt i en del klart definerbare kategorier som igjen har sammenheng med ulike prinsipper for valg av oppslagsform. De mest frekvente av disse er vist i figur 6. Eksemplene her er hentet fra Fritznerns avvik i forhold til ONP, men de samme kategoriene går igjen også i de andre listene.

Fritzner		ONP		
Substantiv nom. pl.	<i>barir</i> <i>bróðrasynir</i>	Substantiv nom. sg.	<i>bör</i> <i>bróðrasonr</i>	'(lik)båre' 'brødresønner'
Verb refleksiv	<i>bakverpast</i> <i>baðast</i>	Verb infinitiv	<i>bakverpa</i> <i>baða</i>	'forkaste - vende ryggen til' 'bade'
Adverb	<i>berbakt</i> <i>bátlaust</i>	Adjektiv	<i>berbakt</i> <i>bátlauss</i>	'med bar rygg' 'uten båt'
Adjektiv (=perf.part.)	<i>borinn</i>	Verb infinitiv	<i>bera</i>	'bære'

Figur 6: Ulike prinsipper for valg av oppslagsform

Fritzner har ført opp flertallsformen av substantiv som han i kildematerialet utelukkende har funnet belagt i flertall, mens ONP har ført opp entallsformen. Enten har ONP funnet belegg for entallsformen, hvilket vil framgå av ordboksartikkelen, eller også er entallsformen konstruert. Fritzner har likeledes ført opp refleksivform av verb han ikke har funnet belagt i aktiv, mens ONP har ført opp infinitiven av verbet, enten de har funnet belegg for bruk i aktiv eller ikke. De to verbene i figur 6 er forøvrig begge funnet belagt i aktiv. Tilsvarende har Fritzner ført opp adverb der han ikke har adjektivformen belagt, mens ONP har ført opp adjektivet, enten det er en reell eller konstruert form.

Det er klart det kan diskuteres hvorvidt man skal konstruere oppslagsformer

som kanskje aldri har eksistert. Men vi hadde som overordnet prinsipp å følge ONP, og da ble det å identifisere denne typen diskrepans og føre den over til ONPs valg av oppslagsform, en relativt enkel oppgave.

4.2.2 Adjektiviske partisipper

Et eksempel på en mer problematisk samkjøring av ordformer på dette trinnet, er adjektiviske partisipper. I ordlistene finner vi blant annet følgende oppslagsformer med utgangspunkt i verbet 'bære':

Fritzner:	<i>bera</i> v.	<i>borinn</i> adj.	<i>arfborinn</i> adj.
ONP:	<i>bera</i> v.	<i>-borinn</i> adj.	<i>arfborinn</i> adj.

Hos Fritzner er slike partisippformer ført opp som selvstendige lemma og klassifisert som adjektiv. I ONP er de samme partisippformene behandlet under artikkelen til verbet. Ut fra vårt prinsipp om å følge ONPs valg av oppslagsform, burde vi føre Fritzners adjektivpartisipper til verbets infinitiv, dvs. *borinn* adj. føres til lemma *bera* v. Men dersom partisippformen inngår i en fast sammensetning i norrønt, har både ONP og Fritzner klassifisert sammensetningen som adjektiv, f.eks. *arfborinn* adj, 'arveberettiget'. I ONP er etterleddet i tillegg ført opp som adjektiv, med henvisning til verbet. Siden vi har valgt å skille mellom selvstendige ord og ledd, førte dette til følgende to lemma på trinn 2, med tilknyttede ordformer:

Lemma 1, <i>bera</i> v.:	<i>bera</i> v. (ONP)
	<i>bera</i> v. (Fri)
	<i>borinn</i> adj. (Fri)

Lemma 2, <i>-borinn</i> adj.:	<i>-borinn</i> adj. (ONP)
-------------------------------	---------------------------

En kompliserende faktor her er at Fritzner ikke nødvendigvis opplyser at en oppslagsform utelukkende opptrer som del av en sammensetning. Vi kan derfor gå ut fra at flere av de adjektiviske partisippene i Fritzner i realiteten er ledd og dermed identisk med adjektivleddene i ONP av typen *-borinn* adj. Men det vil neppe gjelde alle partisippformene i Fritzner, og det virket lite gunstig å behandle adjektivpartisippene ulikt på dette stadiet i prosessen. Derfor valgte vi på trinn 2 å føre alle slike partisipper hos Fritzner til verbets infinitiv.

4.2.3 Oppsummering trinn 2

Diskrepanclistene vi kjørte ut avslutningsvis på trinn 2, viste et ganske oppløftende resultat, særlig for Fritzners vedkommende. Det var nå bare 0,6 % av ordformene i Fritzner som vi ikke hadde klart å føre over til en form i ONP. GNO hadde en noe høyere avviksprosent på 12 %. Dette skyldes hovedsakelig at GNO-listen inneholder mange navn og navneledd, og slike former er ikke tatt med i ONP.

4.3 Trinn 3, en videre sammenslåing av relaterte ordformer

4.3.1 Relaterte ordformer

Etter fullført trinn 2 har vi etablert en formtesaurus som tilsvarende ONPs lemmaliste pluss særformer fra de øvrige listene. ONP er den mest omfangsrike av de norrøne ordlistene, blant annet fordi den har med langt flere variantformer i form av nært relaterte ordformer der de andre listene økonomiserer ordutvalget, se figur 7. En slik ordrikdom er positiv sett fra en bruker av ordbokens side. Når man skal slå opp et ord man støter på i en norrøn tekst, vil man stort sett finne det i ONP. Men denne overflod av nært relaterte ordformer er ikke nødvendigvis like gunstig i en metaordbok, f.eks. med tanke på søk i korpus, eller generering av statistikker over ordforekomster. Dersom man søker etter *barnafaðir*, kan det være av interesse også å få informasjon om relaterte former som *barnfaðir* og *barnsfaðir*. Den ene formen bør ikke utelukke de to andre. Det fører oss til trinn 3 i samkjøringsprosessen, hvor vi går ut over ONPs lemmaliste og slår sammen nært relaterte ord i metalemma, som de tre sammensetningene for 'barnefar' i figur 7. Disse sammensetningene er behandlet i separate oppslag i ONP. Fritzner har kun formen med nullfuge, *barnafaðir*. Ordet forekommer ikke i GNOs materiale eller i LP.

Fenomen	ONP	Fritzner
Substantiv, sterk og svak form	<i>bjartleikr</i> sterk m. <i>bjartleiki</i> svak m. <i>bjarnfell</i> sterk n. <i>bjarnfelli</i> svak n.	<i>bjartleikr</i> sterk m. = 'lys, glans' <i>bjarnfell</i> sterk n. = 'bjørneskinn'
Substantiv med variasjon i grammatisk kjønn	<i>bið</i> f. <i>bið</i> n.	<i>bið</i> n.pl. = 'venting, ventetid'
Svake adjektiv på -a og -i	<i>blóðrisa</i> <i>blóðrisi</i>	<i>Blóðrisa</i> = 'blodsprenget'
Sammensetninger med morfologisk variasjon i 1. ledd	<i>barnafaðir</i> m. <i>barnafaðir</i> m. <i>barnsfaðir</i> m.	<i>barnafaðir</i> m. = 'barnefar'

Figur 7: Eksempler på typer av variasjon i ONP utelatt hos Fritzner

Vi trengte imidlertid kjøreregler for en slik grovere sammenslåing av ordformer. Vi var i den heldige situasjon for bokstaven B at ONP allerede har behandlet den. I ONPs artikler henvises det alltid mellom beslektede ordformer. ONP er dessuten digitalisert og gjort tilgjengelig via internett. Vi kunne derfor hente disse krysshenvisingene fra nettutgaven og generere lister over relaterte former. Men siden ONP ikke skiller mellom henvisninger som angir semantisk slektskap og de som angir formvarians, måtte det en manuell gjennomgang til.

Valg av metalemma har vært gjenstand for en del diskusjon, og vi har til dels også eksperimentert med mulige tilnæringsmåter. Vi kunne velge den etymologisk mest korrekte formen, eller den mest frekvente formen, eller en bestemt fuge i sammensetninger. Men disse framgangsmåtene var ofte svært tidkrevende og passet som regel aldri for alle typer av variasjon mellom relaterte ordformer. I dette prøveprosjektet var det også et poeng å få dataene samkjørt relativt raskt slik at vi kunne få et overblikk over materialet. Derfor valgte vi en enklere og kanskje noe uortodoks metode: Metalemma er den formen som kommer først alfabetisk. Dette er en metode som lar seg klart definere, den er etterprøvbar og kan utøves på alle ordformene. Det må også understrekes at noe av ideen i metaordboken er at artiklene kan gis flere metaoppslag, det vil si kan indekseres etter flere ordlister.

4.3.2 Sammensetninger med variasjon i fugen

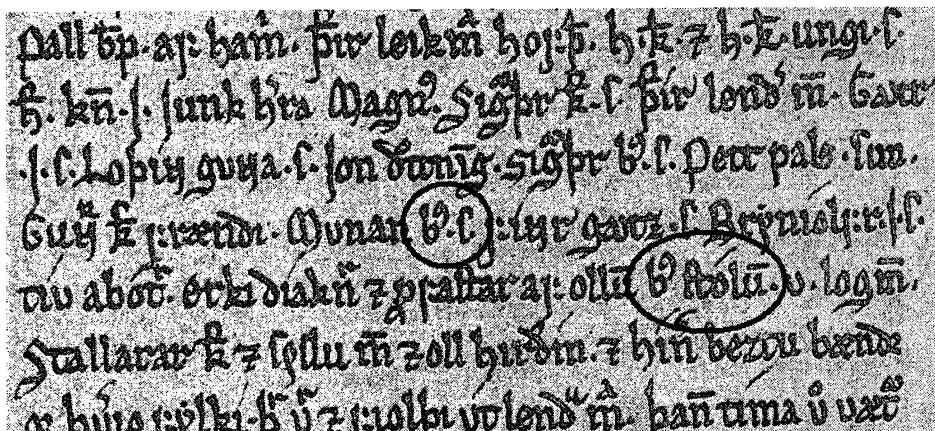
For sammensetningers vedkommende vil den alfabetiske tilnæringsmåten føre til at form med g.pl. i førsteleddet, såkalt a-fuge, dersom den er belagt, blir valgt som metalemma:

metalemma:	<i>barnafaðir</i>		
relaterte former:	<i>barnafaðir</i> ,	<i>barnfaðir</i> ,	<i>barnsfaðir</i>

I en ferdig metaordbok vil ordlistenes opprinnelige ordformer, samt alle samkjøringsfeltene fra trinn 1 og trinn 2, være søkbare, ikke bare det normaliserte metalemma. Det innebærer at et søk på en form som *barnfaðir*, med nullfuge, vil gi direkte treff på Fritznerns og ONPs artikler med denne formen, men også gi tilgang til de to alternative formene i ONP, med a-fuge og s-fuge.

Det faktum at Fritznern i stor grad slår sammen denne typen ordformer, er et argument for at det er forsvarlig også for oss å gjøre det i metaordboken. Et annet argument i samme retning er at ordbøkene, særlig de eldre, dels ekserperer sine ordbelegg fra trykte tekstutgaver av håndskriftene. I middelalderen praktiserte man en utstrakt bruk av forkortede skrivemåter. I trykte utgaver av håndskriftene vil utgiver ofte løse opp forkortelsene. I nyere diplomatariske, dvs. bokstaverette, tekstutgaver er slike oppløste forkortelser som regel tydelig markert. I eldre diplomatariske tekstutgaver er ikke dette nødvendigvis tilfelle, og man kan som bruker av teksten ikke vite hva som er originalens skrivemåte og hva som står for utgivers regning. Som et eksempel kan vi se på følgende skrivemåter for *byskup(s)son* og *byskup(s)stólum* i håndskriftet *Frisbók*³ fra begynnelsen av 1300-tallet, figur nr. 8.

³ AM 45 fol, Codex Frisianus, ca 1300–1325, her gjengitt etter Halldór Hermannsson 1932.



Figur 8: Frisbók s. 123v spalte 452 l. 12 *byskup(s)son* og l. 13 *byskup(s)stólum*. Understreking markerer oppløste forkortelser.

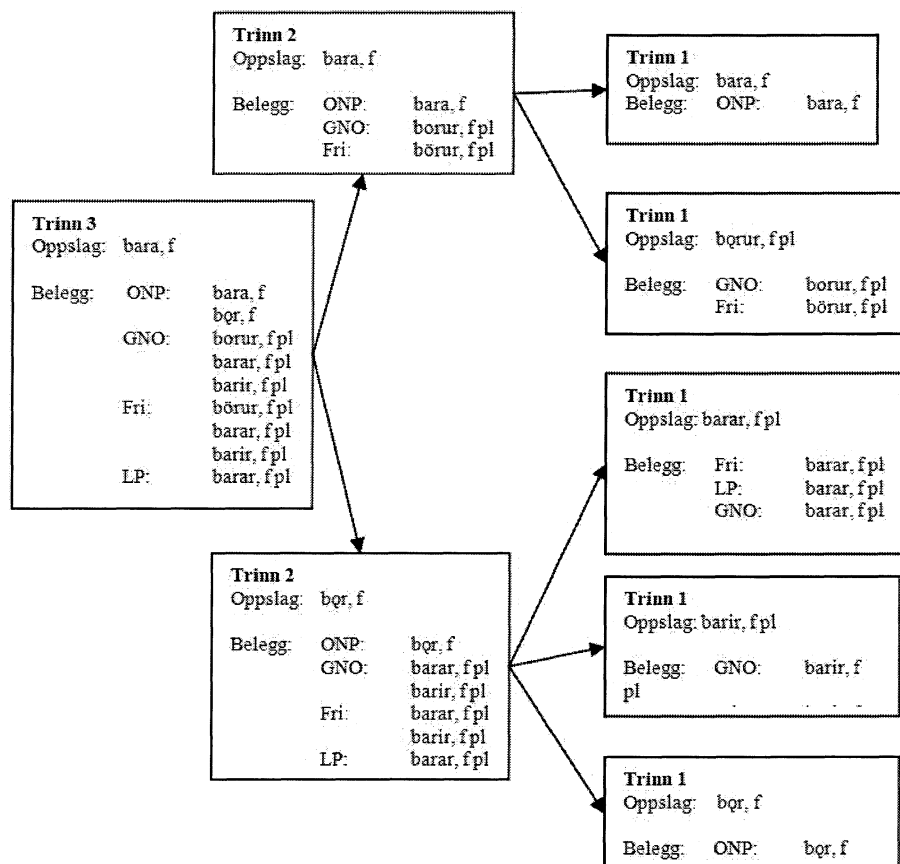
Teksten er her såpass forkortet at vi ikke ut fra skrivemåten kan se hvilken fuge skriveren mon hadde i tankene. I tekstutgaven av Frisbók, Unger 1871, er disse forkortelsene oppløst uten markering som henholdsvis *byskupsson*, med s-fuge, og *byskupstólum*, med nullfuge. Utgiver følger det han oppfatter som håndskriftets norm ved oppløsning av forkortelser. ONP er med alle sine ordformer en deskriptiv ordbok, mens Fritzner kan sies å være en mer normativ ordbok. Men grunnlaget for beleggene i ONPs deskriptive samling av ordformer er i enkelte tilfeller allerede normert i og med utgivers valg av oppløsning. Metaordboken viderefører denne normeringstendensen i den grovere sammenslåingen av ordformer i trinn 3.

4.3.3 Metalemma og relaterte former; et eksempel

I det norrøne kildematerialet finner vi flere relaterte ordformer for det som tilsvarer vårt '(lik)båre'. Begrepet forekommer hovedsakelig i flertall, med tre belagte flertallsformer, *barar*, *barir* og *þorur*. Disse finner vi som tre separate innførsler i GNO og Fritzner, mens LP bare har formen *barar*. I ONP er de to første formene, *barar* og *barir*, begge ført til en konstruert entallsform *þor f*, mens den siste flertallsformen, *þorur*, er ført til en belagt entallsform *bara f*. Dette viser at ONPs tendens til å ta med et stort antall relaterte ordformer til en viss grad dempes noe av ordbokens strengere regler for valg av oppslagsform.

ONPs to oppslagsord er en svak og en sterk form av et feminint substantiv. Det er flere paradigmer knyttet til den sterke formen (nom. pl. *barir* og *barar*), og det er også sammenfall i enkelte bøyingsformer mellom sterk og svak form (f.eks. dat.pl. *þorum*.) ONP krysshensliser mellom de to oppslagsordene. På et metanivå kan vi se ONPs to oppslagsord som varianter av samme lemma, med ulike paradigmer knyttet til seg. På trinn 3 blir alle de ni innførslene for 'båre' ført sammen i én metaordboksartikkel, se figur 9. I metaordboken vil hver ordform

på hvert trinn i prosessen være søkbar og gi tilgang til alle formene tilknyttet metaartikkelen.



Figur 9: Metalemma bara f. og relaterte ordformer i listene.

5 Noen betraktninger om linjene fram til de moderne ordbøkene

Den norrøne metaordboken skal være en felles inngang til ordbøkene over islandsk og norsk i norrøn tid ved å gruppere sammen ulike skriftrealiseringer av det en ønsker å oppfatte som ord eller lemma. Norrøn metaordbok vil gjennom de ulike ordbøkene koble sammen ordbelegg over en periode på rundt 500 år. Til sammenligning dekker *Norsk Ordbok* norske dialekter i drøye 300 år og *Orðabók Háskólans* islandsk i 460 år. Det er lange spenn, men ikke oppsiktsvekkende lange. *Oxford English Dictionary* dekker engelsk over en 1000 års periode.

For de norske historisk-orienterte leksikografiske miljøene er en norrøn metaordbok interessant fordi den vil danne grunnlag for arbeidet med en Mellomnorsk ordbok for perioden 1370–1580 og kunne brukes til å lage en forbindelse

mellom *Norsk Ordboks* beskrivelse av tidlig-moderne og moderne norsk. Det er i så fall et spørsmål hvordan en slik forbindelse skal etableres. I de etablerte norske ordbøkene er det relativt få ord som er markert med gammelnorsk eller gammelislandsk opphav. A. Torps nynorske etymologiske ordbok behandler om lag 15 000 stort sett usammensatte ord. Bare 5 000 av dem har henvisning til norrønt eller gammelislandsk. For bokstaven B har de tre nynorske ordbøkene, *Norsk Ordbok*, A. Torps etymologiske ordbok og *Nynorskordboka*, tilsammen drøye 900 ord med slik henvisning. *Norsk Ordbok* har 18 400 oppslagsord på bokstaven B, mens håndordboken *Nynorskordboka* har 7200. Det er pussig om bare fem prosent av nynorske ord skulle ha norrønt opphav. Motsatt vei er det maksimalt 20 prosent av oppslagsordene i ONP og maksimalt 40 % av oppslagsordene i Fritzner som er referert fra de moderne norske ordbøkene. Dette er kanskje mer naturlig, men er verd et nærmere studium. Denne raske kvantitative testen sier oss at en etter all sannsynlighet må utføre et omfattende leksikografisk arbeid for å få på plass en ordentlig kobling mellom *Norsk Ordbok* og det norrøne materialet. Det foreligger ikke nok ordhistorisk informasjon i de eksisterende ordbøkene til å gjøre en særlig god mekanisk kobling.

6 Oppsummering og videre arbeid

Både den alfabetiske tilnæringsmåten og bruk av krysshenvisingene fra ONP, er nyttige hjelpemidler i arbeidet med å føre sammen relaterte ordformer på trinn 3, men de er ikke absolutte prinsipper. De genererte listene over krysshenvisinger bør derfor brukes til manuell gjennomgang av de i ONP relaterte formene.

Den alfabetisk utvalgte metalemmalisten kan tilsvarende sees som et utgangspunkt for videre bearbeiding. I de digitaliserte ordlistene er hovedform og sideform splittet i separate innførsler for å muliggjøre søk på alle former. Ved feminine på *-an* og *-un* (eks: *skipan*, *skipun*) fører ONP formen med u-vokalisme i endelsen som hovedform, mens den alfabetiske tilnæringsmåten vil velge ut formen med a-vokalisme. En manuell gjennomgang av en alfabetisk generert metalemmaliste er derfor både ønskelig og nødvendig. Men i og med at alle ordformer på alle trinn i samkjøringsprosessen vil være søkbare, er valg av metalemma mindre viktig enn valg av former knyttet til dette metalemma. Det vesentlige er at både *skipan* og *skipun* knyttes sammen, ikke hva vi velger å kalle enheten på toppen. Gjennomgang og redigering av den alfabetisk genererte metalemmalisten kan derfor gjøres på et senere stadium.

Sammenslåingen i trinn 3 har i vårt prøveprosjekt basert seg på informasjon som finnes i ONPs ferdigredigerte artikler på nettet. Disse finnes så langt ONP er kommet med artikkelredigeringen, altså til midten av bokstaven E. For resten av alfabetet er der ingen krysshenvisinger å hente ut fra nettversjonen. I ONPs database for resten av alfabetet er slike krysshenvisinger bare satt inn sporadisk. Det vil si at her må materialet gjennomgås manuelt for å identifisere tilsvarende

relasjoner mellom ord, noe som vil bli en langt større og mer ressurskrevende jobb enn vi har utført i dette prøveprosjektet. Det vil også være en forutsetning at dette kan gjøres i samarbeid med ONPs redaksjon.

Litteraturliste og nettsteder

- Halldór Hermannsson 1932 (ed. facs.): Codex Frisianus (Sagas of the Kings of Norway): MS. No. 45 Fol. in the Arnarnagnæan Collection in the University Library of Copenhagen, *Corpus codicum Islandicorum medii aevi* 4, Copenhagen 1932.
- Jónsson, Finnur 1931: Lexicon Poeticum antiquæ linguæ septentrionalis, Ordbog over det norsk-islandske skjaldesprog oprindelig forfattet av Sveinbjörn Egilsson, S.L.Møllers bogtrykkeri, København, 1931.
- Fritzner, Johan 1886–1896: Ordbog over det gamle norske Sprog I–III, 4. utgave Universitetsforlaget, Oslo, Bergen Tromsø. 1972, (www.dokpro.uio.no, sjekket 270907).
- Hødnebo, Finn 1972: Rettelser og tillegg til Johan Fritzner, Ordbog over det gamle norske Sprog, Universitetsforlaget, 1972, Oslo.
- Menota – Medieval Nordic Text Archive (www.menota.org, sjekket 270907)
- Norrøn metaordbok www.edd.uio.no/menota/metaordbok.html
- Nynorskorboka, tredje utgave, Det norske samlaget, Oslo 2001.
- Norsk Ordbok 1–6, Det norske samlaget, Oslo 1950–2007, (www.no2014.uio.no (sjekket 270907)).
- Ordbog over det norrøne prosasprog, 1–3, København, www.onp.kua.dk (sjekket 270907).*
- Orðabók Háskólans (www.lexis.hi.is/indexny.html, sjekket 270907).
- Ore, Christian-Emil 1999: Metaordboken — et rammeverk for Norsk ordbok? Nordiske studier i leksikografi 6, NFL, Göteborg, 1999.
- Torp, Alf 1919: Nynorsk etymologisk ordbok, Kristiania 1919.
- Unger, Carl Richard 1871 (ed.): Codex Frisianus: En Samling af norske Konge-Sagaer, Christiania, 1871.
- Font brukt for norrøne tegn: *Andron Scriptor Web*, designet av Andreas Stötzner, Leipzig, og distribuert av Medieval Unicode Font Initiative (Mufi) www.mufi.info/fonts (sjekket 270907).

Mette Gismerøy Ekker
f. 1963
Kjeldeskriftavdelingen, Norsk Riksarkiv
mettegek@ulrik.uio.no

Christian-Emil Ore
førsteamanuensis, f. 1957
Institutt for lingvistiske og nordiske studier/EDD
Humanistisk fakultet
Universitet i Oslo, boks 1123 Blindern
NO-0317 Oslo
c.e.s.ore@edd.uio.no

Oddrun Grønvik

Redigeringsprogrammet for Norsk Ordbok 2014

I skjæringspunktet mellom menneskeleg skjønn og automatisering

Norsk Ordbok is to be completed and published by the end of 2014. To meet this demand, the eight remaining volumes are written into an editor programme and stored in a database. All information categories in the dictionary therefore need to be reworked, redefined and tightened up in content, and organised more tightly in relation to other categories, than what is required in a traditional manuscript. This article presents a survey of the process of recreating the dictionary structure, a process in which one must weigh the logical against the practicable, and decide which conventions and procedures can be automated through the editor programme, and which decisions should be left to the human editor. The main premise is that the character of Norsk Ordbok must not be altered. This major dictionary shall, when completed, give a united view of Norwegian dialects and Nynorsk standard language.

Det redigeringsprogrammet for Norsk Ordbok som no har teke form, byggjer på arbeid og drøftingar som redaksjonen for *Norsk Ordbok* har gjort i lag med Eining for digital dokumentasjon (DOK) ved Det historisk-filosofiske fakultetet ved Universitetet i Oslo. Arbeidet har vore mest intensivt i den redigeringsprogram-gruppa som eg er leiar for, men alle som medverkar i prosjektet har på ulikt vis og i ulik grad vore med. Arbeidet med redigeringsprogrammet for *Norsk Ordbok* byggjer også på det tidlegare samarbeidet mellom Seksjon for leksikografi og målføregranskning ved INL og DOK om eit databaseprogram for redigering av *Nynorskordboka*. Dette programmet var ferdigstilt våren 2002 og har verka etter føresetnadene.

Det skal også nemnast at utan den røynsla som ALLEX-prosjektet (NUFU 28/96 og 18/2002)¹ har tilført fagmiljøa våre, ville arbeidet med redigeringsprogrammet vore langt mindre effektivt.

1 Arbeidsdeling

Arbeidsdelinga mellom partane har vore slik:

- Redaksjonen for Norsk Ordbok har hatt ansvar for å definere format for redigeringsprogrammet.
- DOK lagar all programvare som Norsk Ordbok treng, der redigeringsprogrammet er eit berande element.

ALLEX-prosjektet har aldri vore direkte knytt til prosjekt NO 2014, men har i perioden 1992-2002 fungert som testverkstad for idear og eksperiment med ordboksdatabasar som byggjer på digitalisert grunnlagstilfang.

2 To overordna omsyn

Det finst ingen nøytrale løysingar på praktiske problem. Ein må avgjere på førehand kva for omsyn som skal vege tyngst når ein vel løysingar.

Norsk Ordbok høyrer i utgangspunktet til den store tradisjonen med vitskaplege, dokumentariske ordbøker. Dette er ramma for ordboksverket også i framtida, trass i påskunding, rekruttering og auka produksjonstakt.

Vidare skal alt tilfang vegast etter dei same prinsippa gjennom heile verket. Desse prinsippa blir utkrystallisert i redigeringsreglar og konvensjonar. Desse reglane og konvensjonane skal altså gjelde likt for alle ord og alt tilfang i samlingane. Stutt sagt:

- Norsk Ordbok 2014 skal vere materialstyrt
- Norsk Ordbok 2014 skal redigerast konsistent

Kva tyder det at Norsk Ordbok skal vere *materialstyrt*?

Det vil seie at ordartikkelen skal gje eit nøyaktig og systematisk bilete av materialet som ligg attom – definisjonar og bruksdøme frå ordbokskjelder og ordsamlingar frå før 1900, målføretilfang som berre finst i samlingane til Norsk Ordbok og prenta målføresamlingar frå alle kantar av landet, litteraturekserpt og det nye korpuset til Norsk Ordbok (sjå Ridings s 315).

Takk vere Dokumentasjonsprosjektet (1991-97) har Norsk Ordbok det meste av tilfanget sitt i elektronisk form, og vi arbeider stadig med å leggje inn attverande ordbøker og ordsamlingar. Tilfanget er ordna gjennom Metaordboka, som gjer det mogleg å gå attom ordartiklane i databasen og sjå på det underliggjande tilfanget (jf. pk. 13).

Kva vil det seie at Norsk Ordbok 2014 skal redigerast *konsistent*?

Det enkle svaret er at dei same prinsipp og konvensjonar skal gjelde for alt tilfang og alle ordartiklar gjennom heile verket. Dette kan høyrast sjølvstakt ut, men er i praksis vanskeleg å gjennomføre, for arbeidsmengda er enorm, og manuell kontroll av alle detaljar uråd å gjennomføre. Det har derfor synt seg – ikkje minst når store, eldre ordboksverk blir digitalisert – at full konsistens har vore meir eit fromt ønske enn realisert praksis.

I Norsk Ordbok 2014 legg vi mest vekt på følgjande presiseringar i høve til konsistenskravet:

1. *Alt tilfang som har forsvarleg kvalitet, skal vurderast for bruk*

Er eit ord såleis pålitande belagt frå gode nynorsk- eller målførekjelder, kan det ikkje avvistas som «dårleg mål» eller liknande. Om ei tyding eller ein bruksmåte er ukjend for redaktørane, skal den likevel med i ordboka dersom tilfanget tilseier det.

2. *Det skal vere mogleg å finne att tilfanget attom kvar definisjon*

Tilfanget vårt ligg ordna i elektronisk form i Metaordboka. Nytt tilfang skal også inn der. Når redigeringsprogrammet blir kopla saman med Metaordboka, blir det mogleg å sjå nøyaktig kva for belegg som ligg attom kvar ordartikkel og kvar tyding.

3. *Tilfang som ikkje gjev grunnlag for artikkel, skal sorterast ut etter same prinsipp*

Redaksjonen har gjennomgått prinsippa for materialutval og fastslege kva for kjelder som

alltid gjev grunnlag for å ta med eit ord, og kva for kjelder som gjev grunnlag for utelating. Metaordboka har opp mot ein halv million artiklar. Norsk Ordbok 2014 har plass til berre om lag 300 000 ordartiklar, fordelte på tolv band. Om lag halvparten av alle metaordboksartiklar har eitt belegg attom seg (er hapax). Dersom dette belegget t.d. er henta frå ei tospråkleg ordbok, ei terminologisk ordliste, eller det er ei poetisk spontansamansetning, får ordet ikkje artikkel i Norsk Ordbok. Desse utvalskriteria kan automatiserast, ettersom dei er knytte til kjelderegisteret. Redaktørane får då sjå på skjermen om eit ord er frårådd eller tilrådd redigert. Redaktørane kan overprøve denne vurderinga ved å leggje til nytt tilfang som dokumenterer ein vidare bruk av ordet.

4. *Tilfanget skal utnyttast etter same faglege kriterium uavhengig av redaktør*

Eit gjennomgåande særdrag i dei gamle, store ordboksverka er at kvar redaktør sette eigarmerke på det han eller ho redigerte. Dette er ikkje til å undrast over; redigering er tekstproduksjon, og tekstproduksjon er ei svært personleg verksemd. Med redigeringsprogrammet får vi likevel betre høve til å følgje med i produksjonen og syte for at redigeringsreglane blir tolka likt. Ein kan også sjå til at alfabetprogresjonen er jamn – dvs. at tilhøvet mellom manusomfang og den plassen eit ord, ei ordgruppe eller ein alfabetbolk har i språket (vonleg spegla i mengda av underlagstilfang) og – er jamn gjennom heile verket, og lik frå redaktør til redaktør. Eit redigeringsprogram gjer det stutt sagt mogleg å føre ein langt meir effektiv produktkontroll enn tidlegare.

3 Tekstypar i dokumentasjonsordbøker

Ei dokumentasjonsordbok har fire slag tekst.

1. *Redaktørtekst* er skriven av ein av ordboksredaktør på grunnlag av ein materialanalyse. Den omfattar definisjonar, redaksjonelle forklaringar m.m. Denne teksten gjer greie for den semantiske strukturen i tilfanget attom kvart oppslagsord, og den avgjer derfor kvaliteten på ordboka som lingvistisk forskararbeid. Denne strukturen vil få ei fastare ramme i band 5-12 enn han har hatt i band 1-4, av di redigeringsprogrammet gjer det mogleg for oss å halde oversyn over ordartiklane på ein annan måte enn før.

2. *Eksempeltekst* er henta frå kjeldetilfanget. Den består av redaksjonelle bruksdøme, bruksdøme med målføreheimfesting, eller autentiske, kjeldefeste litterære sitat. Med digitaliserte kjelder og ein database til manuskriptet blir dette sambandet både tydelegare og lettare å etterprøve enn før. Det skal ikkje brukast døme eller sitat som ein ikkje finn att i underlagstilfanget, og tilfanget kan sporast direkte frå ordartikkelen til dei digitaliserte samlingane.

3. *Tilvisingstekst* er tekst som peikar utover ordartikkelen. I ei dokumentasjonsordbok er det to slag. Det er kryssreferansar i ymse former, som viser til andre oppslagsord eller ordtydingar i Norsk Ordbok. Og det er kjeldeapparatet, som viser ut av sjølve ordboka og til belegg som ligg under ein definisjon, eller inneheld eit sitat. Igjen er det redigeringsprogrammet, og programpakka rundt det, som gjer det langt på veg mogleg å automatisere både føring og kontroll av tilvisingar og kjelder. Dette er ei stor arbeidssparing, ikkje minst i ferdiggjering av manuskript før prenting.

4. *Usynlege belegg* er det tilfanget som ikkje blir gjort synleg i ordboka, men som ligg i samlingane og kan finnast att gjennom Metaordboka. Dette tilfanget er altså lett tilgjengeleg for manuskriptkontroll fram til utgjeving. Vi kjem også til å halde fram med førsteleddsartiklar (med «sekk»), der vi listar opp samansetningar som ikkje kan få eigen ordartikkel (sjå nedanfor pk. 4). Alle slike ord vil ein finne att i samlingane – det skal ligge tilfang attom også dei orda som blir vraka.

4 Artikkeltypar i Norsk Ordbok

Norsk Ordbok har fire ulike artikkeltypar, som kvar vil få sitt grensesnitt for innskriving:

1. *Innhaldsartikkelen* er den prototypiske ordartikkelen, med fullt utbygde skjema for oppslagsdel, variantdel («hakeparentesen») og tydingsdel. Dei andre tre artikkeltypane er til dels sterkt nedskorne utgåver av av innhaldsartikkelen. I redigeringsprogrammet for Norsk Ordbok har vi derfor gjort ferdig skjemaet for innhaldsartikkelen først. Dei fleste ordartiklane er innhaldsartiklar.

2. *Tilvisingsartikkelen* er ein enkel tabell som gjev ei einaste opplysning, nemleg kryssreferanse innanfor Norsk Ordbok frå oppslagsforma på tilvisingsartikkelen til den ordartikkelen der tilfanget som ligg under tilvisingsartikkelen, er redigert.

I Norsk Ordbok går tilvisingar anten frå jamstelt oppslagsform til oppslagsform for innhaldsartikkel, som t.d.

II **gjorda** v sjå *gyrda*.

eller frå målføreform til normalisert oppslagsform, som t.d.

gleka v sjå *glidka*.

I band 1-4 vil ein finne døme på at andre opplysningskategoriar er tekne med i tilvisingsartikkelen, t.d. tydingsopplysning, men det kjem ikkje til å halde fram i band 5-12.

3. *Førsteleddsartikkelen* er også ei redusert utgåve av innhaldsartikkelen. Førsteleddsartikkelen skal gjere greie for oppslagsform og varianttilhøve for to ordelement som ligg nær kvarandre, nemleg (1) avleiingsprefiks og (2) førsteledd i samansetningar. Vidare skal førsteleddsartikkelen gjere greie for korleis førsteleddet spesifiserer etterleddet, og gje døme på avleiingar eller samansetningar som syner fram den aktuelle bruken. Det er derfor høve til å ha fleire tydingar under førsteleddsartikkelen. Derimot kan det ikkje førast sitat eller kjelder. Eksempeltilfanget er jo einskindord, og for samansetningane i «sekken» dreiar det seg om semantisk gjennomsynlege døme som er belagde i samlingane, men ikkje treng eigen ordartikkel. Redigeringsprogrammet skal varsle dersom nokon prøver å lage innhaldsartikkel på eit ord som står i ei førsteleddsrekkje.

4. *Etterleddsartikkelen* gjer greie for form, tyding og bruk av avleiingssuffiks som er nytta i norsk. Avleiingssuffiksa liknar prefiksa i det at grensa til sjølvstendige ord som fungerer som etterledd i samansetningar, kan vere hårfin. Dømetilfanget er då oppslagsord som eksemplifiserer den aktuelle tydinga av etterleddet. Her let vi programmet godta både dømeord som

har eigen innhaldsartikkel og ord som ikkje har det. Men dømmeorda må vere belagde i språk-samlingane.

Samansetningar som ikkje blir redigerte, kan også behandlast innanfor tydingsbolken, jf. pk. 12.

5 Innhaldsartikkelen i NO

Ein (nummerert) definisjon med tilhøyr (bruksdøme, undertydingar, underoppslag, eksemplifiserande samansetningar; meir om dette nedanfor) blir i NO kalla ein *tydingsbolk*². Innhaldsartikkelen i Norsk Ordbok er sett opp som ein struktur av tydingsbolkar, som då skal syne fram den semantiske trestrukturen til kvart oppslagsord. Innhaldsartikkelen har tre nummererte tydingsnivå, markerte i tekst etter rang med

A

1

a

Skjemaet for tydingsbolken blir utnytta fullt ut på lågaste nivå, anten den aktuelle ordartikkelen har ei eller fleire tydingar og eitt eller fleire nivå.

I tillegg til definisjonsdel med døme, kan den fullstendige tydingsbolken (på lågaste nivå) også ha (1) ei eller fleire *undertydingar* (sjå pk. 10); (2) eitt eller fleire fleiordige *underoppslag* (som kan ha fleire tydingar) (sjå pk. 11); og (3) opp til to *samansetningsrekkjer*, ei som syner oppslagsordet nytta som etterledd, og ei som syner ordet nytta som førsteledd (sjå pk. 12).

6 Tydingsbolken

Den komplette tydingsbolken har dei elementa som trengst for å

- definere og eksemplifisere ei tyding (definisjonsskjema og eksemplifiseringsskjema)
- spesifisere tydinga med undertydingar (undertydingsskjema)
- syne fram og definere leksikaliserte ordsamband som underoppslag (underoppslags-skjema)
- vise korleis ordet i nettopp denne tydinga inngår i samansetningar (skjema for samansetningsrekkjer)

I utgangspunktet kan tydingsbolken utnyttast fullt ut på alle nivå, men i praksis vil det ikkje bli slik. På høgare nivå i tydingsstrukturen blir tydingsbolken nedskoren på ulike måtar. Dette ligg dels i redigeringsprogrammet, t.d. ved at ein ikkje får lage undertydingar eller underoppslag dersom det ligg ein tydingsbolk på eit lågare nivå. Dels er det sett restriksjonar på bruken av tydingsbolken gjennom redigeringsreglane.

I spørsmålet om kvar eventuelle restriksjonar skal liggje, i programmet eller i redigeringsreglane, må det leggjast mykje vekt på kva som tener produksjonen. Norsk Ordbok må bli ferdig etter eit oppsett program, og dermed må arbeidet leggjast til rette slik at ein unngår dødtid. Det som teoretisk kunne vore den logisk beste løysinga, må dermed vike om ei anna løysing gjev raskare og lettare manusproduksjon.

7 Kva er nytt i høve til Norsk Ordbok band 1-4?

Av det som er sagt hittil, framgår det at Norsk Ordbok band 5-12 blir noko endra i høve til band 1-4. Opplysningskategoriane er dei same som før, men dei blir klarare definerte og får ein litt strengare orden. Talet på tydingsnivå er sett til fire og blir gjort synleg med faste markørar, mot tidlegare, då strukturen kunne bli uoversiktleg djup. Føremålet er å ta vare på alt som er bra med Norsk Ordbok slik verket er, men presentere arbeidet på ein ryddigare og meir lettlesne måte, slik at brukarane lettare kjenner att strukturen frå ordartikkel til ordartikkel. I tydingsbolken i Norsk Ordbok band 5-12 blir det høve til å ta med:

- Faste opplysningskategoriar i fast rekkjefølgje
- Éin innleiande definisjon pr tydingsbolk
- Eksemplifisering i form av redaksjonelle døme, målføredøme og litterære sitat (og truleg færre og meir nøye utplukka døme enn før)
- Undertydingar som spesifiserer den innleiande definisjonen
- Underoppslag som viser oppslagsordet slik det inngår i faste leksikaliserte ordsamband
- Samansetningsrekker som viser oppslagsordet i den aktuelle tydinga som etterledd og som førsteledd

Resultatet er at Norsk Ordbok band 5-12 får ein flatare artikkelstruktur enn band 1-4, altså meir sideordning og mindre underordning av tydingar. Retninga innanfor tydingsbolken vil alltid gå frå det meir generelle på høgaste nivå til det meir spesielle på lågare nivå.

8 Definisjonsdelen av tydingsbolken

Definisjonsdelen av tydingsbolken er den viktigaste delen, og den som redaktørane legg mest arbeid i. Det er vanleg godteke at å forme gode definisjonar er ei krevjande oppgåve, så krevjande at det kallar på ei eiga evne til abstrahering og ordlegging. Men denne evna må også trenast opp.

Elementa i definisjonsskjemaet til Norsk Ordbok er dei vanlege. Det er (1) bruksmarkering, (2) definisjonstekst, (3) kjelder (ein eller fleire), (4) tilvising til andre oppslagsord ved redaktør.

Her er eit einfelt døme på ein definisjon under oppslagsordet **fustans-kyrtel**, sett opp i definisjonsskjemaet i redigeringsprogrammet:

bruksmarkering	Definisjonstekst	Kjelder	Tilvising ved redaktør
om norr. tilh	kyrtel av fustan	He.ES 292	jfr <i>fustian</i>

Slik ser dette ut i Norsk Ordbok:

fustans-kyrtel m (om norr. tilh) kyrtel av fustan (He.ES 292); jfr *fustian*.

9 Definisjon i tabellformat

Det er sagt før at Norsk Ordbok er materialstyrt. Redaktørane må presentere det tilfanget vi

har; dei kan ikkje gå ut frå at tydingar har anna utbreiing eller er meir spesifiserte eller generelle enn det som framgår av kjeldene.

Ordboka må derfor ha eit definisjonsskjema som tek høgd for at bruksdøma i tilfanget ikkje alltid refererer presist til ei gjeven avgrensing av definisjonen. Det blir gjort ved å setje opp definisjonsteksten i tabell som nedanfor:

hagfelle f

Definisjonstekst	Kjelder	Redaksjonell tilvising
gjerde, skigard	HdmA, Ød mst, Gbr, ØToten, Eikesdal, Oppdal, Kvikne; Østb VII, 57; Austland 1934	
(laust oppsett) gjerde av tre, ris og kvist	Ød mst, Follidal, NGbr, Hemsedal, Kvikne, Skogn, Grong	
(serl:) gjerde av felte stomnar som ligg slik at toppen av eitt tre ligg over stomnen på det neste	A2; Vardal, Tolga, Skjåk, Heidal, Fåberg, Gausdal, Land1743, INdm sumst, STrl	

Så kjem bruksdøma i samla flokk etterpå. Dei kan vere ganske rundt uttrykte, og høve for alle definisjonsledd, t.d.:

det var dårleg stell med hafellun rundt vangane (Follidal B.Streitlien 81)

Denne måten å ordne definisjonen på er særleg nyttig i høve til (eldre) målføretilfang. I det innsende målføretilfanget kan definisjonane vere fleirledda og sær spesifikke. Men det framgår ikkje av oppstillinga kva for definisjonlekk kvart døme knyter seg til, og mest sannsynleg kan dei stå til alle. Eit drag ved munnleg målbruk er jo at ordlegginga ofte er lite eksplisitt, av di di bae partar veit kva omgrep kvart ord refererer til.

10 Undertydinga

Som nemnt er det tre tydingsnivå i ordartikkelen, og som hovudregel er det nok. Men for mindre tydingsavbrigde med lite tilfang attom har NO også rom for ei undertyding. Undertydinga kan ein berre bruke på lågaste nivå i artikkelen. Har ein undertyding på ein tydingsbolk som er nummerert med **3**), kan ein altså ikkje ha **a**), **b**), osv under tretalet.

I Norsk Ordbok band 1-4 blir undertydinga først og fremst markert ved definisjon; ofte med ei bruksmarkering framfor, men utan tydeleg utskiljing frå annan tekst. Det finst døme på undertyding etter semikolon, enkel eller dobbel skråstrek.

I redigeringsprogrammet blir undertydingar normalt innleidde med ein dobbel skråstrek. Under **gryte 2a** står til dømes underdefinisjonen

// lågt lende mellom fjell el åsar

følgd av kjeldefest bruksdøme.

I NO band 1-4 er det inga streng rekkjefølgje mellom undertydingar og underoppslag. I

redigeringsprogrammet (og dermed band 5-12) kjem undertydingar alltid etter hovuddefinisjonen og før eventuelle underoppslag (jf. pk. 7 ovanfor).

Undertydinga kan ha definisjonstabell, liksom hovuddefinisjonen (jf. pk. 9 ovanfor).

Undertydinga kan også ha eksemplifisering – bruksdøme med fullt utstyr av forklaringar og tilvisingar – liksom hovuddefinisjonen i tydingsbolken.

Normalt vil likevel undertydinga vere mykje mindre omfattande enn ein vanleg tydingsbolk.

11 Underoppslaget

Norsk Ordbok skil seg frå Nynorskordboka og andre handordbøker ved å gje etter måten utførleg handsaming av leksikaliserte fleiordige uttrykk. Dette er eit vanskeleg område innanfor den leksikalske semantikken.

Eit fleiordig uttrykk som skal kvalifisere som underoppslag, må (1) ha fast form, (2) fungere som ord med ordklasse i ei setning, (3) ha ei eller fleire tydingar som ikkje kan dragast ut av samanstillinga av einskildord.

Redigeringsreglane for underoppslaget er ikkje finpusa enno, men det er likevel klart at skildringa av tyding og bruk for eit fleiordig uttrykk krev dei same kategoriane som ein vanleg ordartikkel. Ein kan – som før – føre fleire nummererte tydingar med bruksdøme for underoppslag. Det blir ikkje høve til å føre undertydingar eller samansetningsrekkjer.

Døme på underoppslag er (1) verb med partikkel (for dei store verba samla og ordna alfabetisk under eigne tydingsnummer sist i ordartikkelen, jf. t.d. **føra** 13 a-q), (2) substantiv med modifisering (t.d. *kalka grav* ‘noko som læst vera betre enn det er’)

Fleiordige adverb og konjunksjonar er i NO band 1-4 somme gonger handsama som ordinære oppslag med ordartikkel (på alfabetisk plass), andre gonger som underoppslag under hovudordet. Det er meininga at slike uttrykk skal få ei meir konsistent behandling i NO band 5-12.

12 Samansetningsrekkjene

Metaordboka har om lag dobbelt så mange oppslagsord som det er planlagt at Norsk Ordbok skal ha (jf. pk. 2-3). Det siste elementet i tydingsbolken, samansetningsrekkjene, gjev høve til å lenke ordboksartiklar til det uredigerte tilfanget.

Samansetningsrekkjene er av to slag, etterleddsrekkjer og førsteleddsrekkjer. I programmet har samansetningstypene fast innbyrdes rekkjefølgd og fast plass sist i tydingsbolken.

I ei *etterleddsrekkje* er oppslagsordet etterledd for dei opplista samansetningane. Ei etterleddsrekkje kan sjå slik ut (under **galen** 2c i tyd ‘vill, grådig, svært huga på’):

øg i sms som *bil-*, *danse-*, *film-*, *otball-*, *gifte-*, *gjente-*, *gute-*, *kjøpe-*, *penge-galen*.

Når ein har definisjonen ovanfor, er desse samansetningane sjølvforklarande, og dei styrker definisjonsinnhaldet. I Norsk Ordbok 1-4 er somme av etterleddssamansetningane redigerte, andre ikkje. Det kjem an på tilfanget.

I ei *førsteleddsrekkje* er oppslagsordet førsteledd, og orda høyrer då alfabetisk heime nær

oppslagsordet. Dei orda som står oppførte i førsteleddsrekker, blir ikkje redigerte i eigen ordartikkel. Det dreiar seg om samansetningar som stort sett utbreidde, og som i Norsk Ordbok meir eksemplifiserer ein type ordlagning innanfor eit gjeve semantisk område, enn eit leksikalisert omgrep. Ei førsteleddsrekke kan sjå slik ut (under **glasnost**)

...; hertil *glasnost-bylgje, -politikk*

Dette er spontanlagningar av ein type som alle dei nordiske språka er rike på. Det er interessant å ha dei med, av di dei syner at oppslagsordet er produktivt i ordlagning, og kva slag ord som blir laga, men å definere slike samansetningar fell utanfor NOs arbeidsområde.

Dersom eit oppslagsord er førsteledd i særskilde mange samansetningar, vil ein god del av dei nødvendigvis vere spontanlagningar. Da kan det høve med ein egen førsteleddsartikkel, der ein kan få gruppert samansetningane etter semantiske kriterium. Det er ei oppgåve for redaktørane å evaluere tilfanget slik at sorteringa i +/- redigeringsverdig blir mest mogleg treffsikker.

Den store forbetringa i høve til den tidlegare redigeringspraksisen er at når Norsk Ordbok frå og med band 5 blir redigert inn i database med direkte kopling til Metaordboka, kan den som er interessert i eit uredigert ord, gå rett til tilfanget og sjå på belegg og kjelder. Dermed kan redaktørane halde seg å jour med tilvekst i samlingane, og omvurdering er mogleg fram til ombrekk korrektur.

13 Tilvisingssystemet

Inga ordbok utan tilvisingar! Det er to slag tilvisingar i eit dokumentarisk ordboksverk: (1) verksinterne tilvisingar til oppslagsord og tydingselement, og (2) verkseksterne tilvisingar til tilfanget som ligg attom kvar ordartikkel

Dei verksinterne tilvisingane er det som til vanleg blir kalla tilvisingsapparatet i ei ordbok. Dei kan vere eksplisitte, dvs. utstyrte med ei eller anna form for markering som seier «no kjem det ei tilvising», som «jfr» eller andre innleiingsformlar. Men dei kan også vere innbygde i definisjonen:

I **ap** m, n .. ein som apar (apast), driv, fer med **ap** (II,1) (tilvisinga er utheva)

Her er det vist til ein tydingsbolke under ein annan ordartikkel.

Denne tilvisingsmåten er særleg nyttig – og viktig – når ein definerer eit artsnamn i eit målføre med det offisielle eller dominerande artsnamnet. I ordartikkelen under hovudartsnamnet finn ein så utførleg definisjon, latinsk namn osv. Slike eittordsdefinisjonar blir til vanleg kalla synonymdefinisjonar, og dei fungerer som verksinterne tilvisingar.

I moderne redigeringsprogram for ordbøker (databasar) er det etter kvart blitt standard at tilvisingar blir lenka til den ordartikkelen eller tydingsbolken dei viser til, på ein slik måte at om plasseringa av den artikkelen eller tydingsbolken ein viser til, blir endra, så blir tilvisinga også endra. Det bruker også vere mogleg å gå inn i artikkelen ein vil vise til under redigeringa, slik at ein kan leggje lenka der den skal vere. Slik blir det også for Norsk Ordbok.

Dei verkseksterne tilvisingane til kjeldene er eit velkjent drag ved alle større ordboksverk. Kjeldeføringa viser t.d. til ei bok med sidetal osv. der sitatet står. Om ein vil, kan ein få tak i boka og slå opp, og såleis kontrollere om definisjonen verkeleg høver med sitatet.

I redigeringsprogrammet for Norsk Ordbok 2014 skjer kjeldekontroll via programmet. Som del av redigeringsprosessen må ein hente opp Metaordboka, sjå på belegga, og sortere og tilordne alle belegg ein bruker som underlag for ei tyding. Denne sorteringa blir lagra i databasen. Slik blir også dei belegga som ikkje kjem fram i ordartikkelen som kjeldetilvisingar eller sitat, knytte til rett definisjon, og heile materialgrunnlaget kan etterprøvast.

Denne lenkingsmåten er like gammal som dei dokumentariske ordboksverka. I arkiva til SAOB ligg papirsetlane sirleg sorterte i rett rekkjefølgd, samanhefta etter kva tyding dei er nytta under. Men dette arkivet er berre tilgjengeleg for dei få innvigde. Den elektroniske versjonen av dette dokumentariske prinsippet blir tilgjengeleg for alle med internett-samband til heime-pc-en.

14 Konklusjon

Rammene som er lagde for fullføring av Norsk Ordbok er strenge, men nødvendige, jf. Bakken s. 29. Ordboka må bli ferdig slik at det tilfanget vi har, kan bli ordna og tolka av redaktørar som forstår det. Det er alt ei spesialisert oppgåve å meistre den kunsten, for svært mykje av ikkje minst målføretilfanget er alt 2-3 ættleder gammalt.

Eit godt redigeringsprogram skal frita redaktørane i høgast mogleg grad frå *unødvendig* pirk og *unødvendig* utøving av skjønn. Kreftene må gå til det berre gode ordboksredaktørar kan gjere, nemleg analysere grunnlagstilfanget, gjere ei sortering på leksikalsk og semantisk grunnlag og forme korrekte definisjonar med råkande døme som illustrasjon..

Norsk Ordbok tok i bruk ein førebels versjon av redigeringsprogrammet i juni 2003. Det fungerer både som opplæringsreiskap i ordboksredigering og som reiskap for å ordne tilfanget.

Utviklingsarbeidet med redigeringsprogrammet har også hatt ein annan funksjon. Det er med på å rettferdiggjere det at Norsk Ordbok er plassert ved ein vitskapsinstitusjon.

Ei ordbok gjev ikkje den fulle og heile sanninga om språket som er skildra. Ordboka undertrykkjer i ein viss mon det utydelege og fleirtydige, kvesser definisjonar og set strengare krav til kategorisering av språkleg tilfang enn det ein vil kunne dra ut av levande tekst. Det er ikkje mogleg – eller ønskjeleg – å få med alle språklege gråsonar i ei ordbok.

Ei røynsle frå arbeidet med redigeringsprogrammet er at det strenge formatet gjev ein kraftig tilskuv til fagleg medvetsgjering når det gjeld å avgrense og definere språklege kategoriar – og til å sjå gråsonane og finne måtar å handtere dei på. Såleis kan ein sjå arbeid med eit digitalt redigeringsprogram ikkje berre som ein effektiviseringsreiskap, men også som ein intensiv skole i lingvistisk kategorisering.

Noter

1. ALLEX-prosjektet er eit samarbeidsprosjekt mellom Universitetet i Oslo og Universitetet i Zimbabwe om å lage morsmålsordbøker for dei afrikanske språka i Zimbabwe, og i samband med dette lære opp zimbabwiske morsmålsfilologar i moderne leksikografisk metode, bygge opp korpora for morsmåla og skape eit senter for leksikografi ved Universitetet i Zimbabwe. Prosjektet byrja i 1991 og blir avslutta i 2006, og er finansiert av The Norwegian Council for Higher Education's Programme for Development Research and Education (NUFU).
2. Ordet «tydingsbolke» er ein terminologisk nyskaping som er blitt til av di vi ikkje fann nokon etablert term for dette.

ODDRUN GRØNVIK

Verknader av digitalisering på materialvurdering, redaksjonell metode og opplæring

Effects of digitisation on the use of source materials, editorial method and editorial training. Digitisation of Norway's largest academic dictionary project, Norsk Ordbok 2014, through an interlinked set of relational databases has led to new ways of using source materials in lemma selection, manuscript management and quality control. Editorial method is radically changed by increased transparency, both in looking at source materials and in handling dictionary and entry formats. These changes are prerequisites for training linguists to become productive editors within a minimal time span, with six months for basic training and full editorial responsibility after one year. The aim of the exercise is to complete seven volumes of Norsk Ordbok in nine years, with no loss of, and preferably an increase in, quality.

1. Innleiing

Norsk Ordbok (NO) er eit vitenskapleg ordboksverk over nynorsk skriftmål og norsk talemål, der forfattarskap til ordartiklar er akademisk meriterande for kvar einskild redaktør. Redigeringsstad er Universitetet i Oslo, der redaksjonen er organisert som prosjekt under Institutt for lingvistik og nordistikk. Digitaliseringa av NO 2014 skjer i tett samarbeid med Eining for digital dokumentasjon ved Historisk-Filosofisk fakultet.

I perioden 2003-2014 er ordboksprosjektet refinansiert og omorganisert med det føremålet at alle tolv band skal vere utgjevne ved utlauptet av 2014, som del av 200-årsjubileet for de norske Grunnlova. Verket skal (1) vere konsistent i materialutnytting og redigeringspraksis, (2) halde høg vitenskapleg standard når det gjeld lingvistisk kategorisering og innhald, og (3) vere dokumentert ned til detaljnivå, slik at alle opplysningar lett kan etterprøvast.

Dette er bakteppet for det store hamskiftet i Norsk Ordbok, som er full digitalisering av heile prosjektet, kombinert med rekruttering og opplæring av eit tjuetals nye redaktørar. Den følgjande drøftinga tek for seg verknader av digitalisering på materialvurdering, redaksjonell metode og opplæring, sett frå redaktørsynsstad og frå leiinga si side.

2. Materialvurdering

Konsistent materialvurdering i høve til verket som heilskap er vesentleg mellom anna for lemmaseleksjon og omfangsdimensjonering. Første oppgåve for prosjektet var derfor å skaffe seg oversyn over underlagsmaterialet, både i høve til lemma (mengd, type; normaliseringsgrad) og kjelder.

2.1. Underlagsmateriale for Norsk Ordbok

NO er ei dokumentasjonsordbok som opphavleg vart bygd opp på

1. nedskrivne talemålsopplysningar (ikkje lyd)
2. litteraturekserpt

Samlingane, som Norsk Ordbok byggjer på (i hovudsak skapte før prosjektstart i 2001) er svært ueinsarta. Eit slåande drag ved dei, som samlingar, er mangelen på retningslinjer for oppbygginga, og den svært ujamne utnyttinga av ulike kjelder. Dette heng saman med den ujamne sjangerdekninga i nynorsk skriftmål – noko som er typisk for minoritetsspråk – og den like ujamne dokumentasjonen av norske målføre. I dei eldre samlingane har ein måtta bruke alt tilgjengeleg tilfang som var av verdi (Grønvik 1997:33).

Det gjeld generelt for ekserperte ordsamlingar at dei har underdekning av frekvent ordtilfang, og av ordtilfang frå nyare kjelder. Det galdt også dei nynorske ordsamlingane. NO 2014 har derfor bygd opp eit supplerande korpus av nynorsk tekst som i 2005 er på om lag 23 mill. ord. Om lag 80 % av tekstmengda er frå etter 1975. Hovudvekta ligg på sakprosa, men der er også skjønnlitteratur og ein del tekstar frå klassisk nynorsk, mellom anna den første fullstendige nynorske bibelutgåva.¹

Korpuset er særskilt nyttig til å gje betre og meir nyansert dekning av det frekvente ordtilfanget. Lemma frå korpuset som ikkje finst i samlingane, er for det meste samansetningar – utan at dei derfor er uinteressante. Korpuset gjev dessutan høve til å utnytte ein del av dei eldre litterære kjeldane på ein mykje grundigare måte, i og med det gjev tilgjenge til heile tekstar, ikkje berre ekserpt.

2.2. Materialoversyn – Metaordboka

Det manglar førebels oversyn over kva for tilfang som er utnytta, og korleis, i den delen av Norsk Ordbok som var redigert på gamlemåten (band 1-4, utgjevne 1965-2002, og band 5 harm-håvøyll). Fullt oversyn over dette får vi ikkje før band 1-4 er ferdig digitaliserte.² Men den gamle redaksjonen var medviten om at samlingane var ujamne og til dels tynne. Parolen var derfor å utnytte tilfanget til siste detalj. Bruksprinsippet var (noko forenkla) at alt av verdi skulle inn i ordboka. Mal for omfangsrekning av manus mangla. Resultatet vart redaktøravhengig, og dermed inkonsistent, materialutnytting og redigering.³ I tillegg mista ein kontroll over totalvolumet av manus (Bø 1989:90).

For å skaffe oversyn over grunnlagstilfanget vart det i perioden 2000-2003 bygt

¹ For korpusstatistikk sjå <http://folk.uio.no/danielr/no2014/nn/dist25.html>.

² Planlagd ferdig i 2006.

³ Som døme kan det nemnast at materialutnyttinga for verba *gjera*, *gjeva*, *ha*, *halda*, *gå* varierer frå om lag 120 til om lag 280 belegg per prenta side.

opp eit sams indeks for dei nynorske språksamlingane, den såkalla Metaordboka (MO), der tilfanget vart gjennom-normalisert ein gong til (om dette sjå Svardal 2003). MO synte at språksamlingane har om lag 550 000 oppslagsord. Om lag halvparten av desse oppslagsorda har eitt einaste belegg, men det belegget kunne vere alt frå ein etter måten stor artikkel i Grunnmanuskriptet⁴ for NO til eit einstaka litterært ekserpt. Lemmatal kunne derfor ikkje vere einaste utvalskriterium.

2.3. Omfang og styringsbehov

NO 2014 band 1-4 har om lag 100 000 artiklar. Plassomsyn og omsyn til kvaliteten på redigeringa tilseier at eitt band (på 800 sider) ikkje bør ha meir enn 25 000 ordartiklar. Dette gjev eit totalanslag på 300 000 artiklar fordelte på tolv band.

NO 2014 trong derfor retningslinjer for utsjalting av uaktuelle oppslagsord ut frå mengd og kvalitet på det underliggjande tilfanget. Vidare trongst det ein mal for artikkelstorleik innanfor totalramma på tolv band. Både retningslinjer for ordutval og dimensjoneringsmal måtte kunne automatiserast gjennom redigeringsprogrammet, men det måtte også vere mogleg for redaktørane å overprøve resultatet dersom utslaget vart vurdert som fagleg urimeleg.

2.4. Materialvurdering som grunnlag for lemmaseleksjon

Lemmaseleksjon i eit verk som NO må dels byggje på generelle omsyn, dels på omsyn som er spesifikke for verket. Generelle omsyn tilseier at grunnord får grundigare behandling enn avleingar og samansetningar, og at det nordiske ordtilfanget får ei grundigare handsaming enn importord (Venås 1989:162 f). Dei verksspesifikke omsyna har å gjera med nynorskens historie, status og tilblivingsgrunnlag, der relasjonen til norske målføre er sentral. Nynorsk er eit ungt skriftmål som ikkje kan seiast å vere fullt ut standardisert.⁵ Gjennom arbeidet med NO kjem redaktørane stadig over ordformer i målføra som ikkje er bestemte. NO har derfor eit særleg ansvar for å behandle eldre ordtilfang og unormert målføretilfang, og gje relativt større plass til mindre frekvente norske grunnord som ein ikkje finn i vanlege bruksordbøker.

Lemmaseleksjonen vart frå 2003 knytt til ei validering i kjeldegrunnlaget for NO. Vi hadde alt ein bibliografidatabase over alle skriftlege kjelder. Her var det høve til å leggje inn tilleggsopplysningar som i sin tur kunne nyttast i lemmaseleksjonen. Alle skriftlege kjelder for NO 2014 vart derfor sjangermerkte i bibliografidatabasen. Sjangermerkinga, kombinert med tal belegg, blir så nytta til å skilja mellom tre lemmagrupper, (1) lemma som skal redigerast, (2) lemma som kan redigerast, (3) lemma som er frårådde.

⁴ Sjå omtale på <http://no2014.uio.no/tekster/sok/index.php>.

⁵ Jf. det omfattande systemet med parallellformer og tillatne sideformer (klammeformer) i nynorsk rettskriving.

Obligatoriske lemma er (1) Lemma frå Grunnmanuskriptet, Skards rettskrivingsordliste og *Nynorskordboka*, og (2) Lemma frå målføretilfanget som ikkje er med i andre normerte ordbøker.

Dei viktigaste gruppene av frårådte oppslagsord er (1) lemma med tospråklege ordbøker, skuleordlister eller spesialordbøker (frå slang til yrkesterminologi) som einaste kjelder ("ordboksort"), og (2) lemma som er hapax frå skjønnlitterære omsetjingar eller som framstår som tilfeldige (skjønnlitterære) spontanlingar. I den siste gruppa er det svært mange samansette ord.

Redaktøren kan overprøve automatikken ved å leggje til nytt tilfang, til dømes frå korpuset.

2.5. Dimensjonering av band og artiklar

Etter at sjølv redigeringskjemaet med utskriftfunksjon var på plass i 2004, var det dei redigeringsadministrative funksjonane som stod høgst på prioriteringslista. Denne programmodulen er ikkje komplett enno, men eitt av dei viktigaste elementa, dimensjoneringsmål pr. artikkel, er på plass.

Det tilfanget som må redigerast, er ordna gjennom MO, og dette tilfanget er gjennomarbeidd og normalisert.⁶ Derfor er det MO som er lagt til grunn for dimensjoneringa av dei attståande banda av NO. Dimensjoneringa for einskildartiklar byggjer på ei manuell utrekning av setelprogresjon og artikkeltypar i dei ferdigredigerte banda, som så vart brukt til å lage ein utrekningsmodell for programmet, med vektning for beleggmenge pr. artikkel. Modellen gjev noko større plass til lemma med få belegg attom seg, enn lemma som byggjer på eit større tilfang. Modellen vart testa og justert, og er no teken i bruk. Kvar gong ein artikkel blir generert opp, kjem det fram eit framlegg om manslengd i tal linjer for denne artikkelen.

Dei resterande 6 banda av NO er også rekna gjennom i bolkar på 20 sider, slik at ein har første og siste lemma pr. tjuesiders bolk frå og med band 6 og ut band 12.

Redaktørane får tildelt manus til redigering i alfabetbolkar på 3-4 bokstavar (t.d. *kly-*, *kok-*). Programmet kan enno ikkje gje eit linjetal for ein alfabetbolk, så det må redaktørane rekne ut sjølv ved å summere linjetal for artiklar. Innanfor linjetalet for ein alfabetbolk kan redaktøren omdisponere artikkellengder – leggje til litt her og knipe litt der – men det er ikkje høve til å gå utover manslengd for den tildelte alfabetbolken.

Dimensjoneringmodulen kom inn i redigeringsprogrammet ved påsketiden 2005, og verknaden vart synleg med ein gong. Før dette hadde mange redaktørar store problem med å halde omfangsgrensene, med påfølgjande lange retterunder og omskrivingar. Etter at dimensjoneringsmodulen kom inn, held dei fleste linjetalet sitt i første omgang, og dei kjem dermed fortare fram til leveringsklart manus.

⁶ MO i-å vart normalisert i 2000-2003. Bokstavane a-d er no normaliserte, medan normalisering pågår for bokstavane e-g. I gjennomsnitt blir tal oppslagsord pr. bokstav redusert med 15 %.

3. Redaksjonell metode

Medan redaktørane før produserte ordboksmanuskript som seriell tekst, skal dei no tilordne tilfang og skrive inn i felt med førehandsdefinerte rammer. Kva følger har dette hatt for arbeidsmetodar, produksjon og kvalitetskontroll?

3.1. Redigering i *Norsk Ordbok før prosjektstart*

Produksjonen av ordboksmanus har tre komponentar, uavhengig av hjelperåder. Den første, og morosamaste, er den som er rein filologi og lingvistikk, som omfattar vurdering og ordning av tilfang, utforming av definisjonar, utveljing og kommentering av døme osv. Den andre fasen er utforming og nedskrivning av manus, med kjelder, kryssreferansar, spesialteikn og konvensjonar – og dei er det mange av i eit akademisk ordboksverk. Den tredje fasen er klargjering av manus for produksjon, med rett sats og typografi osv. Ein stad imellom fase 2 og 3 ligg ansvaret for at manuskriptet er feilfritt, både tekstleg og reelt (kontroll av tilvisingar osv.).

Den tradisjonelle arbeidsdelinga i mange akademiske ordboksredaksjonar har vore at filologar har ansvar for den første fasen, medan ymse hjelpepersonell har teke seg av mest mogleg innanfor fase 2 og 3. For tjuve år sidan vart NO-manus dels skriva for hand, dels på pc, av redaktørane. Alt den gongen vart manuskriptet feltinndelt (etter ei blanding av typografiske og leksikografiske kriterium) med kodar før setjing (sjå t.d. Hjulstad 1986; Worren 1992). Men det vart gjort av ein sekretær, og det var berre ho som kunne kodesystemet. Redaktørane var mest opptekne av at typografien vart rett i korrektoren, og ville neppe oppdage innhaldsmessig feilkoding dersom det ikkje vart avslørt av feil i typografien.

Dei fleste ordboksverk har hjelpedokument til støtte for redigeringa. For NOs del er det vel rett å seie at hjelpedokumenta før 2003 var nokså einsidig innretta mot akribien i verket, dvs. tekstkonvensjonar, pluss kjeldekunnskap og -kontroll.

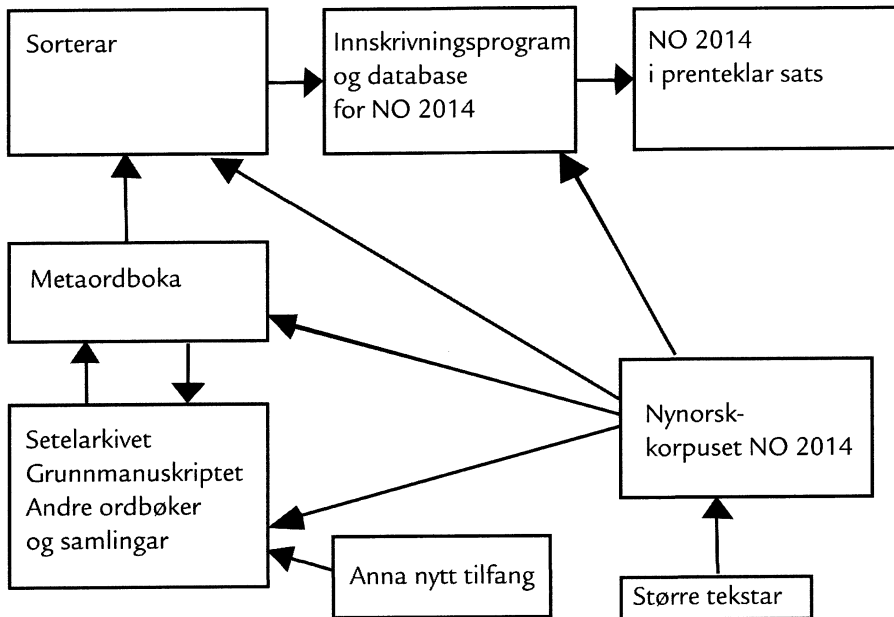
Dette heng dels saman med kva det er mogleg å kontrollere manuelt i ein flat tekst som har mange forfattarar. Det skal også seiast at NO har uvanleg krevjande konvensjonar, som skal ta høgd for både tale og skrift, og både synkron og diakron formvariasjon. Redaksjonelt fokus og redaksjonell metode kom derfor til å bli prega av behovet for korrekt handtering av ordboksakribien, medan dei lingvistiske kriteria nok var føresette, men kom meir i bakgrunnen.

3.2. *Krav til redigeringsprogrammet for NO 2014*

For NO 2014 var det eit overordna mål at innskrivingsprogrammet så langt mogleg skulle fristille redaktørane til å samle seg om det lingvistiske og leksikografiske i arbeidet, medan konvensjonane skulle handterast av programmet. I høve til kjeldeføring og tilvisingar var det også ønskeleg å ikkje berre yte hjelp til effektiv redigering, men også førebyggje og hindre feil.

Både behovet for datatryggleik og behovet for kontroll tilsa at ordboka skulle

skrivast inn i ein relasjonell database, feltinndelt for ulike opplysningskategoriar, og at heile dokumentasjonssystemet rundt ordboka (som alt var lagra i andre databasar) skulle lenkast mest mogleg direkte til relevante felt i artiklane (meir om dette i Grønvik og Tvedt 2006). Figur 1 syner korleis dei viktigaste komponentane i redigeringsapplikasjonen er kopla saman.



Figur 1. NO 2014: Materialorganisering mellom databasar og andre komponentar

Eit anna krav var at det skulle vere minst mogleg nytt å lære for kvar ny programdel som kom på plass. Grensesnitt, ikon og basisfunksjonar er derfor mest mogleg like, og likt plassert, frå programdel til programdel og skjermbilete til skjermbilete. For eit meir IT-basert perspektiv på dette, sjå Grønvik og Tvedt 2006.

3.3. Innskrivningsprogrammet som modell av ordartikkelen

Dei fleste ordartiklane NO er stutte og etter måten enkle, men dei store artiklane kan vere kjempedigre og sær innfløkte. Det var opplagt at innskrivningsprogrammet måtte spegle ordboksartikkelen i maksimumsformat. Det har vore drøfta om det i ettertid skulle lagast eit redusert ”kvardags”-skjema, men sidan kategoribehovet kan variere sterkt frå små-artikkel til små-artikkel har det ikkje blitt følgt opp.

Det første som vart gjort, var å analysere tidlegare praksis i band 1-4, og skilje ut

hovuddrag i struktur. På grunnlag av dette vart det laga framlegg til artikkelskjema, som utgangspunkt for programmeringa (om dette sjå Grønvik 2005).

3.4. *Endringar i arbeidsmetode*

Akademisk leksikografi dreiar seg om å sortere, systematisere og kommentere språklege data. Når artikkelskrivinga går føre seg ut frå papirarkiv, og manus blir skrivne som laupande tekst, vil arbeidet arte seg slik:

- sortering av papirbelegg i haugar
- ordning (systematisering) av setelhaugane (utover skrivebord eller i sorteringskasser) (gjerne med påhefta lappar med ”overskrift” og nummerering)
- forming av tekst ut frå kvar liten bunke, i rett rekkjefølgd
- plassering av tilfang i arkiv igjen, helst ordna slik at det speglar artikkelstrukturen, med dei siterte setlane først og i rett rekkjefølgd.

Det vil seie at ein øvd redaktør skal ha i hovudet (1) heile kategorisystemet, (2) ordninga av kategoriane i hierarki, og (3) alle konvensjonane som gjeld for verket.

Med redigeringsskjema som inngang til ein ordboksdatabase blir dei viktigaste endringane i arbeidsmetode desse:

- Kategorisystemet (artikkelstrukturen) er synleg på skjermen ned til detaljnivå før ein tek til å arbeide
- Ordninga (systematisering) av kategoriar i hierarki er innebygd i systemet
- Rekkjefølgja i innskriving er langt mindre relevant enn den var – det som tel, er å skrive inn teksten i rett felt. Ranging og rekkjefølgje kan takast i ein separat prosess til slutt.
- Tilfanget er digitalisert, og kan sorterast og tilordnast digitalt ned på sitatnivå, slik at kven som helst kan kontrollere materialordninga attom kvar artikkel. Ein slepp å lagre sorteringa i ein eigen prosess etter redigeringa.

3.5. *Grensesnittet*

Som nemnt kan dagens NO-redaktørar sitje og sjå på det som førre generasjon måtte ha i hovudet. Det blir her gjeve ein kort gjennomgang av korleis innskrivingsdelen av applikasjonen er ordna, visuelt og innhaldsmessig:

Grensesnittet i innskrivingsprogrammet er delt i tre, med (1) ein sams toppdel for heile Oracle-systemet (ikonrekkje for basisfunksjonar osv.), (2) artikkelen sett opp i trediaqram til venstre, og (3) innskrivingsdelen til høgre. Innskrivingsdelen har form av ei rekkje skjema, eitt for kvar hovuddel ein artikkel kan ha. På skjerm er desse skjema ordna slik at det ser ut som arkivkort, jf. tabell 1.

[trediagram]	Artikkel	Oppslag	Hakeparentes	Tydingsbolk	Tilvising	Sorterar
--------------	----------	---------	--------------	-------------	-----------	----------

Tabell 1.

Under kvart av desse "korta" finn ein så kortsystem som dekkjer dei aktuelle underkategoriane. Tabell 2 viser underkategoriar pr. kort.

1 [trediagram]	2 Artikkel	3 Oppslag	4 Hakeparentes	5 Tydingsbolk	6 Tilvising	7 Sorterar
[tre – viser artikkel-element ned til kjelde. Viser også kva for artiklar som har tilvising til den artikkelen ein står i – jf. kort for Tilvising]	Dimensjonering Status-historie Redaktør-historie Merknad	Homografnr Oppslagsord Vokalparentes Leddanalyse Ordklasse Ordklassetilstand Tabell for bøyingsssystem	Uttale Eldre kjelder Avvikande skriftformer Målføreformer Etymologi	Definisjon Sitat/Døme Samansetting s-rekkjer Undertydingar Underoppslag	[skjema for vising av artikkel som viser til den artikkelen ein står i]	[område for eigendefinerte tabellsett over tilfanget bak kvar ordartikkel, der redaktøren kan sortere og lagre sorteringar]

Tabell 2. *Underkategoriar pr. kort*

3.5.1. Trediagram

I utgangspunktet var det ein viss skepsis til tredigrammet som arbeidsreidskap, truleg av di det var uvant. Den store styrken ved tredigrammet er at det gjev godt oversyn, og er lett å omorganisere. Flyttar ein rundt på tydingar, så følgjer alt "utstyr" med. Alle endringar synst med ein gong. Denne føremonen har stadig fleire kome til å verdsetje i det daglege arbeidet.

3.5.2. Artikkel

Her ligg dei redigeringsadministrative opplysningane pr. artikkel. Dimensjonering er kommentert ovanfor. Statushistorie syner kvar ein artikkel er i produksjonsgangen, redaktørhistorie syner om ein eller fleire har vore innoom artikkelen osv. I merknadsfeltet kan redaktøren notere særskilte tilhøve ved ein artikkel, og eventuelt sende desse vidare til gruppeleiar eller hovudredaktør.

3.5.3. Oppslag

Her ligg skjema for oppslagsdelen av artikkelen. Dette skjemaet skil seg frå dei andre ved at det er menyar eller føringsrestriksjonar for samtlege felt utan feltet for oppslagsord og bøyingsformer, og oppslagsordet blir generert opp. Berre bøyingsformer blir skrivne inn av redaktøren.

3.5.4. Hakeparentesen

Fem opplysningskategoriar har plass her, og tre – eldre kjelder, avvikande skriftformer og målføreformer – krev kjeldeføring. Kjeldeføringa er menystyrt, og heimfestingane for målføreopplysningane har no også fått automatisert hierarki, slik at områdenemne overtek for punktheimfestingar når dekningsgraden er god nok. Det er ikkje mogleg å føre både områdenemne og punktheimfesting. Uttaledelen er ikkje menystyrt, men dei fleste aktuelle føringmønster er opplista i redigeringshandboka, for å sikre konsistens.

Formatet for hakeparentesen vil bli justert ein gong til i samband med at band 1-5 skal leggjast inn i ordboksdatan. Da blir det høve til å leggje ei siste hand på etymologiskjemaet, der formatet førebels gnissar litt mot materialet.

3.5.5. Tydingsbolken

Her blir hovudtyngda av det skapande redaktørarbeidet lagt inn. Kjeldefesting, krysstilvisingssystem og tilvisingar til lemma som ikkje blir redigerte i papirutgåva (semantisk gjennomsynlege samansettingar med oppslagsordet som førsteledd) er også her lagt inn som menyar og lenker, med innebygde restriksjonar mot t.d. sirkelkoplingar. Under redigeringa har ein fullt oversyn over desse artikkel-elementa så langt ein vil, dei kjem opp som peikarar i tredigrammet etter kvart som dei blir hekta på.

Ein kan sjå på papirordboka som ein simulert database, der tilvisingane skal uttrykkje lenker m.a. i det semantiske hierarkiet som finst i språket. Men papirordboka kan aldri bli verkeleg stringent, for feilkoplingar straffar seg ikkje direkte. Med NO som database, straffar det seg straks å kople feil. Lagar ein sirkelkoplingar, får ein ikkje lagre. Blir artikkelar feilplasserte i førsteleddsrekke, forsvinn dei frå oppgenerert sats. Ingen redaktør gjer den typen feil mange gonger.

Tydingsbolken var ferdig spesifisert alt i 2003, og er ikkje endra når det gjeld innhald, men det har vore somme justeringar av funksjonaliteten. Det nye (2005) er at materialsorteringar og korpus er kopla inn i applikasjonen, slik at sorteringresultat kan lagrast etter kvart, anten samla på artikkelnode, eller spesifisert på tyding (definisjonselement) og sitat eller døme. Det viser seg at tilfanget attom ulike lemma kan invitere til ulike kategorisering. Kvar redaktør rår over egne sorteringar og utval, ut frå sitt faglege skjønn. For somme artikkelar blir det laga førebelse tydingskjema i trestrukturen med direkte innsortering av tilfang, for andre blir det sortert etter meir empiriske kriterium (t.d. kollokat). Ein kan lage – og lagre – eit uavgrensa tal sorteringar for kvar artikkel.

3.5.6. Tilvising

På dette ”kortet” kan ein sjå den artikkelen som har ei tilvising til det lemmaet ein arbeider med. Tilvisingskoplingane ligg i tredigrammet som ei liste med peikarar

under lemmaet, men er også distribuert på tydingsbolkane i treet, slik at ein skal kunne gå rett på ved første forsøk.

3.5.7. *Sorterar*

Dette er som før nemnt det ferskaste modulen i redigeringsprogrammet. Sorteraren er tilgjengeleg via innskrivingsapplikasjonen, men er ikkje del av den. Den er meir å likne med ei handlekorg, der ein kan hente inn tilfang frå MO, frå korpuset, og etter kvart, om vi skulle ønskje det, frå andre kjelder. Sorteraren er forma som eit rekneark der tilfanget ligg i første kolonne, og så kan ein leggje til kolonner, subkategorisere og ordne av hjartans lyst. Her kan ein også skilje homograr og kopiere over sorteringar til andre artiklar.

Sorteraren var eit svært etterlengte hjelpemiddel i arbeidet med dei store artiklane, og har alt synt seg svært anvendeleg i fleire fasar av redigeringsprosessen. NO 2014 har ein satelittredaksjon ved NTNU i Trondheim, som bere har tilgjenge til samlingane over nett. Her var sorteraren særskilt etterlengta, ettersom den fristiller redaktørane frå omfattande kladding og sortering i vanlege tekstbehandlingsprogram.

3.6. *Eit ord om søkjesystemet*

I og med at heile NO-systemet er eit sett av databasar, er også alt søkbart. Dette gjer det lett å finne modellar og leite etter parallellar når ein står fast. Ein kan t.d. søkje opp alle definisjonar under alle lemma på ”-skap m” og sjå om det finst mønsterformuleringar (det bør det jo vere). Søkjesystemet blir nytta i redigering, til administrasjon og i konsistenskontroll – og med god effekt. Når manglande stringens kan finnast med eit par tastetrykk, er det bryet verdt å rydde godt opp etter seg.

3.7. *Oppsummering*

Dei store endringa i redigeringsmetode som applikasjonen med tilhøyrande databasesystem har ført med seg, er at redaksjonen no har eit samla overblikk over tilfang, artiklar og dokumentasjon, ikkje berre frå ein synsvinkel, men frå mange. Alle kan gå inn kvar dei vil i systemet og sjå på kva dei vil. Sett på spissen kan ein seie at medan kvar redaktør før sat aleine og neddykka i tilfanget ”sitt”, så sit redaktørkollegiet no høgt oppe og ser på korleis kvar og ein ordnar opp med sin teig. Medan spørsmålsstillinga før ville vera ”korleis skal eg handtere dette ordet?” er spørsmålet no (i større grad i alle fall) ”korleis handterer NO denne typen ord?”

Den endra spørsmålsstillinga aktualiserer dermed meir generelle faglege problemstillingar, som i sin tur gjer det lettare å sjå tilfanget frå ein språkvitskapleg ståstad. Det er lett å opprette eit artikkelskjelett (trestruktur) på grunnlag av ein hypotese, og prøveplassere tilfang inn ut frå ei generell lingvistisk tilnærming til ei gjeven ordgruppe. Dermed er det også t.d. lettare å sjå møtepunkta mellom grammatikk og semantikk, og ikkje minst sjå det semantiske innhaldet i dei større gram-

matiske kategoriane. Det er sannsynleg at dette vil gje NO meir systematiske, og derfor betre, ordartiklar over dei store funksjonsorda (t.d. preposisjonane, ein del store verb og adverb) enn det som var mogleg å få til på gamlemåten. Også behandlinga av ordlagingselement som prefiks og suffiks vil få eit naudsynt lyft på denne måten.

Produksjonen av ordartiklar blir langt meir målretta og effektiv når tvil om konvensjonar osb. er rydda vekk. Berre det at band 5 vart ferdig til rett tid (desember 2004), med ca. 270 sider redigerte av for det meste nyttilsette redaktører, syner det.

Når det gjeld kvalitetskontroll, må den vera tvifelt. At konsistensen i verket er blitt betre, veit vi, men konsistens er ikkje nok. Påstanden at ordboksmanuset held minst like høgt fagleg nivå som før byggjer på intern kontroll, ved at gruppeleiarar og hovudredaktør les og korrigerer fortløypande, og på ekstern kontroll, ved at redaktørar presenterer ord og ordgrupper som dei har arbeidd med i vitskaplege føredrag og artiklar – m.a.o. normale akademiske kvalitetsprøvingar.

4. Opplæring

Før har det vore vanleg å rekne med at det tek to til fem år å lære opp ein redaktør i eit vitskapleg ordboksverk. Norsk Ordbok 2014 set nye redaktørar i funksjon etter svært kort opplæringstid, ved hjelp av innskrivingsprogrammet og databasesystemet.

Kva følgjer har dette for fokus og innhald i redaktør opplæringa? Lat oss opne med konklusjonen: I denne perioden har NO 2014 tilsett opp mot tjue nye redaktørar og redaksjonsassistentar. Det representerer ein stor og variert lingvistisk kompetanse, men det dei har sams er at ingen av dei hadde arbeidd i eit ordboksverk tilsvarande NO 2014 før dei byrja i stillingane sine. Alle måtte derfor ha opplæring frå botnen – og alle er blitt produktive innanfor stipulert tid.

4.1. Tradisjonell redaktør opplæring

Den tradisjonelle opplæringsmetoden i dei akademiske ordboksverka har vore at ein har kome inn i redaksjonane på ein slags lærlingstatus, og så har trødd i fotefara til meistrane, utan å få noka organisert opplæring. Opplæringsinitiativet låg hos nybyrjaren, som helst skulle lære ved å lese det som var skriva før, og så absorbere detaljane under marsjen. Ein skulle redigere, gjere feil, få feilen påpeikt, rette den, få nye feil påpeikte, osb. Tida frå første manus vart skriva og til utgjeving kunne vera svært lang (åtte-ti år er ikkje uvanleg), noko som igjen fører med seg tidkrevjande revisjon og innarbeiding av nytt tilfang i samband med korrektur og utgjeving. Med ei slik tilnærming er det ikkje rart at folk treng minst fem år for å bli varme i trøya.

4.2. Kunnskapskrav før og no

Vi kan jamføre kunnskapskrav til ordboksredaktørar i akademiske ordboksverk før og no, for å sjå kvar skilnadene ligg. Før måtte ein ha

1. Solide lingvistiske kunnskapar, ensyklopediske interesser og god formuleringsevne
2. Svært godt minne for detaljar (tilfang, redaksjonelle konvensjonar, lister, namn, ortografi ...)
3. Svært god evne til grannsemd i detaljar, også utan streng ytre kontroll
4. Evne til å tenke strukturelt er ein føremon
5. Grenselaust tolmod!

Etter kvart sat ordboksredaktøren på ein enorm detaljkunnskap om eit grunnlagstilfang som få andre kjende.

Krava i dag er dels dei same, men ikkje heilt. No trengst det:

1. Solide lingvistiske kunnskapar, ensyklopediske interesser og god formuleringsevne
2. Kunnskap om korleis ein handterer og vurderer store datamengder
3. Eit skarpt blikk for det meiningsberande i detaljar, men mindre krav til minne (mykje av det som før måtte hugsast er no bygd inn i redigeringsprogrammet)
4. Evne til å tenkje strukturelt er heilt nødvendig
5. Godt syn og god finmotorikk – elles blir det mange uhell i applikasjonen!

Ein kunne leggje til at der ordboksredaktøren før måtte tole å sitje myke aleine, må han no tole kritisk innsyn i alt han eller ho gjer.

Punkt 1, nøkkelkravet, det som kan lyfte definisjonsskriving frå handverk til kunst, er ikkje endra med tida. På dei andre punkta er det endring, og dei endringane speglar at vi lever i eit informasjonssamfunn med overflod på data og godt med pengar, men veldig lite tid. I dagens universitetssamanheng verkar det mest komisk å snakke om prosjekt som varer i meir enn ti år, slik NO 2014 gjer.

4.3. Krav til redaktøropplæring

I NO 2014 diskuterer vi ikkje lenger om det er behov for organisert opplæring av nye redaktørar. Spørsmålet er kva slag opplæring dei treng, som gjev prosjektet dyktige, produktive redaktørar på stuttast mogleg tid. Maksimal opplæringstid og prøvetid har til no vore eitt år, etter det er det normale produktivitetskrav til alle. I NO 2014 har dei nye redaktørane gått denne løypa, med mindre variasjonar:

1. Først ei økt med materialarbeid (normering i MO i 2-3 månader)
2. Deretter rett på redigering og definering
3. Får tildelt alfabetbolk
4. Får opplæring i gruppe, individuelt og gjennom seminar (førelesingsrekke knytt til redigeringshandboka, ved kolleger eller innhenta ekspertise)

5. Får støtte i detaljert redigeringshandbok, redigeringsprogram med manual og i open arbeidskultur: Det er greitt å spørje – og alle treng korrigerering.
6. Får oppmuntring og oppfølging når det gjeld å fylle faglege lakuner, t.d. ved å følgje aktuelle kurs eller seminar på UiO, konferansedeltaking osv.

4.4. Progresjon i redaktør opplæringa

Dei fleste nye redaktørar byrjar med å redigere innhaldsord, og går laus på funksjonsord først når dei har noko røynsle. Men dette gjeld ikkje i same grad for alle. Mellom rekruttane er det fleire utprega grammatikarar som raskt tek tak i større funksjonsord (som dei fekk tildelt som særskilde oppdrag).

Generelt har ordninga likevel vore slik: Opplæringa har gjeve ein gjennomgang av leksikalske særdrag og tilsvarende leksikografisk behandling ordklasse for ordklasse. Dei nye redaktørane skriv dei første artikkane sine om konkrete substantiv med 5-20 belegg og minst ein godt avgrensa sakdefinisjon. Deretter følgjer adjektiv, så verb. Ved verb bør ein ta seg tid til å gjennomgå heile verbfrasen med valens og ulike typar kollokat, av di dette er tydingsskiljande kriterium som ber strukturen i verbartiklar. Gjennomgangen av ordgrupper tek til med grunnord, deretter følgjer avleiingar og samansettingar.

Dette høyrer kanskje ut som ein repetisjon av grunnkurs i leksikalsk semantikk, men i høve til ordboksarbeid blir perspektivet annleis. Det er ikkje spørsmål om å mestre omgrepa eller forstå tenkinga. Perspektivet er ”korleis operasjonaliserer vi denne kunnskapen i materialbehandlinga, og kva følgjer får den for sortering, systematisering og framstilling i ordartikkelen?” Dette er ikkje triviell kunnskap, og det er ikkje purt handverk. Ein skal vite kva ein gjer, og korleis, men ein må også vite kvifor – og kva for alternativ som kunne vore aktuelle. Dermed er ein langt inne i generell lingvistisk teori, som jo bør tole slik materialtesting.

4.5. Prioriterte mål for NO 2014

Redigeringa av NO i database tok til våren 2003. Dei måla prosjektet då hadde for staben, var (1) meir effektiv opplæring, (2) auka tempo i manusproduksjon, (3) auka kontroll med resultatet, (4) auka konsistens i redigeringa, (5) tettare samarbeid innanfor redaksjonen og med Eining for digital dokumentasjon.

Dette har gått om lag som planlagt. Ikkje utan motbakkar, men motbakkar gjev god treningseffekt. Utan redigeringsprogrammet med tilliggjande modular, endra redaksjonell metode og effektiv opplæring hadde NO 2014 aldri fått ut band 5 til rett tid. Redigeringsprogrammet med tilhøyr blir karakterisert som sjølvforklarande av dei nyaste redaktørane, som har kome til ferdige program og slik sett dekkja bord. NO 2014 meiner derfor å kunne slå fast at redigeringsprogrammet står for ”den gode systemtvangen”.

5. Konklusjon

Digitalisering av NO 2014 har gjeve (1) konsistent materialbehandling som grunnlag for lemmaseleksjon, dimensjonering og kjeldeføring, (2) meir systematisk og kollektiv leksikografisk metode på grunnlag av lingvistisk skoloring, med prosjekt-målet som ramme og redigeringsapplikasjonen som hjelpemiddel, og (3) grunnlag for effektiv redaktør opplæring – og dermed synt at ordboksredigering er eit fag som kan lærast.

Litteratur

- Bø, Reidar 1989: Arbeidet med Norsk Ordbok. I: Almenningen, Olaf og Grønvik, Oddrun (red.): *Ord og mål. Festskrift til Magne Rommetveit*. Oslo: Kringkastingsringen, 80-94.
- Grønvik, Oddrun 1997: Om kjeldegrunnlaget for Norsk Ordbok. I: Vikør, Lars S.: *Norsk Ordbok – nynorskens leksikografiske kanon?* Universitetet i Oslo, Institutt for nordistikk og litteraturvitskap, 23-37.
- Grønvik, Oddrun 2005: Redigeringsprogrammet for Norsk Ordbok 2014. I skjeringspunktet mellom menneskeleg skjøn og automatisering. I: Fjeld, Ruth Vatvedt, og Worren, Dagfinn (red.): *Nordiske studier i leksikografi 7. Rapport fra Konferanse om leksikografi i Norden Volda 20.-24. mai 2003*. Oslo, 156-165.
- Grønvik, Oddrun og Tvedt, Lars Jørgen 2006: Norsk Ordbok 2014 – presentasjon av eit komplekst leksikografisk verktøy. I denne rapporten.
- Hjulstad, Håvard 1986: Mikro FELTED. Feltinnskrivingssystem for mikromaskin. Universitetet i Oslo, Avdeling for namnegransking, 16 s.
- Svardal, Terje 2003: Norsk Ordbok – om normering av Metaordboka. I: Hansen, Zakaris Svabo og Johansen, Anfinnur (red.): *Nordiske studier i leksikografi 6. Rapport fra Konferanse om leksikografi i Norden Tórshavn 21.-25. august 2001*. Tórshavn, 351-356.
- Venås, Kjell 1989: Kva ventar vi oss av Norsk Ordbok? I: Almenningen, Olaf og Grønvik, Oddrun (red.): *Ord og mål. Festskrift til Magne Rommetveit*. Oslo: Kringkastingsringen, 153-166.
- Worren, Dagfinn 1992: Om feltinndelinga av ordartiklane i Norsk Ordbok. I: Vikør, Lars S. og Worren, Dagfinn: *Feltskjema for innskriving av ordmateriale*. Universitetet i Oslo, Institutt for nordistikk og litteraturvitskap, 18 s.

DET NORSKE LITTERÆRE ORDBOKSVERK I 25 ÅR, 1921—1946 (1922—1947).

Av Trygve Knudsen.

I.

I fjor (1946) var «Norsk Stadnamnarkiv» 25 år, og det fikk sin historikk av Per Hovda i siste dobbelhefte av «Maal og Minne» for 1946. Formelt er «Det norske litterære ordboksverk» like gammelt, søsterinstituttet som holder til under samme tak på Universitetsbiblioteket. Reelt begynte Ordboksverket sin tilværelse i første del av 1922. Ragnvald Iversen (den gang universitetsstipendiat) ble ansatt som sekretær 1. juli 1921, men oppholdt seg om høsten i København for å studere plan og metoder for det store danske ordboksarbeid, slik at han tok til med ekserperingen først i 1922. Derfor er det vel begrunnet å gi en oversikt i år, i 1947, over virksomheten i det første kvart århundre.

Som idé og plan var «Det norske litterære ordboksverk» godt og vel et halvt snes år da virksomheten ble satt i gang. Etter underhåndsdrøftelser i interesserte kretser fikk tanken første gang form i en uttalelse fra et privat møte 26. mai 1911. Her var tilstede professorene Gerhard Gran, Marius Hægstad, Halvdan Koht, Moltke Moe og Alf Torp, rektor M. Nygaard og adjunkt I. A. Refsdal, og noen dager senere gav professor Hjalmar Falk skriftlig sin tilslutning til den plan som var skissert: å ta opp snarest mulig arbeidet med å forberede «en videnskabelig fuldstendig ordbok over norsk litteratursprog». Arbeidet skulle knyttes til Det norske samlaget (Bymåslaget og Landsmåslaget i forening) og ledes av et utvalg oppnevnt av begge lag sammen. Dette arbeidsutvalg ble utpekt 11. november 1911 og kom til å bestå av deltagerne i mai-møtet supplert med dr. Amund B. Larsen og Didrik Arup Seip, og med Alf Torp som formann; medlem var fra først av også overlærer Steinar

Trygve Knudsen.

Schjøtt, som var forhindret fra å delta i møtene; rektor Nygaard døde kort tid etter oppnevningen. Komitéen arbeidet meget grundig, dels samlet, dels i særutvalg som fikk spesielle sider av oppgaven å tilrettelegge. I et avsluttende fellesmøte 28. oktober 1913 ble vedtatt i endelig form en ansøking til Kirkedepartementet om statstilskudd til den planlagte ordbok.

Her er oppgaven tegnet i hele sin bredde og storhet. Ordboken skal avspeile norske menns og kvinners litterære sprog fra begynnelsen av den nyere tid frem til nutiden, «opta hvert eneste ord som er blitt brukt i hele den norske litteratur i de sidste fire hundred aar.» Hvert ord skal få sin monografi, følges fra den eldste forekomst i ulike former og betydninger, og belyses med rikelige litteratursitater. Sammen vil disse sitater gi en prøvesamling fra ordbruken på skjønnlitteraturens forskjellige felter, i det religiøse sprog og på de mange fagområder, også innbefattet dialektlitteraturen. I ansøkingen er sterkt understreket den særlige rang dette tiltak har, hevet «over alle dagens mindre krav paa dette omraade». Utad vil en slik ordbok tre frem som et monument over vårt folks åndelige gjerning i nyere tid, på linje med de uttømmende ordbøker andre kulturfolk, store og små, har skapt eller holder på å skape over sine sprog. Innad vil et verk som det planlagte bli både et kart over den sproglige vei vi har gått og en retningsviser for utviklingen fremover; midt i alle dyptgripende motsetninger vil den nasjonale fremgangsvei tegne seg klarere enn før, linjen fra dansketidens dialektdiktning frem til det nynorske litteraturmål i det 19. og 20. århundre, og fornorskingslinjen i det skriftsprog som har sin tradisjon fra felleslitteraturen. Når 1914 ble valgt som startår for arbeidet, var det fordi man da kunne se tilbake på en stor avsluttet litteraturepoke, men ikke mindre fordi dette år markerte oppgavens samlende nasjonale verdi; den falt dessuten godt i linje med de samarbeidstanker som hadde vært fremme i sprogdiskusjonen i tiden nærmest forut. Oppgaven føyde seg også naturlig inn i andre store planer som nylig var blitt skissert for den nasjonale forskning. I sitt foredrag på det 2. norske historikermøte (9. juli 1912) om «Det historiske centralinstitut» hadde Oluf Kolsrud satt som en post på Instituttets arbeidsprogram «Historisk ordbok til det norske bokrike»; han henviste til de forberedende drøftelser som alt hadde funnet sted.¹

¹ Jfr. skriftet «Det historiske centralinstitut» (1912) s. 16 og 26.

Det norske litterære ordboksverk i 25 år, 1921—1946 (1922—1947).

I ansøkningsen til Kirkedepartementet ble lagt frem en detaljert plan for organiseringen av arbeidet. Driftsmidlene skulle sikres ved en årlig statsbevilgning på kr. 10 000,00 og ved årlige kvoter av honoraret (kr. 30 000,00) fra H. Aschehoug & Co. (William Nygaard), som hadde påtatt seg å forlegge verket. Det norske samlaget (Bymåslaget og Landsmåslaget) skulle være det forretningsmessige mellomledd mellom det offentlige og redaksjonskomitéen; denne komité skulle oppnevnes av Kirkedepartementet etter forslag fra de to styrer i Det norske samlaget, og selvstendig lede hele virksomheten med en hovedredaktør i spissen. Med dette økonomiske og faglige grunnlag for driften regnet man å ha fullført ordboken på maksimum 20 år; men man mente det lot seg gjøre å greie det på 15 år, og det var for dette tidsrom Arbeidsutvalget bad Staten om årlige bevilgninger. At man tenkte seg å kunne ta til med redigeringen forholdsvis snart, viser også den grundige drøfting av spørsmålet om målformen i ordbokens tekst. Et særutvalg (Hægstad, Torp, Falk) hadde behandlet dette problem og foreslått to alternative løsninger: to utgaver, én i hver målform, eller én utgave med redaksjonssprog fastsatt av statsmyndighetene (Stortinget); ansøkningsen bygger på det siste alternativ. Når Arbeidsutvalget straks pekte på disse mulige løsninger av dette nokså ømtålige spørsmål, var det naturligvis også for å hindre at det skulle volde vanskeligheter hos de bevillende myndigheter.

Det var da heller ikke strid om redaksjonssproget, og enda mindre uenighet om oppgavens høye nasjonale rang, som stengte veien for statsstøtte. Både hos Kirkedepartementet og i budsjettkomitéen var det finansielle betenkeligheter som skremte fra å binde Staten til et foretagende av dette omfang. I den korte stortingsdebatt om saken (april 1914) nevnte dessuten statsråd Bryggesaa at professor Torp, den utpekte hovedredaktør, enda en tid fremover kom til å være opptatt av arbeidet med sin nynorske etymologiske ordbok. Da Arbeidskomitéen gjentok ansøkningsen i 1914 og 1915, raste den første verdenskrig. Avgjørende for en utsettelse av starten ble det dessuten at professor Torp døde i 1916.

I 1920 ble ordbokssaken tatt opp på nytt grunnlag, men av et arbeidsutvalg hvor kontinuiteten fra den første komité var bevart: Halvdan Koht, Magnus Olsen, Didrik Arup Seip, I. A. Refsdal og Ragnvald Iversen. Med Magnus Olsen

Trygve Knudsen.

som talsmann lyktes det å sikre ordboksarbeidet plass på programmet for «Statens videnskapelige forskningsfond av 1919» som en av Fondets tre store humanistiske oppgaver, ved siden av stedsnavnstudier og undersøkelse av norske minner på de britiske øyer; i 1921 ble bevilget i alt kr. 14 000,00. Samtidig sa H. Aschehoug & Co. seg villig til å effektivere den støtte forlaget tidligere hadde stilt i utsikt, med en årlig sum på kr. 3 000,00 i 10 år. For disse midler ble det mulig å sette arbeidet i gang i 1922. Forskningsfondet gav nye bevilgninger i de nærmest følgende år. Men det var en uttrykkelig forutsetning at hjelpen fra denne kant bare skulle være til starten, «til planlegging og forberedelse». I 1923, 1924 og 1925 søkte derfor Ordbokskomiteén Kirke departementet om fast statsbevilgning til sikring av utvidet og kontinuerlig drift, kr. 25 000,00 i 15 år. Den humanistiske avdeling av Forskningsfondet støttet (i 1924) denne ansøking; den fremholdt sterkt at et tiltak av denne alment nasjonale og kulturelle art bør bæres økonomisk av Staten, og at Fondet selv i lengden vanskelig maktet byrden. Allikevel ble det Forskningsfondet som i hele 10 nye år stilte til rådighet midler til å holde virksomheten gående, inntil 1932 sammen med Aschehougs forlag. Det var disse to institusjoner som satte Ordboksverket i sålen og hjalp det til å ri det første lange og tunge stykke av veien.

I 1934 og 1935 mottok Ordboksverket en mindre sum (hvert år kr. 2 000,00) fra Nansenfondet til arbeid med spesielle oppgaver (utskrivning av Wergeland og Vinje). Og i 1935 fikk endelig Forskningsfondet full avløsning da Kirke departementet tildelte Ordboksverket kr. 10 000,00 av Pengelotteriets overskudd. Denne årssum har siden vært det økonomiske grunnlag for driften; i de siste år er beløpet utdelt av Departementet som statstilskudd.

Ordboksverket har hele tiden holdt til på Universitetsbiblioteket. I de første år fikk seddelesker og annet materiale plass i hyller og skap som Den Norske Avdeling velvillig avstod. Siden disponerte Ordboken i mange år et lite rom ved siden av bibliotekets hovedinngang. Fra 1942 har den hatt utmerket lokale i første etasje i den nybygde vestfløy.

Formann i Ordbokskomiteén har helt siden 1920 vært D. A. Seip. Ragnvald Iversen, som var den første medarbeider Ordboksverket knyttet til seg, har i alle år besørget ekserpering, i den lengste tid som eneste ekserptor. Få år etter starten ble Anne

Det norske litterære ordboksverk i 25 år, 1921—1946 (1922—1947).

Holtsmark og forfatteren av denne oversikt med i arbeidet; Anne Holtsmark har hatt mindre anledning til å delta i virksomheten etterhvert som ledelsen av Gammelnorsk Ordbok har krevd mer og mer av hennes tid. Blant andre filologer som gjennom flere år har vært medarbeidere i Ordboksverket, er å nevne Valborg Erichsen Lynner, Margarethe Kjær, Eva Fürst, Ragna Smith, Ingerid Gjøstein, Marius Sandvei, Kirsten Hansteen og Ludvig Holm-Olsen; de har særlig besørget avsluttende kontrollrevisjon og klassifisering av seddelmaterialet. For tiden utføres hovedparten av dette arbeid av Hallfrid Christiansen, Ingeborg Hoff og Vemund Skard. De to første holder på å gjennomarbeide den nynorske seddelmasse, med normalisering av oppslagsord og med henvisningsapparat for variantformer; tidligere har Eivind Vågslid ført sammen på sedler stoffet fra nynorske ordbøker og ordlister. Karen Seip, Anthone Boeck (fra 1933) og Alette Scheel (fra august 1945) har besørget den kontormessige administrasjon ved siden av sitt arbeid med avskrivning, seddelrevisjon og alfabetisering; sitatkontroll og ordning av materialet ble i mange år også utført av Alfild Slomann.

II.

I den plan Arbeidsutvalget la frem i sin ansøking høsten 1913 stod selve redigeringen av den vordende ordbok i forgrunnen som et forholdsvis nært mål. Det var nok også gjort klart rede for det materialgrunnlag man ville bygge på, en skatt av sitater fra de århundrer verket skulle spenne over. Men selve innsamlingen, tilrettelegningen og kontrollen av dette stoff var viet mindre oppmerksomhet. Man regnet med å greie dette grunnarbeid på relativt kort tid og med beskjedne utlegg (kr. 12 000,00 pr. år, iberegnet lønn til hovedredaktør). Det er vel ikke tvilsomt at både tidsfristen og budsjettet ville blitt mange ganger overskredet om tiltaket var satt ut i livet på denne basis.

Da Ordbokskomiteén søkte om statsstøtte i 1923, så den arbeidet i et noe annet perspektiv. Som i 1913 ønsker komitéén en bevilgning sikret i 15 år. Men den ber om den dobbelte sum, kr. 25 000,00, og den gir intet løfte om tiden for fullføringen av den planlagte ordbok; derfor innlater den seg heller ikke på å drøfte spørsmålet om redaksjonssproget. Det er den første del av

Trygve Knudsen.

oppgaven, materialinnsamlingen, komitéen vil sikre midler til på lengere sikt; det igangværende arbeid med den store danske ordbok hadde vist at man trenger milliontall av sitatsedler for å kunne gå trygt i vei med redigeringen, og i de par år det norske ordboksverk hadde vært i virksomhet, hadde man lært hvor tidkrevende oppbyggingen av en fyldig og pålitelig materialsamling er. Samtidig satte komitéen videre mål for denne samling. Den pekte på at verdien av den sitatmasse som skulle skapes, ikke måtte være bundet ensidig til den vordende ordbok. Det man fra første stund måtte sikte mot, var et mest mulig fullstendig ordarkiv på sedler for litteratur av alle slag etter 1500, et arkiv med selvstendig eksistensrett både før ordboken er blitt virkelighet og etterat den foreligger fullført; bare en mindre del av sitatstoffet vil komme med der, og innsamlingen må gå videre ut over den tidsgrense ordboken stanser ved. Steg for steg må Ordboksverket for sitt spesielle arbeidsområde — ordforråd og ordbruk i nyere tid — bli et sentralsted, hvor en kan finne stoffet uten møysommelig og sinkende leting i mange kilder og uten den utrygge følelse at en går glipp av viktig materiale.

Skal sitatmassen være noe mer enn bare tilfang til éngangsbruk for en ordbok, stilles der særlige krav til den. Den må være fysisk holdbar, helst uforgjengelig, så langt det står i menneskelig makt; derfor har Ordboksverket hele tiden brukt sedler av utvalgt papirkvalitet og syrefritt blekk. Både for ordbokens og ordarkivets skyld må materialet fra de utskrevne verker ikke bare være representativt, men fyldig. Det er ikke nok at hele ordforrådet er med; sitatmassen må registrere ordenes betydningsnyanser, deres faste fraseologiske forbindelser, deres eldste (og eventuelt yngste) kjente forekomst, og helst også gi et inntrykk av hyppigheten i bruk til ulike tider, hos ulike forfattere og på ulike fagområder. Er tekstmaterialet fast avgrenset og overkommelig, vil en sitatutskrivning etter thesaurusprinsippet dekke alle disse krav på ideell vis. For Ordboksverket har denne vei ikke vært farbar. Arbeidsområdet er for veldig, og den store grunnstamme av felles ordforråd innenfor de forskjellige faggrupper og perioder gjør heller ikke en slik innsamlingsmåte nødvendig; bare ganske få spesielle tekster har vært og vil bli skrevet ut thesaurus-messig. Praktisk talt hele seddelkartoteket er da blitt til ved utvalg (ekserpering) av ordforrådet i de enkelte verker, men et utvalg foretatt av ordboksvante fagfolk etter de retningslinjer som

Det norske litterære ordboksverk i 25 år, 1921—1946 (1922—1947).



Et hjørne av Ordboksrommet med to av kartotekskapene.

ovenfor er nevnt. Tekstene er gjennomtrålet med et nett så finmasket at der skulle være nesten full visshet for at ikke bare alle ord, men også alle betydninger og faste vendinger er fanget inn. Seddelarkivet er minst av alt en kuriositetssamling: de «alminnelige» ord, bl. a. formord av alle slag, er rikelig belyst i sine ulike funksjoner, ofte med hundretall eller tusentall av sedler. Det er et hovedprinsipp som aldri fravikes, at hvert ord presenteres på sin seddel ikke alene med sidehenvisning, men i et sitat som gir hele konteksten, og som danner en avsluttet mening, om nødvendig med hjelp i forklarende innskudd; eventuelle andre opplysninger som teksten selv gir, f. eks. i fotnoter, kommer også med. Hvor materialet gjør det mulig, er dessuten registrert, som innskudd i sitatene, synonyme ord som er brukt i andre utgaver eller i varianter av samme tekst, også under selve utformningen slik den kan følges i forfatterens forarbeider og rettede manuskripter. Målet er å samle på hver seddel alle de opplysninger det utskrevne tekststykke byr, ut fra et ideelt krav til kvalitet, men også ut fra det praktiske syn at på lang sikt er denne fyldighet både rasjonell og tidsbesparende; ordboksredaktørene og de mange andre som kommer til å bruke materialet, skal bare i rent sjeldne

Trygve Knudsen.

tilfelle behøve å slå opp på det trykte sted seddelen henviser til. Etter samme grunntanke legger Ordboksverket stor vinn på å trygge påliteligheten i sitatavskriftene i alle enkeltheter; hver skreven seddel revideres to ganger.

De store krav til fyldighet og pålitelighet kan selvsagt ikke tilfredsstilles uten økonomiske offer; når en har en trang bevilgningsramme å røre seg i, betyr det at det en setter inn på kvalitet, må en avskrive i fart. Til de løpende utgifter til ekserpering, avskrivning, revisjon og alfabetisering kommer med visse mellomrom dryge uttellingar til seddelkjøp og kartotekskap; en fast årlig utgift er husleien til Universitetet. Det årsbeløp Ordboken i størstedelen av tiden har hatt til å balansere utgiftsiden, kr. 10 000,00, må en nesten kalle mikroskopisk når en sammenligner det med de midler som står til rådighet for de store tilsvarende ordboksforetagender i Danmark og Sverige. På denne bakgrunn tør en vel si at arbeidsresultatet for Det litterære ordboksverk i de første 25 år har vært respektabelt. Den samlede seddelmasse (reviderte og alfabetiserte sedler) er i dag godt oppe over millionen, omkring 1 100 000; nesten halvdelen av seddelmengden, omkring 450 000, er på nynorsk og bygdemål. En fortegnelse over de ekserpererte verker følger sist i denne artikkel.

Ordboksverket har hatt den glede å se sine samlinger flittig brukt.

Først og fremst har de naturligvis vært til nytte for ordbøker over litteratursprog i Norge i gammel og ny tid. Den Ibsen-ordbok som er i arbeid, og som skal avslutte den store hundreårsutgave av hans verker, bygger utelukkende på materialet i Ordboksverkets skuffer. Med «Nynorsk ordbok» har der vært stadig kontakt, slik at ekserperingen av landsmålstekster delvis har fulgt ønsker fra redaksjonen av denne ordbok og vært samordnet med utskrivningsarbeidet der. «Norsk riksmålsordbok» har sitater og henvisninger fra Ordboksverkets seddelarkiv i hver eneste spalte, ja nær sagt i hver eneste større artikkel; særlig har Ibsen-samlingen vært til uvurderlig hjelp, med en rikdom av stoff fra et forfatterskap som strekker seg over et halvt århundre og favner over nesten alle stilarter.

Det har vist seg at Ordboksverkets materiale har interesse for mange andre enn bare for ordboksarbeidere. Seddelsamlingene har vært brukt av studenter som stoff for hovedfagsavhandlinger,

Det norske litterære ordboksverk i 25 år, 1921—1946 (1922—1947).

bl. a. til studier i Ibsens sprog, ved begrepshistoriske og ordhistoriske undersøkelser, av folkeminneforskere. Ikke sjelden har jurister tydd til Ordbokens sitater for å få fastslått innholdet av omstridte ord og uttrykk i rettslige dokumenter. Stadige sproglige spørsmål fra institusjoner, aviser og privatpersoner har Ordbokens medarbeidere kunnet besvare med hjelp av det materiale de har under sine hender, lettere og sikrere enn det ellers ville vært mulig.

Tross det i og for seg høye seddeltall er den kunnskap Ordboksverket kan gi, ennå begrenset. Kartoteket må opp i det dobbelte eller tredobbelte omfang før en kan regne med at det avspeiler litteraturen etter 1500 så fyldig som en må kreve; så stort må materialet også være for den planlagte historiske ordbok. Det sier seg da selv at det høyeste ønske for Ordboksverket idag er å komme ut av den trange økonomiske ramme det hittil har måttet virke i. Det har hatt en tilværelse nærmest på eksistensminimum, med midler som er forsvarlige til forberedelse og start, men ikke til å drive en institusjon som for lenge siden er ute over forsøksstadiet, og som står fullt rustet med trente medarbeidere til å utvide virksomheten til rasjonelt omfang. I den aller siste tid har Ordboksverket prøvd nye metoder i sitt arbeid, fotostatering istedenfor avskrivning av tekster; begynnelsen er gjort med Jonas Lies verker. Holder fremgangsmåten hva den synes å love, vil den i sterk grad forenkle arbeidet og dermed øke farten. Men for å bli fullt effektiv forutsetter den nye ordning at større tekstpartier fotostateres samtidig; den vil derfor gjøre det til et enda mer påtrengende krav å trygge en bredere økonomisk basis. Det samme gjelder stensileringsmetoden, som også byr lokkende muligheter, og som også skal prøves; Gammelnorsk Ordbok bruker den med gode resultater.

Ordboksverket har flere ganger søkt om en fast årssum på kr. 25 000,00. Med en slik budsjetttramme og med bruk av nye tekniske hjelpemidler ville den årlige økning av seddelmengden sikkert kunne mangedobles; det kunne f. eks. bli mulig å virkeliggjøre den gamle plan om å registrere ordforrådet i våre eldste aviser. Samtidig ville der også kunne bli større rom for andre arbeider: fortsatt fagkyndig behandling av seddelmaterialet, bl. a. med ordning av sitatene på de enkelte ord etter betydninger, dessuten utarbeidelse av fullstendige seddelhenvisninger for alle stikkordenes suffikser. På denne måte vil seddelmassen få sterkt øket verdi

Trygve Knudsen.

for alle som bruker den, samtidig som en slik systematisk tilrettelegning i virkeligheten er første steg i redigeringen av den vorde ordbok.

Ved å bygge opp sitt materiale som et permanent seddelarkiv har Ordboksverket sikret det mangesidig verdi, uavhengig av spesielle formål. Men institusjonen ville svike den oppgave den er grunnlagt på, og som har gitt den navn, om den stadig vedblir å skyve fra seg tanken på å nyttiggjøre samlingene i en stor historisk ordbok (eventuelt også i flere ordbøker, for mindre tidsavsnitt og for visse enkelte forfattere). Ordbokstanken har i skiftende form vært fremme hos sproginteresserte nordmenn flere ganger i løpet av mer enn to hundre år før den fikk fast skikkelse i 1911; ingen skal kunne si at den ikke er moden til å virkeliggjøres. For beskjedne midler vil en nå kunne øke tilveksten i seddelsamlingene og tilrettelegge stoffet så raskt at ordboken om ikke altfor mange år kan rykke inn på det aktuelle arbeidsfelt. Den må ha sin sikre plass på programmet for uavviselige kulturelle tiltak, like selvfølgelig som planen om en ny og tidsmessig gammelnorsk og mellomnorsk ordbok. Sammen vil disse verker bli sproglige skattkammer og veiledere for hele den periode da litteratur har vært skrevet og trykt i Norge.

Tekster som er utskrevet til seddelsamlingene.

Bokmål.

Absalon Pederssøn: Om Norgis Rige (G. Storms utg.).

Peder Claussøn Friis: Samlede Skrifter (G. Storms utg.).

Petter Dass: Samlede Skrifter. — Dassiana (A. E. Erichsens utg.). Materialet suppleres med nye lesemåter og varianter fra T. Hannaas' utgave av «Den norske Dalevise» og Seips utgaver av «Nordlands Trompet» og av «Viser og Rim».

«Topographisk Journal» hefte 1—4, 7, 8, 10, 16, 17, 20, 21.

«Det Norske Nationalblad» 1816—1821 (en del nummer mangler).

«Saga» I og III.

Henrik Wergeland: Samlede Skrifter (Jægers og Seips utg.). I, 1.

Henrik Ibsen: Alle dikterverker; dessuten utkast, forarbeider og endringer i de forskjellige utgaver, etter den store hundreårsutgave. Fra denne utgave er også utskrevet artikler og taler, foruten brev (så langt de er utkommet).

Bjørnstjerne Bjørnson: Samlede Digterverker, Samlede Digte

Det norske litterære ordboksverk i 25 år, 1921—1946 (1922—1947).

(begge i Francis Bulls utgave, og begge med varianter). Dessuten «Kong Eystejn» og brevsamlingene «Gro-tid», «Brytningsår» og «Kamp-liv» (alle ved Halvdan Koht).

Fra eldre og yngre aviser ca. 7000 sitater som skyldes tilfeldig utskrivning av Ordboksverkets medarbeidere.

Under arbeid er Tullin: Samtlige Skrifter (1770—73), og Jonas Lie: Samlede Digterverker (standardutgave 1920).

Nynorsk.

For å samle de spredte leksikalske opplysninger er Aasens og Ross' ordbøker (med tillegg) klippet opp og klebet på sedler (ett stikkord på hver seddel). På samme måte er hittil innordnet i seddelkartoteket tilfanget fra andre ordbøker og ordsamlinger: «Ældre norske sprogminer» II, III, IV (Ordsamling fra Røbyggjelaget, Christen Jensøns Dictionarium, Knud Leems samlinger), Hallager, Schytte (Lofoten), Vidsteen, Skulerud (ordlister i festskf. til Torp og i «Studier tillågnade Axel Kock»), Hveding (håloygsk ordsamling), Opedal (Hardanger, ordliste i «Makter og menneske»), Rise (Opdal), Skjørvo (Vik i Sogn), ordlister i «Gamalt frå Vestfold» (I og II).

A. O. Vinje: Dølen. — Diktsamling (1864). — Blandkorn. — Skrifter i samling.

Henrik Krohn: Ei liti Ferd. — Svein og Gudveig.

Christopher Janson: Fraa Bygdom.

O. J. Fjærtøft: Folkelesna 1—3. — Ignatius (oversettelse). Fedraheimen 1877—78. Utskrivningen fortsetter.

Johannes Skar: Gamalt or Setesdal I—VIII.

Ivar Kleiven: I Heimegrendi. — Lom og Skjaak. — Lesja og Dovre. — Østre og Vestre Gausdal. — Ringbu. — Fronsbygdin.

Knut Loupedalen: Eventyr og Segnir fraa Telemarki.

Denne oversiktsartikkel bygger på materiale i Ordboksverkets arkiv. En del ansøknings og årsmeldinger har vært trykt i «Maal og Minne» (1914, 1915, 1924, 1925, 1928, 1929, 1930). En kort oversikt over Ordboksverkets plan og arbeidsmåte er gitt i «Norsk tidsskrift for sprogvidenskap» III, 366—369.

Digital sats eller digital satsing?

Presentation or content? The Unit for Digital Documentation at the University of Oslo has been making dictionary writing systems (DWS) for more than ten years. In this paper we will present some of our ideas on how a modern DWS should deal with the source material, other dictionaries and cross references inside the dictionary, and what implications this will have on the implementation of such a system.

1. Innleiing

1.1. *Om Eining for digital dokumentasjon*

Eining for digital dokumentasjon (EDD) ved Universitetet i Oslo er direkte underlagt det Humanistiske fakultet (HF) og har som mål å hjelpe fageiningane ved HF med å lage digitale dokumentasjonssamlingar, tekstsamlingar, korpus og andre databasar.

EDD er eit produkt av Dokumentasjonsprosjektet som var eit stort samarbeidsprosjekt mellom dei humanistiske fakulteta ved dei fire universiteta i Noreg. Oppgåva til Dokumentasjonsprosjektet var å digitalisere dei store samlingsarkiva ved dei ulike HF-einingane.

Etter at Dokumentasjonsprosjektet vart avslutta, vart det naudsynt å etablere ein driftsorganisasjon som skulle vedlikehalde og vidareutvikle databasane som vart bygd opp i prosjektperioden, samt å utvikle nye verktøy for brukarane. EDD vart oppretta som ei permanent eining under HF. EDD har i tillegg til dei faste oppgåvene for HF også ansvaret for gjennomføringa av Museumsprosjektet, eit nytt stort nasjonalt digitaliseringsprosjekt som skulle bygge på dei røynslene som var gjort under Dokumentasjonsprosjektet.

Som ei følgje av dette har EDD ein forholdsvis liten fast tilsett stab som har hovudansvaret for oppgåvene som ligg til HF, og ein stor prosjektstab knytt til Museumsprosjektet. Trass i dette skiljet, utviklar EDD bevisst løysingar som er felles for HF og Museumsprosjektet, og ein prøver å utnytte kompetansen mellom dei to stabane til beste for begge partar.

1.2. *Leksikografi*

Arbeidet med dei leksikalske arkiva har vore ein sentral del av EDD sine oppgåver heilt frå tida i Dokumentasjonsprosjektet. Setelarkivet til Norsk Ordbok var mellom dei første arkiva som vart digitalisert tidleg på 1990-talet, og skanning og registrering av 3 millionar setlar tok mykje ressursar.

EDD var tidleg ute med å lage løysingar for å bruke dei leksikalske arkiva, og

utvikla også tidleg applikasjonar for redigering av ordbøker. Taggarprosjektet som var eit samarbeid mellom Tekstlaboratoriet ved HF og EDD, utvikla tidleg ein taggar basert på data frå Nynorskordboka og Bokmålsordboka. EDD laga også verktøy for redigering av dei første einspråklege ordbøkene i Zimbabwe (ALLEX-prosjektet), og har også stått bak redigeringsystema for Nynorskordboka og Bokmålsordboka. I dag jobbar EDD mellom anna med redigeringsystemet til Norsk Ordbok, og eininga har utvikla integrerte korpussystem knyta til redigeringsapplikasjonane basert på CQP/Corpus Workbench.

1.3. EDD sin filosofi

Gjennom arbeidet med samlingane har EDD fleire gonger påvist små og store manglar både ved måten arkiv er organisert på, og ved innhaldet i arkiva. Arkiva har ofte vore utforma på ein slik måte at dei berre kan brukast av ei lita gruppe menneske, og ikkje sjeldan er det store problem med oppdatering av arkiva. Ofte kjem ein over arkiv der tidlegare tolkingar av eit materiale vert kasta når det vert gjort nye tolkingar. Andre arkiv kan vere organisert slik at det ikkje er rom for nytolkingar. Eit døme på det siste er setelarkivet ved Norsk Ordbok som i papirutgåve var sortert etter ei normalisert skriftform basert på 1938-rettskrivinga av nynorsk.

Dei første digitale arkiva EDD laga, var nok på mange måtar prega av dei same problema, til dømes i høve til det å kunne ha rom for nytolking av materialet. Etter kvart har EDD lagt stor vinn på å forbetre dette, og ein har i fleire år vore aktiv deltakar både i nasjonale og internasjonale fora der ein prøver å lage standardar for datautveksling, samt diskutere løysingar på ulike problem knytt til forskingsdatabasar.

EDD har på bakgrunn av dette lagt vinn på å utvikle arkivløysingar som på alle vis kan karakteriserast som vitskaplege. Med dette meiner vi mellom anna at ein må lagre grunnlagsmaterialet i så lite bearbeidd form som mogleg. Dei mest atomære storleikane i eit arkiv bør vere så fri for tolking som mogleg. Dernest må alle hypotesar som er basert på dette grunnlagsmaterialet handsamast med like stor respekt. Arkivsystema må ta omsyn til at nye tider og nye forskarar vil ha nye tolkingar av kva som ligg i eit materiale. Og sist, men kanskje viktigast, arkivsystema må kople saman grunnlagsmaterialet og hypotesane slik at det er enkelt for nye forskarar å etterprøve resultat.

Ut over dette har EDD ønske om å kunne knytte saman informasjon i ulike arkiv slik at forskarane kan dra nytte av kvarande sine resultat i langt større grad enn i dag.

2. Produksjon av tradisjonelle ordbøker

Produksjon av ordbøker vart tidleg ei viktig oppgåve for EDD, og EDD har vore med på å flytte fleire store prosjekt inn i framtida. Norsk Ordbok er absolutt det største, og vi vil bruke denne ordboka som eit døme på korleis dette har gått til.

2.1. Fokus på utsjånad

Før EDD kom inn i biletet, vart artiklane i Norsk Ordbok skrivne i WordPerfect og Word. Ein nytta eit sjølvutvikla kodesystem for å merke opp teksten, og all oppmerking hadde som mål å styre utsjånaden på trykk.

Problemet med denne framgangsmåten er at ordboka sjølv ikkje gjer det lett for ein brukar, eller ein forskar, å etterprøve informasjonen. Rett nok har ein i alle vitenskaplege ordbøker eit referansesystem til litteratur eller arkivmateriale som underbygg påstandane, men reelt sett er det ikkje mogleg for ein brukar å sjekke desse kjeldene med mindre brukaren sit i arkivet til ordboksredaksjonen. Sjølv om ein skulle sitte i arkivet, ville det vere omtrent umogleg å finne fram til dei setlane som har danna grunnlaget for ein artikkel, og ikkje minst finne ut kva setlar redaktøren ikkje brukte. Ordboka ville med andre ord formelt sett tilfredstille vitenskaplege krav, men reelt sett ville ho i avgrensa grad bli utsett for vitenskapleg kritikk fordi det ville vere eit alt for stort arbeid å gå til kjeldene.

Frå redaksjonen sin synsstad har også denne gamle produksjonsmåten mange ulemper. Mellom anna er handteringa av krysstilvisingar mellom artiklar eit evig problem. Det er ingenting som garanterer at ein definisjon som ein viser til, faktisk eksisterer. Stringens i form og innhald vert også meir komplisert enn det ein gjerne vil. Det einaste som styrer dette er redaktørane sin godvilje, og redaksjonen sin vilje til å overprøve einskildredaktørar sin måte å skrive artiklar på. Handtering av mange redaktørar vert også vanskeleg. Dess fleire redaktørar, dess vanskelegare vert det å halde styr på dei to føregåande punkta.

Gjenbruk av data frå ordboka er heller ikkje særleg lett med denne måten å redigere artiklar på. Det er svært lite struktur i innhaldet, og den som måtte ønskje å bruke ordboka til noko anna enn å slå opp definisjonar, måtte bruke mykje tid på å finne det ein leita etter, trass i at teksten var tilgjengeleg digitalt.

2.2. Fokus på struktur

Ut over 1990-talet vart det opplagt for dei fleste at dei gamle måtane å redigere ordboka på, ikkje gav samfunnet nok attende i høve til innsatsen. Krav til tilgjenge for ålmenta, gjenbruk av data, vitenskapleg kvalitetssikring, og ikkje minst effektiv produksjon av ordboksmanuskript, sette fart i planane om å endre måten Norsk Ordbok vart redigert på. Alle krava ein stilte til den nye ordboka, peika i retning av meir fokus på struktur, mindre på utsjånad, samt at det måtte vere mogleg å kople ordboka sine artiklar mot grunnlagsmaterialet, og at ein trong administrative rutinar

bygd inn i systemet for å auke effektiviteten. I dag er det to teknologiar som peiker seg ut, strukturert tekst (SGML/XML) eller ei eller anna form for databaseløysing.

2.2.1. SGML/XML

EDD har lang erfaring i bruk av strukturert tekst. Sidan starten av 1990-talet har EDD arbeidd med SGML i oppmarkeringa av tekst. Då XML overtok for SGML, vart dette ein teknologi som vart tilgjengeleg for eit større publikum, og absolutt ein mogleg teknologi å bruke i utforminga av ordbøker.

Den største styrken til XML som format for eit ordboksmanuskript er kombinasjonen av stringens og fleksibilitet. XML gjer det mogleg å lage klare reglar for korleis strukturen skal vere, men kan tillate stor grad av valfridom i innbyrdes plassering av tekstelement. XML kan også på ein enkel måte opne for at einskilde element kan dukke opp på fleire plassar i eit dokument utan at strukturen lir under dette.

Det største problemet med XML er at dette ikkje er eit databaseformat. I utgangspunktet er eit XML-dokument ei samling tekstar som alle har fast struktur. For store dokument vil ein rein XML-struktur gjere referanseintegritet mellom delar av dokumentet tungt og tidkrevjande, og kopling til element som ligg utanfor dokumentet sjølv vil vere enno tyngre. Med XML er det ikkje mogleg å garantere at referansar til grunnlagsmaterialet i form av til dømes ordsetlar, er korrekte. XML kan heller ikkje hjelpe til med administrasjonen av ein stor redaksjon. Eit større ordboksprosjekt er avhengig av desse funksjonane.

2.2.2. Databaseløysingar

Hovudalternativet til XML er å lagre artikkelteksten i ein database av eit eller anna slag. Databasen sin største styrke er innebygde rutinar for dataintegritet. Dataintegriteten gjer det mogleg å sikre at krysstilvisingar til ein kvar tid er korrekte, at kopling til kjeldematerialet er på plass, osb. Databasesystem utan slik integritetskontroll er til lita nytte for eit ordboksprosjekt.

Databasen tilbyr struktur, men kan tidvis også stille for store krav til struktureringa av artikklane i ei ordbok. Dette er den største innvendinga mot å bruke ein database som basis for sjølv artikkelteksten. Databasen har ikkje innebygde system for å markere at delar av ein laupande tekst har ei spesiell mening.

2.2.3. Kombinasjonsløysingar

Konklusjonen er at korkje ei rein XML-løysing eller ei rein databaseløysing fungerer for dei fleste større ordboksprosjekt. Ein må ty til kombinasjonsløysingar.

Det finst mange verktøy som i ulik grad løyser problema med XML ved å innføre ein eller annan form for databasestruktur rundt XML-dokumenta, og dette kan ofte vise seg svært fruktbart. Ofte nyttar ein XML som internt databaseformat i eit slikt

databasesystem, og dette kan gjere desse databasesystemet ope og lett å dokumentere, men dette er på langt nær opplagt. Mange av desse systema er også svært nye og lite utprøvde, dei er laga for mindre organisasjonar, og skalerer ikkje godt i store miljø.

Mange databasesystem i dag kan handtere XML eller XML-fragment i tekstfelt i databasen. Dette løyser litt av problema knytt til databasen sine ekstreme strukturkrav.

Om ein vel å basere seg på ei utvida XML-løysing, eller ei utvida databaseløysing, kjem an på kva eigenskapar ein meiner er viktigast. Eit spesialutvikla databasebasert system vil vere mest teneleg dersom ein skal ha store redaksjonar og god kontroll med referansar og tilvisingar i ordboka. Men dette krev at ein kan definere ein streng struktur for artiklane utan at dette går ut over den faglege kvaliteten.

Ei mindre ordbok med ein mindre redaksjon og mindre krav til stringens, vil tene på å bruke eit enklare, meir generelt XML-basert verktøy.

3. Den nye ordboka

I dei føregåande avsnitta presenterte vi dei krava vi meiner ein må stille til verktøy for produksjon av dei tradisjonelle ordbøkene. Men teknologien fører ikkje berre til at ein kan gjere tradisjonelle oppgåver på ein betre måte. Det kan like gjerne føre til at oppgåva eller det endeleg resultatet totalt endrar karakter. Tenkjer ein visjonært kan ein ofte med same arbeidsinnsats, få eit anna og mykje meir nyttig resultat. Dette ser ein tydeleg innanfor mange humanistiske fag i dag, og ikkje minst innanfor språkforskinga.

Teknologien kan og gje eit omgrep eit heilt nytt innhald, og innhaldet knytt til omgrepet endrar krava publikum stiller. Dette ser ein tydeleg i samband med bøker generelt, og ordbøker spesielt. I Norsk Ordbok, band 1, som vart utgjeve i 1966, er ordet "bok" definert som ei samling innbundne, prenta ark. I dag kan vi lese på web-sidene til Norsk Ordbok at ein har som mål å kome med ei elektronisk utgåve av ordboka. Både Norsk Ordbok, men også store delar av det norske folk, meiner at ei samling av ord og definisjonar som ein kan finne på internett, er ei ordbok. I dagleg tale er det ikkje lenger uvanleg å høyre om elektroniske bøker, eller lydbøker. Teknologien har gjort at ei bok i dag er noko heilt anna enn det ho var for 40 år sidan, og våre krav til bøkene speglar teknologien det er mogleg å bruke for å presentere innhaldet av boka. Den beste ordboka er ikkje berre den med dei mest korrekte definisjonane, men også den som utnyttar den eksisterande teknologien best til å formidle innhaldet i definisjonane. Leksikografen har ikkje gjort jobben sin dersom han ikkje har utnytta teknologien optimalt når han formidlar kunnskapen sin om eit ord.

Dette legg eit stort press på dagens leksikografar når dei skal lage vitskaplege ordbøker. Frå Gutenberg og fram til 1990 måtte leksikografen ta stilling til

sidestorleikar, skrifttypar, spalter, osv. Dette var nokre av dei parametrane som styrte formidlinga. I tillegg hadde formidlingskanalen klare grenser for kva ein kunne ta med. Mangel på teknologiske løysingar førte til at ein sjølvstekt ikkje kunne prente alt kjeldemateriale, og ein kunne heller ikkje legge ved lyd eller video. Dersom ein kunne gjort dette ville den vitskaplege verdien av ordboka auka.

I dag set ikkje teknologien slike grensar. Det er fullt mogleg å lenke oppslagformer og dialektformer direkte til lydkjelder. Definisjonar og døme kan peike til tekst i eit korpus og ein kan legge inn formell syntaktisk informasjon i tillegg til definisjonar og døme. Det er mogleg å kople mot kjelder for ensyklopedisk informasjon, og ein kan legg inn informasjon om bøyning av ord, informasjon om samansetting av ord, osv. Tekst, lyd, bilete og film i uavgrensa mengde kan knytast til artikkelane. Leksikografen si oppgåve blir å organisere all denne informasjonen slik at det er mogleg å finne fram i han.

Utfordringa er at brukarane av ordboka no venter å finne denne typen informasjon. Forskarar vil døme ordboka etter kor lett det er å etterprøve informasjon ved å følgje ordboka sin eigen dokumentasjon av påstandar. Oppdragsgjevarar og andre venter at det skal vere mogleg å trekke informasjon ut av ordboka, og bruke dette til talesyntese, maskinomsetting og dikteringssystem. Når det er teknisk mogleg å lage ei slik ordbok/leksikalsk database, så vil alle forvente at dette vert gjort.

Vi veit alle at det ikkje er mogleg å lage denne altomfattande databasen i dag, og sannsynlegvis vil ein heller aldri nå heilt i mål. Men det er ei plikt for oss å skjele til det som ligg i dagens og morgondagens teknologi når vi lagar løysingar. Ein kan ikkje forsvare å lage ordbøker i dag som ikkje kan inngå i ein større heilskap ved eit seinare høve. Det vil vere ein hån både mot fagmiljø, brukarar og dei som betalar for arbeidet.

4. Vegen mot vitskaplege leksikalske ressursar

Er det så i praksis mogleg å lage slike ressursar på eit slikt vis at alle vert nøgd? EDD har forsøkt, og trass i einskilde feilskjer undervegs, har vi kome eit godt stykke på veg.

Men utfordringa ligg ikkje primært i å lage gode system for spesifikke problem. EDD og Norsk Ordbok har i fellesskap skapt fleire ulike databasar som på ulikt nivå tilfredsstillar dei krava vi har sett til det vi kallar vitskaplege databasar. Dess enklare basane er, dess lettare er det å oppfylle desse krava. Databasane over setelarkiva til Norsk Ordbok er døme på banale basar i denne samanhengen. Basane er berre ein digital representasjon av gamle papirbaserte arkiv.

EDD har også utvikla databasane som Norsk Ordbank mellom anna brukar til å registrere bøyingsinformasjon for meir enn 100.000 grunnformar for dei to norske målformene, bokmål og nynorsk, men heller ikkje her ligg det særleg referanse til dømes til empiriske data.

EDD har også utvikla fleire ulike redigeringsapplikasjonar for spesifikke ordbøker. Dei fleste av desse er applikasjonar som ikkje gjer anna enn å registrere den teksten redaktøren vil ha på trykk, og på det viset framstår dei, etter våre krav til vitskap, ikkje som vitskaplege databasar. Det ligg lite eller ingen informasjon som gjer det mogleg å gå til kjeldematerialet.

Databasen for Norsk Ordbok har vi derimot forsøkt å bygge på ein slik måte at alle våre krav til vitskaplege metoder er oppfylte.

Metaordboka er også ein svært enkel konstruksjon, som enkelt forklart er ein indeks til alle dei andre leksikalske ressursane EDD og Norsk Ordbok rår over. Men sidan vi har brukt Metaordboka til å kople saman mange ulike arkiv, ordbøker og ordlister, står samlinga av mindre og enkle databasar i større grad fram som eit vitskapleg heile. Det er i dag mogleg å gå frå ein artikkel i Norsk Ordbok, via Metaordboka, til Norsk Ordbank, og få ut alle brukte bøyingsformar av ordet. Ein kan også gå frå handordbøkene, Nynorskordboka og Bokmålsordboka, via Metaordboka til Norsk Ordbok, og der få utfyllande informasjon om bruksmåtar, dialektformer, osb., og sjølvstgt vidare derifrå til å sjå på kjeldematerialet, anten i form av ordsetlar, eller i form av konkordanseliner henta frå eit korpus. Samankoplinga gjer det enkelt å navigere mellom dei ulike databasane, og dermed står dei einskilde basane fram med mykje større vitskapleg tyngde enn før.

5. Samarbeid

Frå dette forholdsvis enkle systemet for samankopling av data, ser vi at det er svært mykje å hente ved å kople saman ressursar. Vi har så langt prioritert å kople saman dei ressursane vi sjølv rår over. Vi har også planar om å legge inn eit større tal dialektordbøker i dette systemet, for ytterlegare å auke nytta av databasane.

I framtida ser vi at det kunne vore nyttig å vurdere alle dei store ressursane i Norden med omsyn på ei slik samankopling. Det er svært mange tunge og viktige databasar i Norden, og ved å kople dette saman vil sannsynlegvis leksikografar i heile området få langt meir ut av også sine eigne databasar.

Det er klart at dette vil krevje ein god del avklaring, både på den tekniske sida, men også når det gjeld spørsmål om opphavsrett, kommersielle rettar, osb. Vi ønskjer likevel å kome i gang med dette arbeidet, spesielt på den tekniske sida. Det hadde difor vore ønskjeleg at personar med teknisk kompetanse frå dei ulike fagmiljøa i Norden kunne samlast for å diskutere korleis ein skal gå fram for å utvikle metodar for utveksling av data. Eit mål må vere at ein slik samling blir gjennomført før neste konferanse om leksikografi i Norden.

Litteratur

- Ore, Christian-Emil Smith 1998a: Hvordan lage fagdatabaser for humanistiske fag. I: *Fra skuff til skjerm. Om universitetenes databaser for språk og kultur*. Oslo: Universitetsforlaget. ISBN 82-00-12670-6. 30 s.
- Ore, Christian-Emil Smith 1998b: Making multidisciplinary resources. I: Lou Burnard, Marilyn Deegan and Harold Short (eds.) *The Digital Demotic, A Selection of Papers from Digital Resources in the Humanities 1997*. Publication 10, Office for Humanities Communication, King's College, London, 1998, ISBN: 1-897991-12-7.
- Ore, Christian-Emil Smith 2001: Metaordboken – et elektronisk rammeverk for Norsk Ordbok? I: Martin Gellerstam et al. (eds.): *Nordiska studier i leksikografi 5. Rapport från Konferens om Lexikografi i Norden, Göteborg 26.-29. maj 1999*. Göteborg, 202-216.
- Ore, Christian-Emil Smith; Tvedt, Lars Jørgen; Bjørnstad, Tone 2002: *The Meta Dictionary ALLC/ACH 2002*; 24.07.2002-28.07.2002.
http://www.edd.uio.no/artikler/leksikografi/meta_dictionary.html
- Text Encoding Initiative Consortium, C.M. Sperberg-McQueen and L. Burnard (red.) 2002: *TEI P4: Guidelines for Electronic Text Encoding and Interchange. Print Dictionaries*, Chapter 12, XML Version. Oxford. <http://www.tei-c.org/P4X/>

Web-sider (alle sider sjekka 22. september 2005):

- ALLEX, <http://www.dokpro.uio.no/allex/allex.html>
- CIDOC, <http://www.willpowerinfo.myby.co.uk/cidoc/>
- Corpus Workbench, <http://www.ims.uni-stuttgart.de/projekte/CorpusWorkbench/>
- Dokumentasjonsprosjektet, <http://www.dokpro.uio.no/>
- Eining for digital dokumentasjon, <http://www.edd.uio.no/>
- Museumsprosjektet, <http://www.muspro.uio.no/>
- Norsk Ordbok, <http://no2014.uio.no/>
- Tekstlaboratoriet, UiO, <http://www.hf.uio.no/tekstlab/>

Ålov Runde, Terje Svardal, Oddmund Vestenfor

Metaordboka som reiskap for lemmaseleksjon og ordboksproporsjonering

Metaordboka (The Metadictionary) is a search index to the databases (digitalized archives) of Norsk Ordbok. In addition to the digitalized original paper slip archive a digital corpus of written Norwegian is being developed. This is the new digital platform for the further edition of Norsk Ordbok.

In this paper two components of dictionary making, i.e. lemma selection and dictionary volume planning, are focused upon. The Metadictionary gives an exact survey of the number of lemmas in the database. This, combined with our knowledge about different types and number of sources connected to each lemma gives us information that enables us to set up a plan for the contents of every volume of the dictionary.

Innleiing

I dette innlegget vil vi komme inn på fire hovudpunkt:

- 1) Kva er Metaordboka (MO)?
- 2) Korleis kan Metaordboka vere til hjelp i arbeidet med lemmaseleksjon?
- 3) Korleis kan vi nytte Metaordboka i arbeidet med proporsjoneringa av Norsk Ordbok?
- 4) Korleis kan bruken av korpus supplere (det allereie eksisterande) arkivmaterialet til Norsk Ordbok?

1 Metaordboka

Metaordboka er, enkelt sagt, ein systematisert søkjeinngang til dei ulike elektroniske samlingsane til Norsk Ordbok (Ore 1999, Grønvik 1997). Ho er normert på grunnlag av Skard (1939) som er ei ordliste over rettskrivinga for nynorsk, vedteken i 1938. Vi valde denne norma for Metaordboka, fordi ho er nytta som redaksjonsmål for Norsk Ordbok. I Metaordboka er variantar av lemma førde under eitt og same oppslag.

Metaordboka har no følgjande tilfang: Der finn ein for det første den digitale versjonen av setelarkivet til Norsk Ordbok. Det er ca. 3 millionar setlar som ligg føre som faksimile, og dessutan er opplysningar på setelen om oppslagsord (lemma), ordklasse, kjelde, heimfesting og innsamlar skrivne inn i eit eige skjema. Grunnlaget for søkjeindeksen i MO er oppslagsorda frå setelarkivet.

På same måte er arkivet til Trønderordboka behandla, og desse setlane er også kopla til MO. Dessutan har MO elektroniske utgåver av NOB, Skard 1939 og det såkalla Grunnmanuskriptet. Grunnmanuskriptet er det første utkastet til Norsk Ordbok, som blei ferdigstilt i 1940, men aldri utgjeve.

Ein typisk MO-artikkel kan illustrerast med t.d. artikkelen 'idrott', som vi ser nedanfor

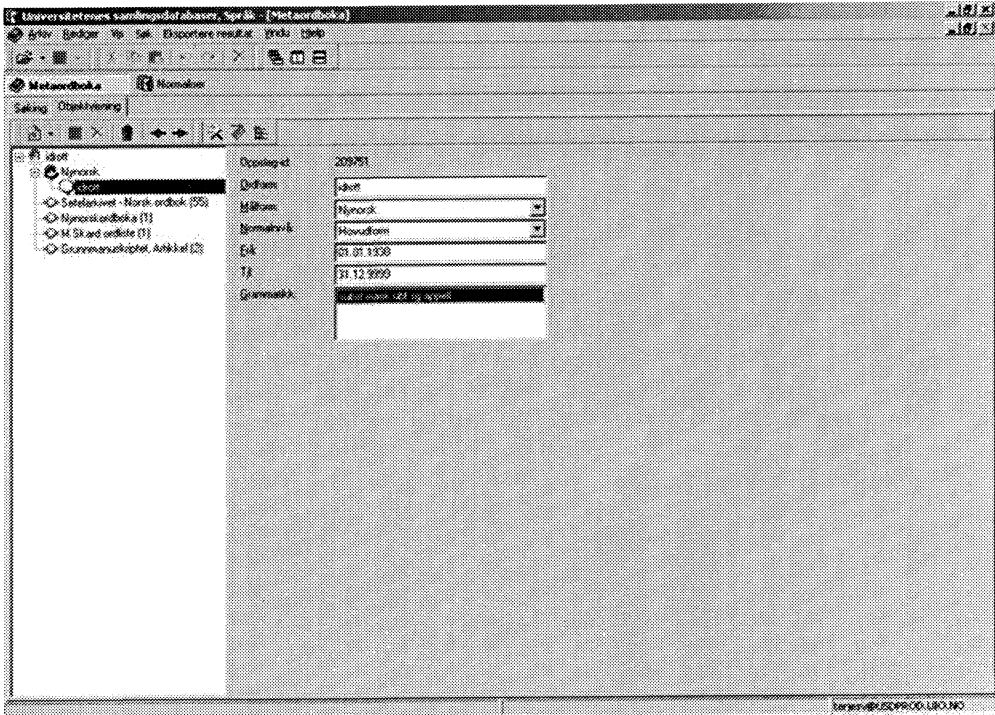


Fig. 1

(Fig. 1). Den er normert til maskulinum på grunn av normeringa i Skard. Til denne MO-artikkelen er det knytt setlar frå arkivet til Norsk Ordbok (55 i alt), og artiklar frå NOB, Skard 1939 og Grunnmanuskriptet.

I samarbeid med Eining for digital dokumentasjon ved HF-fakultetet (EDD, tidl. DOK) arbeider vi kontinuerleg med å byggje ut MO. Mellom anna skal Alf Tørp *Nynorsk etymologisk ordbok* (1919) koplast til MO, og på planen for neste år står også å leggje inn elektroniske utgåver av Ivar Aasens ordbøker frå 1850 og 1873.

2 Lemmaseleksjon

Metaordboka er ein nyttig reiskap i arbeidet med Norsk Ordbok, mellom anna når det gjeld spørsmålet om kva for oppslagsord som skal vere med i ordboka. Her skal det kort nemnast nokre sentrale punkt knytte til dette spørsmålet. Kriteria som ligg til grunn for lemmautvalget, er både positivt og negativt definerte.

Metaordboka skal vere grunnlaget for oppretting av artiklar i Norsk Ordbok. Desse skal i første omgang genererast utifrå dei kriteria som blir nemnde nedanfor.

Metaordboka gir oss for det første eit fullstendig oversyn over talet på oppslagsord i databasen. Det ligg no på 554 275 ord. I Norsk Ordbok, som er planlagt til 12 band, skal om lag 300 000 ord beskrivast. Vi har med andre ord eit overskot på om lag 250 000 ord.

Av oppslaga i Metaordboka er ca. 50 prosent hapax legomenon, altså ord som berre har éi kjelde knytt til seg. Her er hovudregelen at ord med berre ei kjelde bak seg ikkje skal beskrivast i Norsk Ordbok, men der finst nokre viktige unntak frå regelen. Målførebelegg skal som ein hovudregel beskrivast i Norsk Ordbok, sjølv om dei er hapax. Artiklar frå Grunnmanuskriptet blir også tekne med i Norsk Ordbok, sjølv om dette er einaste belegget vi har på ordet. Grunnen er at Grunnmanuskriptet byggjer på Aasen og Ross.

Eit anna kriterium for om eit lemma frå Metaordboka skal få vere med i Norsk Ordbok, er kva litterær kjelde det er henta frå. I samlingane våre finst det ekserpt frå ordbøker laga relativt tidleg i nynorskperioden, der eitt av måla tydeleg har vore å komme med framlegg til (ny)norske avløyсарord. Ikkje alle desse orda har slått igjennom og gått inn i skriftspråket. Dersom ei av desse ordbøkene er einaste belegg på ordet, vil vi velje det bort, for ikkje å beskrive reine «skrivebordsord». Eit døme på dette er ordet 'innkrevjande', som er henta frå Ola Raknes: *Fransk-norsk ordbok* (1939-1942).

Ein annan type lemma vi vel å sjå bort frå på grunnlag av Metaordboka, er samansetningar frå omsett litteratur som berre har éi kjelde. Faksimilen nedanfor, eit ekserpt frå Henrik Rytter: *Shakespeare skodespel I-X* (1922-1933), har eit døme på denne typen samansetning, adjektivet 'isterdiger':

Ein tredje type oppslagsord vi kan sjå bort frå på bakgrunn av kunnskap om kjeldene til Norsk Ordbok, er oppslagsord innleidde med bindestrek. Suffiks er beskrivne i andre ordbøker og blir tekne med i ordboka, medan resten av «bindestreksorda» er oppslagsord som er komne til på grunn av såkalla etterleddssetlar frå det tradisjonelle arkivet. Vi vil berre generere oppslagsord av «bindestreksord» (artiklar innleidde med bindestrek) som byggjer på ordboksartikkel. Dvs. at dei er beskrivne anten i Skard 1939, Grunnmanuskriptet eller NOB. Til dømes vil det bli skrive ein artikkel om suffikset *-sam*.

Til saman gir kombinasjonen Metaordboka og kjennskap til kjelder eit godt grunnlag for lemmaseleksjon. Samstundes er det viktig å understreke at all informasjonen i databasen er intakt og tilgjengeleg for redaktørane i arbeidet med Norsk Ordbok. Det er ikkje slik at vi kvittar oss med informasjon, men vi gjer eit utval av det vi vil vise i ordboka.

3 Ordboksproporsjonering

Av dei 12 banda som NO skal omfatte, er 4 allereie utgjevne¹ og åtte står att å skrive fram til 2014. Kwart band i Norsk Ordbok omfattar om lag 25 000 ordartiklar.

Som nemnt tidlegare i artikkelen, inneheld arkivet i alt om lag 555 000 MO-artiklar. Bolken h-å, som står att å redigere, omfattar 367 000 MO-artiklar. Deler vi dette talet på to, utfrå kriteriet om at hapax ikkje skal redigerast, er vi nede i 180.000 ord, som gir eit snitt på 23 000 artiklar per attståande band.

For planlegginga sin del kan vi altså bestemme nokså nøyaktig innhaldet i og omfanget av kvart band. Kvar bokstav er kartlagd; bokstaven *s* har til dømes ca. 85 000 ord. Av desse skal ca. 50 prosent beskrivast (dvs. om lag 42.000 ord), som det går fram av utvalskriteria. Dette utgjer ca. 2 heile band. På den andre sida er det på bokstaven *i* knapt 8 000 MO-artiklar og den vil omfatte knapt 4 000 artiklar i Norsk Ordbok. Dette utgjer grovt rekna 1/6 ordboksband.

4 Korpus som supplement

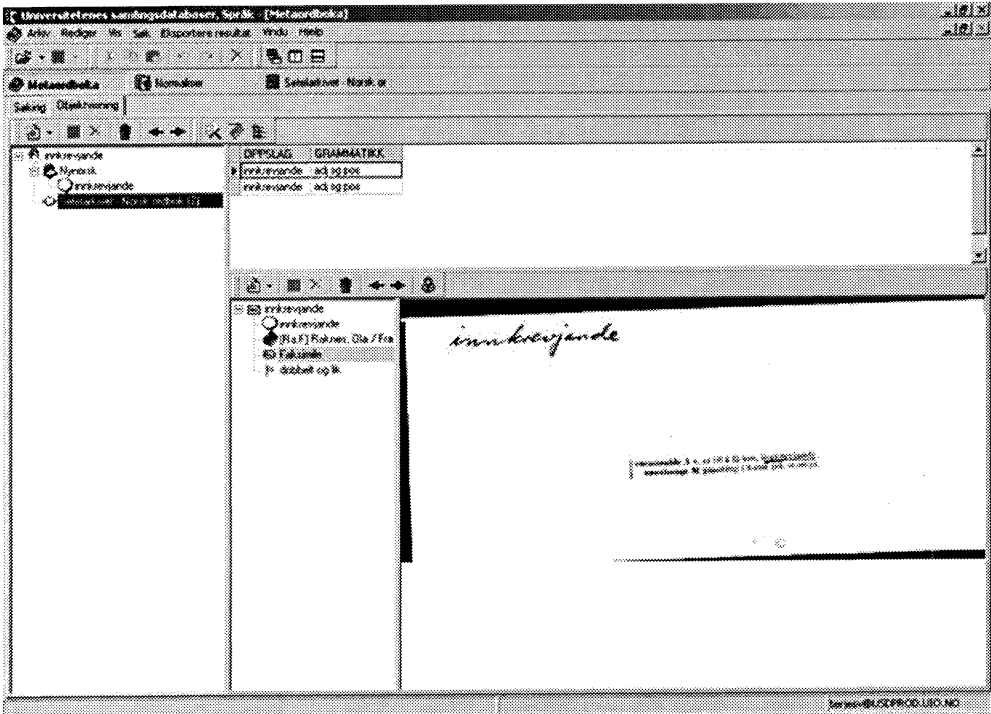


Fig. 2

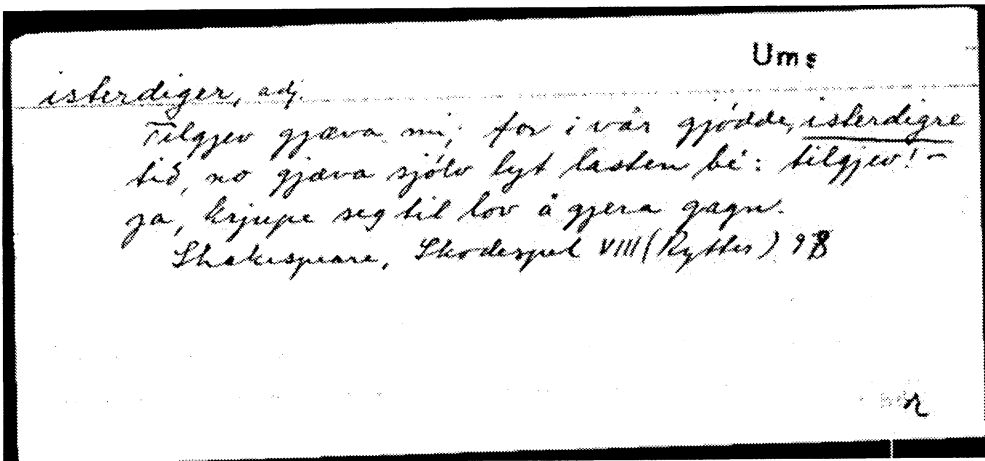


Fig. 3

Nynorskorpuset, som blir bygd opp ved Seksjon for leksikografi og målføregransking, Universitetet i Oslo, skal knytast til Metaordboka. Vi vil knyte nokre kommentarar til korleis det kan hjelpe oss i ordboksarbeidet. Arkivet til Norsk Ordbok har ein viss skeivleik. Det spesielle er til dels betre dekkja enn det generelle og eldre språk er betre dekkja enn yngre.

Korpuset vil kanskje særleg vere til hjelp i arbeidet med å finne kollokasjonar og gi betre grunnlag for semantisk inndeling av tilfanget. Dessutan kan ein bruke korpuset til å generere nye oppslag i Metaordboka, der det er manglar i arkivet til Norsk Ordbok.

5 Oppsummering

Metaordboka er ein svært nyttig reiskap i arbeidet med Norsk Ordbok på fleire måtar; her er det nemnt to av dei. Vi har sett korleis ein ved hjelp av Metaordboka, kombinert med kjennskap til kjeldene til Norsk Ordbok, har eit godt grunnlag i arbeidet med lemmaseleksjon. Dessutan kan vi ved hjelp av Metaordboka bestemme nokså nøyaktig innhaldet i og omfanget av kvart band som står att å redigere. Og sist, men ikkje minst: Metaordboka supplert med korpusbruk gir eit rettare og meir fullstendig bilete av det moderne nynorske skriftspråket.

Litteratur

Grønvik, Oddrun 1996: *Om kjeldegrunnlaget for Norsk Ordbok* i Norsk Ordbok – nynorskens leksikografiske kanon?, Oslo, 23-37

Ore, Christian-Emil 1999: *Metaordboken – et rammeverk for Norsk Ordbok*, Gøteborg 1966-: Norsk Ordbok – ordbok over det norske folkemålet og det nynorske skriftmålet, Oslo

Raknes, Ola 1939-1942: *Fransk-norsk ordbok*, Oslo

Rytter, Henrik 1932-1933: *Skodespel / William Shakespeare; i norsk umskrift ved Henrik Rytter*, Oslo

Skard, Matias 1939: *Nynorsk ordbok for rettskriving og litteraturlesnad*, Oslo

Torp, Alf 1915-1919: *Nynorsk etymologisk ordbok*, Kristiania

Aasen, Ivar 1850: *Ordbog over det norske Folkesprog*, Christiania

Aasen, Ivar 1873: *Norsk Ordbog : med dansk Forklaring*, Christiania

Notar

1. Band 5 blir utgjeve i 2005.

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 dokumentasjonsleksikografisk 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 bygge opp elektroniske ordbasar. Språkleg dokumentasjonsverksemd og dokumentasjonsleksikografiverksemd 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 forskningen 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 tyde-legare 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 globaliseringspråket"

(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. Ei utfordring er at det etter Dokumentasjonsprosjektet ikkje har skjedd tilsvarende systematisk digitalisering av nyare materiale. Ei anna utfordring er at korkje bokmålstilfanget eller ordtilfang frå gammal- og mellomnorsk tid har vorte digitalisert tilsvarende dialekt- og nynorsktilfanget. Dessutan er det uklårt kva som vil skje med dialekt- og nynorsktifanget 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 samkøyring 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 fyldigare 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 bygge 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 mellomaldertekster og -handskrifter. Det er først og fremst eit digi-taliseringssprosjekt.

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 sjølvsagt 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 sjølvsgt 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 tvinga 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 prosjektsøklarar òg kan trekkje lærdom frå og bruke i sitt arbeid. Men ein viktig føresetnad her er sjølvsgt at dei universitetsbaserte leksikografamiljøa 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* resultatata.

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 bygge 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>

LARS S. VIKØR

Norsk Ordbok ved ein krossveg. Rapport frå eit omstillingsprosjekt

On the fourth conference on Nordic lexicography (1997, see Vikør 1999), *Norsk Ordbok* ('Norwegian Dictionary') was presented as an example of a dictionary project spanning several generations. In 2000, the Norwegian Ministry of Culture declared itself willing to strengthen the dictionary project in order to make completion possible already in 2014. This will necessitate a thorough reform process, which is now in the preparation phase. The paper gives a brief description of the latest developments and the current situation at the end of 2001, and discusses some problems and dilemmas which have to be solved in the process.

På den fjerde nordiske leksikografikonferansen, på Hanaholmen i 1997, vart Norsk Ordbok presentert som eit eksempel på ei "fleirgenerasjonsordbok", og dei særlege problem og dilemma som knyter seg til eit slikt verk, vart drøfta (Vikør 1999). I 2000 sette det norske Kulturdepartementet seg som mål å styrkje Norsk Ordbok så radikalt at ho skulle kunne fullførast alt i 2014. Dette har ført til at eit stort omstillingsprosjekt er blitt sett i gang. Det er enno ikkje avklart kva det vil føre til, men eg skal her først skissere det som er hendt dei siste par åra og kva situasjon vi står i no. Deretter skal eg ta utgangspunkt i dei dilemmaene eg skisserte i 1997 og gi ei vurdering av korleis dei fortonar seg i dag.

På Hanaholmen nemnde eg *framdrifta* som eitt av dei store problema for Norsk Ordbok. Både forlaget vårt, Det Norske Samlaget, og Kulturdepartementet, som støttar ordboka med ei årleg bevilgning, ønskte ei monaleg forsering av utgjevinga. Historia til Norsk Ordbok er framstilt andre stader (t.d. Vikør 1996b: 706-707, Vikør 1997: 6-7 og 13), så eg nemner her berre at ordboka i utgangspunktet var eit sjølvstendig prosjekt med statsstøtte, som starta opp i 1930. I 1972 vart ho lagd inn under Universitetet i Oslo, og alle redaktørar og andre medarbeidarar, som det etter kvart vart åtte av, fekk faste stillingar der. I 1987 kom Kulturdepartementet inn med si årlege bevilgning, som dekte tre stillingar, med tanke på ei framskunding. Sidan da har altså ordboka for størstedelen vore finansiert over universitetets budsjett, mens Kulturdepartementet bidrog med eit mindre tilskot. Tanken på å ha eit slikt tilskot, om enn ikkje så stort, på budsjettet i fleire generasjonar, har nok bekymra byråkratane i departementet. Eg fekk forklart frå ein av dei at det dreiar seg om elementær politikarpsykologi: Små bevilgningar i det uendelege er meir uakseptabelt for ein politikar enn store summar til formål med ein tydeleg definert og tidfesta slutt. Derfor fekk vi på slutten av nittitalet følarar frå departementet med spørsmål om kor mykje ekstra midlar vi trong for å fullføre heile verket i 2014. Årstalet 2014 er symbolsk: Da er det 200 år sidan unionen med Danmark vart oppløyst og nordmennene vedtok ei grunnlov som var utgangs-

punktet for den grunnlova vi enno har - fødselsåret for den nye nasjonalstaten Noreg, kan vi seie.

Målsetjinga 2014 verka ganske utopisk på oss. Vi meinte at eit realistisk overslag over kor mykje som skulle til, ville vere nok til å setje punktum for heile ideen. Ein medie-kampanje for friske millionar til Norsk Ordbok såg vi ikkje for oss. Utan særleg håp sette vi likevel opp eit reknestykke og sende over til departementet.

Denne sida av saka skal eg ikkje gå altfor mykje inn på her, men tjue nye stillingar og like mange millionar pr. år fram til 2014 var det vi kom til og la fram. Da fekk vi også pressedebatten, i og med at avisa Dagens Næringsliv slo stort opp at staten hadde tenkt å bruke 260 millionar på - av alle ting - ei ordbok. Det skjedde i april 2000, men da hadde det skjedd viktige ting i mellomtida. Departementet hadde lagt fram ei stortingsmelding om "Kjelder til kunnskap og oppleving. Om arkiv, bibliotek og museum i ei IKT-tid". Der vart Norsk Ordbok nemnd som eitt av fire nasjonale oppslagsverk som departementet ønskte å sjå ferdig i 2014. Dei tre andre var alle langt mindre av omfang enn Norsk Ordbok. Departementet - dvs. regjeringa, sjølvsagt - skreiv klart i meldinga at ho var innstilt på å bidra med det som trongst av bevilgningar for å få realisert dette, utan å talfeste nokon ting. Denne stortingsmeldinga kom ved årskiftet 1999-2000, og var den direkte årsaka til at somme aviser begynte å interessere seg for saka.

No vart det ingen pressestorm mot Norsk Ordbok. Tvert imot, vi fekk mykje positiv interesse frå folk som aldri hadde hørt om oss før. Og etter eit år, ved starten av 2001, gjekk også Familie-, Kultur- og Administrasjonskomiteen i Stortinget inn for målsetjinga om fullføring i 2014. Da hadde vi hatt regjeringsskifte i mellomtida, frå sentrumsregjeringa Bondevik I til Arbeidarparti-regjeringa Stoltenberg. Det førte ikkje til noka endring for Norsk Ordbok. I stortingskomiteen var alle parti med meir enn eitt mandat i Stortinget representerte, og det var ingen dissens.

Vi arbeidde vidare i forhold til styresmaktene, og fekk eit gjennombrøt under eit møte med departementet den 13. juni 2001, der både statsråden, kulturminister Ellen Horn, og Universitetets rektor, Kaare Norum, var med. Møtet var kort, men godt forberedt. Statsråden lova der konkret å finansiere dei nye stillingane som trongst etter det behovsoverslaget vi hadde gitt, og rektor Norum lova å bidra med det som trongst når det gjaldt infrastruktur m.m. Finansieringa skal skje etter ein opptrappingsplan, og da Arbeidarpartiregjeringa la fram statsbudsjettet sitt i oktober 2001, hadde dei ført opp den summen som denne planen føresette. Det hadde da vore stortingsval, som vart eit uvanleg kraftig nederlag for Arbeidarpartiet, og kort tid etter overtok ei borgarleg regjering, Bondevik II, med Høgre som det dominerande partiet. Men summen til Norsk Ordbok vart ståande også etter at den nye regjeringa hadde gjennomgått og revidert budsjettet. I skrivande stund (30.11.01) er avklaringa enno ikkje hundre prosent, men det er likevel usannsynleg at det blir noka nedskjering for Norsk Ordbok.

Dette var den ytre hendingsgangen. Parallelt med dette har det vore arbeid intens internt, i ordboksredaksjonen og på institutt- og fakultetsnivå for å førebu denne omlegginga. Blant anna vart det leigd inn ein ekstern utgreiar for å lage ei tilråding, eit såkalla forprosjekt (Tungesvik 2000), som eit grunnlag å arbeide vidare på. Denne tilrådinga har ulike sider. Delvis har det handla om dei praktiske sidene: å presisere stillingsbehovet og drøfte korleis så mange nye folk skal kunne rekrutterast, drøfte problema med sjølve organiseringa av prosjektet, stillingsstruktur osv. Denne prosessen er i gang, og skal halde fram no i haust og vidare.

Men det viktigaste er sjølvsagt korleis ein skal makte dette reint fagleg, og kva slags konsekvensar denne omlegginga får for ordboka som ordbok. I resten av denne artikkelen vil eg konsentrere meg om dette. Like lite som på Hanaholmen i 1997 kan det leggjast fram klare konklusjonar no, men diskusjonane og tenkinga er naturleg nok komme eit stykke lenger sidan da.

Eg tek utgangspunkt i dei seks dilemmaene eg skisserte i Vikør 1999: 404-405, nemleg *effektivitet vs. kvalitet, representativitet vs. uttømmelegheit, allsidigheit i informasjonstypar vs. konsentrasjon om det vesentlege, einskaplegheit over tid vs. tilpassing til endra krav, forskning vs. formidling, papir vs. elektronikk*. Eg skal seie litt om korleis vi no vurderer desse punkta jamført med situasjonen for fire år sidan.

1) *Effektivitet vs. kvalitet*. Dette er eit brennande spørsmål, som faktisk er like opent no som i 1997. Rett nok kan ein seie at det ikkje burde vere ei motsetning her, men at det er ei *potensiell* motsetning, må vi i det minste ha opne auge for. Ei så kraftig forsering av ordboka føreset ei langt strammare redigeringsstakt, der tida til å fordjupe seg i vanskelege problem nødvendigvis må vere avgrensa. Dette nødvendiggjær igjen eit langt meir detaljert og eksplisitt sett redigeringsreglar, som frittek redaktørane for val så langt det er mogleg både i framstillingsforma - sjølv redigeringskonvensjonane - og i utvalet av informasjon som blir teke opp. Avgrensinga til tolv band (Vikør 1999: 397) nødvendiggjær i alle fall eit langt strammare utval av eksempel osv. Problemet med motsetninga effektivitet vs. kvalitet er at effektivitet er enkelt å måle - det er berre å rekne ut produsert tekstmengd pr. tidseining, mens kvaliteten derimot eigentleg ikkje kan målast. Det blir ei svært viktig oppgåve å setje opp tydelege og handterbare kvalitetsnormer. Der må redaksjonen og ikkje minst redaksjonsleiinga ha hovudansvaret, men forskarsamfunnet og brukarsamfunnet rundt oss bør også ha eit medansvar; vi vil altså gjere dette i dialog med dei ordboka er til for. Effektiviteten må deretter målast på ein slik måte at desse kvalitetsnormene blir ein del av kriteria å måle etter.

2) *Representativitet vs. uttømmelegheit*. Norsk Ordbok har tradisjonelt strevd etter å nærme seg mest mogleg det prinsipielt uopnåelege målet om uttømmelegheit, og både plass- og tidsbruken kan langt på veg forklarast ut frå dette idealet. Ei prioritering av det representative på bekostning av det heilt uttømmmande vil måtte liggje til grunn for omlegginga vår - det følger allereide av avgrensinga til tolv band. Det vil innebere både eit strengare ordutval, basert på klare og gjennomsiktlige kriterium (Vikør 1996a), og ei deprisering, t.d. når det gjeld framlegginga av dialekttilfanget for dei einskilde orda. Her ligg det ei fallgruve av same slag som vi hadde for "kvalitet": Korleis skal vi definere "representativitet"? Det må - i sjølv omleggings- og planleggingsfasen - setjast opp klare kriterium for dette òg, der ein tek høgde for at representativitet ikkje nødvendigvis alltid må innebere høg frekvens, til dømes. Original språkbruk hos sentrale forfattarar må såleis bedømmast annleis enn hos meir perifere. Men uansett er det klart at Norsk Ordbok i framtida sterkare må framstå som ei selektiv ordbok, og vise eintydig også for brukarane kva kriterium det er selektert etter.

3) *Allsidigheit i informasjonstypar vs. konsentrasjon om det vesentlege*. Dette er eit meir problematisk punkt enn det førre. Ein ting er å stramme inn på den informasjonsmengda ein tek inn, ein annan ting er å endre på sjølv omlegget og malen for artiklane. Eg sa i Vikør 1999: 404 at vi kanskje måtte prioritere definisjonar og bruksopplysningar, syntaktisk og stilistisk informasjon over opplysningar om ulike former, da særleg målføreformer, belegg-opplysningar osv. Det kan hende, i den forstand at vi depreisierer dei sistnemnde typane av

plass- og kapasitetsomsyn og er mindre utførlege i å gjere greie for dei. Dette kan faktisk vere ei fagleg vinning, fordi særleg dialektopplysningane med dei nøyaktige heimfestingane vi har i dag, lovar meir enn dei held. Dei gir inntrykk av å vere eit fullstendig og detaljert oversyn som baserer seg på presise grenser mellom dialektane frå bygd til bygd, mens det i verkelegheita er glidande - og over tid skiftande - overgangar som blir betre markerte med færre "typeformer" og ein meir romsleg heimfestingspraksis, altså regionar og landskap heller enn enkeltbygder. Men så langt eg kan sjå, bør det framleis vere ei av normene for ordboka at vi skal ha det same spekteret av opplysningstypar som vi har i dag. Det bør, etter mi meining, vere den ramma vi gjer dei nødvendige endringane innanfor. Det fører rett over til det neste viktige dilemmaet:

4) *Einskapleghet over tid vs. tilpassing til endra krav.* Dette omfattar mange andre problem enn det eg nettopp har nemnt, for det er uunngåeleg at ein alltid vurderer kritisk både det ein har gjort og det ein arbeider med, og at ein kan komme til at både dette og hitt kunne ha vore gjort på ein annan måte. Det gjeld typografi og andre trekk som gjeld ytre form, som rettskrivinga. Det gjeld altså sjølve malen for artiklane, og det gjeld sjølvstilt ikkje minst endringar innanfor sjølve faget, som i grammatisk terminologi og lingvistisk beskrivingsmodell, synet på kva ein dialekt er og på kva litterære kjelder som er meir eller mindre vesentlege osv. Ikkje minst gjeld det slike ting som definisjonsutforming og prinsipp for tydingsinndeling. På slike punkt har det vore ein tommelfingerregel inntil no at vi skal halde oss til det som har vore praktisert før, f.eks. skal vi ved skrivinga av artiklar om bestemte typar ord, t.d. frå eit visst semantisk felt, jamføre med tidlegare redigerte ord frå det same feltet og følgje malen der. Eg må seie at slike prinsipp berre delvis blir følgde; det som er komme ut, er produkt av tre ulike redaktørgenerasjonar, og det har vore uunngåeleg at yngre redaktørar i ein del tilfelle har gjort ting annleis enn dei eldre. Total einskap over tid er truleg uoppnåeleg for eit fleirgenerasjonsverk.

Andre har hatt same problemet, og eitt kollegaprojekt vi har hatt ein del kontakt med er WAT, *Woordeboek van die Afrikaanse Taal* i Sør-Afrika. Dei har gjennomgått ein tilsvarende overgang som vi må ta mellom to heilt ulike måtar å redigere og publisere ordboka på, i retning kraftig effektivisering. Sjefredaktøren der, Dirk van Schalkwijk, sa tydeleg da han besøkte oss i fjor og drøfta desse tinga at hos dei hadde dei skore igjennom og gjennomført ei rekkje fundamentale endringar mellom to band, og han meinte det ville bli akseptert når det berre vart gitt gode nok grunnar for det, også overfor brukarane. For vår del står den situasjonen for døra, men mi personlege haldning er nok meir konservativ enn Van Schalkwijks. I den trykte versjonen ligg det mest i retning av at vi vil prøve ein mellomveg. Den ytre forma meiner vi bør vere einskapleg gjennom heile verket, det gjeld bl.a. typografi og rettskriving. Norsk Ordbok bruker ei rettskriving som var offisiell mellom 1938 og 1959, men den ligg ikkje altfor langt frå den noverande rettskrivinga, så for dei fleste fortonar nok redaksjonsspråket i ordboka seg som ein litt konservativ variant av den nynorsken dei er vane med, heller enn som ei historisk utdatert rettskriving. Vi er heldigare stille slik enn kollegaene våre i WNT, *Woordenboek der Nederlandsche Taal*, som var fullført i 1998. Dei valde også å halde på ei gammal rettskriving, avskaffa i 1947, gjennom heile verket, bl.a. med ei kasusbøying av substantiv som i moderne språk har falle bort. Den verkar nok meir "historisk" enn Norsk Ordbok. Både vi og WNT har altså gjort eit anna val enn SAOB, som gjekk over til den Fritjuv Bergska rettskrivinga ikkje altfor mange år etter 1906.

Dette er ikkje gjennomdrøfta, men mi innstilling er at vi i NO bør satse på endringar som ikkje "synest" i det ytre eller i oppbygginga av artiklane. Reint innholdsmessig og fagleg vil vi følgje med tida, slik at vi i framstillinga av grammatiske forhold, etymologiar osv. vil dra inn nye kjelder og byggje på ny kunnskap. For etymologien inneber det at vi tek omsyn til ei heilt ny etymologisk ordbok som kom ut i 2000, Harald Bjorvands og Fredrik Otto Lindemans *Våre arveord*, mens hovudkjelda vår inntil no har vore Alf Torps *Nynorsk etymologisk ordbok* frå 1919. Einskaplegheit i form, men tilpassing til tida i innhald er målet vi etter mitt syn bør streve mot.

5) *Forsking vs. formidling*. Dette – konflikten mellom dei akademiske og dei samfunns-messige interessene ordboka skal tene – er inngåande drøfta i Vikør 1999: 401-403, for det var eit viktig diskusjonsemne da. Eg trur ikkje dette dilemmaet er så stort lenger, for det kan ikkje bli nokon annan konklusjon enn at ordboksprosjektet er begge delar, og at begge komponentane er gjensidig avhengige av kvarandre. Om det ikkje ligg solid og seriøs forskning i botnen, vil prosjektet bli mislykka, og om denne forskinga ikkje blir formidla slik at kunnskapen når ut i samfunnet og styrkjer bevissteita om norsk språk og språkdyrking, er det også mislykka. Dette siste kan oppnåast ved å la sjølve den papirbaserte ordboka bli ein kjerne i eit større kompleks av ulike formidlingsformer, og her kjem det siste dilemmaet inn, som eg berre var innom med ein harelabb på Hanaholmen, men som står uhyre sentralt i dag og i framtida, nemleg:

6) *Papiret vs. elektronikken*. Kjernen i Norsk Ordbok-prosjektet er framleis ei papirutgåve i tolv band; det er alle aktørane einige om, frå Stortinget til sjølve ordboksredaksjonen. Men her må det blir eit både – og, ikkje eit anten – eller. Eigentleg er vi komne forbi den fasen da vi såg på dette som eit dilemma. Det skal satsast på moderne teknologi for alt han er verdt. Vi er i gang med å utvikle elektroniske reiskapar som langt på veg vil kunne kompensere for dei avgrensingane ei papirutgåve set. Éin ting er at store delar av det kjeldearkivet vi bygger redigeringa på, no er elektronisk tilgjengelege og søkbare, både for oss i redigeringa, og for andre interesserte som ønskjer å finne informasjon om ordbruk i nynorsk og dialektane. Det vil gjere det mogleg for redaktørar å sitje andre stader i landet og arbeide med det same kjeldematerialet som vi, ein klar fordel når vi snart, om det går bra, må etterspørje dobbelt så mange nye medarbeidarar som vi har i dag, ja, enda fleire, fordi dei noverande redaktørane nærmar seg pensjonsalderen.

Sjølve Norsk Ordbok vil også finnast elektronisk saman med ei rekkje andre nynorske ordbøker, og det gjer det mogleg for brukarar i forskarsamfunnet og elles i samfunnet å kople informasjonen i desse kjeldene og søkje i dei på kryss og tvers. Det blir ei samankøpling av ulike kjeldetypar med ein felles inngang som består av ei normert skriftform med grammatisk markering for kvart einaste oppslagsord eller søkjeord. Det er det vi kallar metaordboka, sjå artiklane av Ore o.a. og Svardal i denne rapporten.

Dette vil også gjere det mogleg å oppdatere ordboka i framtida, noko som kan trengast når dei første hefta vart produserte i 1950-åra. Den avgrensinga som ei fleirgenerasjonsordbok alltid har vore underlagd, nemleg at første og siste band avspeglar ulike språksteg i ulike samfunn, kan såleis kompensereast. Ei sak for seg er det at slik oppdatering av dei første banda må vente til vi har ferdig resten, fordi all tilgjengeleg kompetent arbeidskraft må setjast inn på dei kommande delane. Slik sett er det arbeid nok med Norsk Ordbok etter 2014 òg, sjølv om det neppe blir laga supplementsband på papir. Det er berre rett og rimeleg, for språket skal halde fram med å endre seg og utvikle seg vidare etter at NO er ferdig,

og eit vitskapleg leksikografisk institutt må kontinuerleg følgje med i og registrere denne utviklinga.

Teknologien vil òg hjelpe oss på andre måtar, bl.a. ved at vi får utvikla eit redigeringsprogram som vil gjere ei konsekvent og systematisk redigering meir mogleg enn den gamle manuelle måten kunne. Dette programmet er berre i startgropa i dag, men arbeidet med det blir noko av det viktigaste i det kommande året - 2002.

Eg vil konkludere så langt med at elektronikken vil kunne tilføre Norsk Ordbok nye kvalitetar som var utenkjelege for dei som sette i gang prosjektet - så sant prosjektet får dei ressursane som er nødvendige, og så sant vi som skal arbeide der, gjer ein tilstrekkeleg god jobb og i tillegg til alle moderne hjelpemiddel mobiliserer like mykje god gammaldags hjernekraft og flid som forgjengarane våre gjorde.

Litteratur

- Tungesvik, Steinulf 2000: *Norsk Ordbok 2014*. [Upublisert forprosjektrapport, Institutt for nordistikk og litteraturvitskap, UiO, Oslo].
- Vikør, Lars S. 1996a: Problem kring ordutvalet i Norsk Ordbok. I Ásta Svavarsdóttir o.a. (red.): *Nordiske studier i leksikografi* 3, s. 403-412.
- Vikør, Lars S. 1996b: Norsk Ordbok – the Crown of Nynorsk Lexicography? I M. Gellerstam o.a. (red.): *Euralex '96. Proceedings I-II*, s. 705-713.
- Vikør, Lars S. (red.) 1997: *Norsk Ordbok – nynorskens leksikografiske kanon?* Rapport frå eit seminar på Blindern 31. mai 1996. Institutt for nordistikk og litteraturvitskap, Universitetet i Oslo.
- Vikør, Lars S. 1999: Fleirgenerasjonsordbøker og tida. I Peter Slotte o.a. (red.): *Nordiska studier i leksikografi* 4, s. 395-403.

Norsk Ordbok og NO 2014-prosjektet

Åse Wetås

1. Innleiing

Denne artikkelen er ein presentasjon av Norsk Ordbok og av materialet og arbeidsmetodane som ligg bak dette store ordboksverket. Artikkelen startar med ein innleiande presentasjon av sjølve dokumentasjonsordboka Norsk Ordbok. Etter kjem eg nærare inn på historikken til ordboka, med særleg fokus på det arbeidet som blei gjort i den tidlege arbeidsperioden, då verket var under plan-legging. Så vil eg presentera prosjektorganisasjonen Norsk Ordbok 2014 og demonstrera utviklinga og endringane som har skjedd i arbeidsmetodane i og med den moderne prosjektorganisasjonen. Eg vil i denne samanhengen også gå nærare inn på grunnlagsmaterialet for ordboksverket og prinsippa bak lemma-utvalet, før eg avrundar med å gje ei skildring av korleis arbeidet og kvalitets-sikringa i den store moderne ordboksredaksjonen skjer frå materialstadiet og fram til dei endelege ordboksartiklane.

2. Norsk Ordbok

Norsk Ordbok – med undertittelen *Ordbok over det norske folkemålet og det nynorske skriftspråket* – er ei dokumentasjonsordbok som når ho blir ferdig, skal bli på i alt 12 band. Som tittelen på verket viser, er dette ei ordbok som doku-menterer det norske folkemålet, altså alle dei norske dialektane. I tillegg dekkjer ordboka det nynorske skriftspråket frå og med den tidlegaste tekstproduksjonen på landsmål og heilt fram til i dag. Verket dekkjer både skrift- og talespråk, og dette er bakgrunnen for at Kultur- og kyrkjedepartementet i 2008 karakteriserte ordboka som “eit sjeldan verk i europeisk samanheng” (St.meld. 35, 2007–08, kap. 8.3.5.3). Dei tilsvarende store dokumentasjonsordboksverka over nabo-språka våre, *Svenska Akademiens Ordbok* (SAOB) og *Ordbog over det Danske Sprog* (ODS), har ei litt anna innretning, og dekkjer primært skriftspråk.

Den prenta versjonen av Norsk Ordbok kjem når verket er ferdig til å innehalda ein stad mellom 300 000 og 350 000 ordartiklar. Kvart av dei att-ståande banda skal i utgangspunktet innehalda om lag 25 000 artiklar, men det faktiske talet på artiklar har lege noko

høgare og vore litt skiftande frå band til band. Band 7 av ordboka (*l-mugetuft*), kom ut i 2008 og inneheld 29 000 artiklar. Band 8 (*mugg-ramnsvalt*) kom i november 2009, og inneheld om lag 30 000 ordartiklar, mens band 9 (*ramost – skodda*), lansert i februar 2011, inneheld ”berre” drygt 27 000 artiklar.

Norsk Ordbok er ei dokumentasjonsordbok, og ho utgjer eit omfattande vitenskapleg arbeid. Det dåverande Kultur- og kyrkje-departementet (KKD) skreiv i stortingsmeldinga om norsk språk frå 2008 (*Mål og meining*, St.meld. 35, 2007–08, kap. 8.3.5.3) at verket i fullført stand vil representera “ein etter måten fullnøyande leksikalsk dokumentasjon av den nynorske målforma og det nynorske skriftspråket som vil stå seg i lang tid framover.” Men departementet var òg nøye med å understreka 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 verdi-full dokumentasjon av ein felles, nasjonal språkarv som direkte vedkjem ein mykje større del av norske språkbrukarar enn dei som nyttar nynorsk som skriftmål (op.cit.).

For den store redaksjonen som til dagleg arbeider med ordboka, er understrekinga av det tosidige mandatet og den vitenskaplege tilnærminga til materialhandsaminga grunnleggjande viktig. At Norsk Ordbok dekkjer både talt og skrive norsk, har ein sjølvsgad samanheng med historia til det nynorske skriftspråket. Prinsippet bak integreringa av dialektord utan landsdekkjande utbreiing i ordboksverket er at alle ord frå det norske talemålet kan normerast til ei nynorsk skriftform. Utgjevingsarbeidet blei starta rett etter 2. verds-krigen, og språket i ordboksverket følgjer 1938-rettskrivinga. Dette inneber at alle opp-slagsord (lemma) blir normerte til ei form som følgjer denne normalen, slik han kjem til uttrykk i Matias Skards *Nynorsk ordbok* frå 1954.

3. Historikk

Arbeidet med Norsk Ordbok blei starta i 1930, etter ei løyving frå staten i 1929 på kr 75.000. Desse midlane skulle fordelast over 7 ½

år. Det var nynorsk-forlaget Det Norske Samlaget som tok initiativet til ordboksarbeidet, og frå starten var det forlaget som styrte arbeidet med ordboka. Det blei sett opp ei eiga tilsyns-nemnd for ordboka, og denne nemnda var det forlaget som peika ut medlemmar til. Nemnda for Norsk Ordbok hadde fram til 1963 litteratur-professoren Olav Midttun som leiar. Etter dette overtok professor Olav T. Beito som nemndleiar. Svært mange sentrale nynorskfolk har sete i denne tilsyns-nemnda for ordboka. Mellom dei kan nemnast Leiv Heggstad, Gustav Indrebø, Sigurd Kolsrud og Knut Liestøl.

Frå 1947 var Alf Hellevik hovudredaktør for Norsk Ordbok, og under han gav ordboksredaksjonen i 1950 ut første heftet av band 1. Arbeidet med å ferdigstillast resten av band 1 tok 16 nye år, og då det kom i 1966, dekte bandet alfabetbolken *a-doktrinær*. Fire redaktørar, Gunnar Pedersen, Reidar Bø og Arnbjørg Hageberg, i tillegg til hovudredaktøren Hellevik, arbeidde då i redaksjonen. Ordboka var opphavleg planlagd til å strekkja seg over 2–3 band, men ved ferdigstillinga av band 1 såg redaksjonen at oppgåva kom til å bli mykje meir omfattande, og dei meinte då at ordboksverket samla kom til å bli på 8–9 band.

Til band 2, som kom ut i 1978, var det åtte redaktørar som hadde levert artiklar. Band 2 strekte seg over alfabetbolken *dokument-flusken*. Seks år tidleg-are hadde redaksjonen og dei store samlingane til Norsk Ordbok gått inn i eit nyoppretta Norsk leksikografisk institutt ved Universitetet i Oslo.

Etter dette skulle det gå 16 nye år før band 3 (*flusker-gigla*) blei gitt ut i 1994. Grunnen til at arbeidet med dette tredje bandet tok så lang tid, var primært at redaksjonen prioriterte å få ferdig den eittbands handordboka *Nynorskord-boka*. Første utgåva av denne ordboka kom i 1986. Band 4 av Norsk Ordbok, som dekte alfabetbolken *gigle-harlemmerolje*, kom i 2002, og blei gitt ut nett ved omorganiseringa av arbeidet til eit tidsavgrensa prosjekt med planlagd avslutning i 2014.

3.1 Utgangspunktet for Norsk Ordbok

I 1932 skreiv Sigmund Skard om Norsk Ordbok at ordboka skulle fylla tre oppgåver. Meir spesifikt skulle ho

- “arbeida saman til eitt alt som er bokført av norske ord; og tyda deim paa norsk”
- “taka upp mest mogeleg av nytt tilfang fraa folkemaalet”

- “fylla det ut med ord fraa den nynorske bokheimen”

I dette låg det både eit program og ein arbeidsmetode for grunnlagsarbeidet. Den første oppgåva munna ut i utarbeidinga av eit grunnmanuskript som skulle ut-gjera fundamentet for det store nye ordboksverket. Dette arbeidet utgjorde òg eit naudsynt grunnlag for å kunna føra i mål dei to andre prioriterte oppgåvene.

Arbeidet med grunnmanuskriptet starta i 1930, og manuskriptet omfatta ord-artiklar frå Ivar Aasens *Norsk Ordbog* frå 1873, Hans Ross' *Norsk Ordbog. Tillæg til “Norsk ordbog” af Ivar Aasen* frå 1895 og Steinar Schjøtt's *Norsk ord-bok med ordtyding paa norsk-dansk* frå 1914. Vidare inneheldt grunnmanu-skriptet Alf Torps *Nynorsk etymologisk ordbok* frå 1919, Chr. Vidsteens *Ordbog over Bygdemaalene i Søndhordland* frå 1900, ved sida av nokre mindre ord-bøker, eldre prenta ordsamlingar og større handskrivne samlingar frå eldre tid (dvs. frå 1600-talet og frametter). I grunnmanuskriptet er alle definisjons-tekstane skrivne på nynorsk.

Den neste oppgåva var å auka tilfanget til ordboka av ord frå dei norske dialektane. Dette måtte gjerast gjennom ei omfattande innsamling av folkemåls-materiale. Innsamlinga var eit stort og krevjande arbeid som måtte gjerast med hjelp av uløna med-arbeidarar ute i distrikta. Første oppropet til ålmenta om å delta i dette arbeidet blei sendt ut i 1931. Frå då av og fram til utgangen av april 1932 hadde det meldt seg så mange som 450 friviljuge med-arbeidarar. Ofte kom kontakten med den sentrale ordboks-redaksjonen i stand gjennom mållag eller gjennom at det blei etablert særlege ordboksnemnder lokalt.

Skard (1932) var tydeleg på at det var to slag utfyllingar innsamlinga av folkemålsmateriale skulle bidra med. Den eine var å få samla inn og systema-tisert ord som aldri før hadde vore ned-skrivne, men som var i levande bruk i folkemålet. Den andre var å få på plass nye heimfestingar av ord, saman med nye tydingar av ord som alt var kjende. Skard understreka særleg at rettleiinga til innsamlarane var streng på at innsamlinga ikkje måtte avgrensast til “rare” og uvanlege ord i dei lokale dialektane. Ønsket var å få inn eit så breitt og represen-tativt materiale som mogleg.

Dette var ei naudsynt og viktig presisering. Innsamling av ord kan fort bli prega av eit fokus på det partikulære, og på ord og former som informant og/eller innsamlar oppfattar som uvanlege og sær-prega for det lokale talemålet. Når målførebelegg blir sette i system, slik dei blir i Norsk Ordboks elektroniske setelarkiv og metaordbok,

viser det seg ofte at ord eller ordtydingar som i ein einskilddialekt framstår som særprega, ofte finst også mange andre stader innan-for eit større eller mindre geografisk område. I tillegg er det slik at eit leksiko-grafisk dokumentasjonsverk ikkje bare treng dokumentasjon på det mest uvanlege og særprega i språket, men like mykje på bruken av meir høgfrekvente ord og ordformer i språket.

Den tredje oppgåva som skulle dekkjast, var innsyn i det nynorske skrift-språket, slik dette kom til uttrykk i litteraturen. Dette måtte skje gjennom inn-samling og ekserpering av skriftspråkleg materiale. Det litterære materialet skulle omfatta vitskapleg litteratur og faglitteratur frå alle typar fagområde, det skulle omfatta skjønnlitteratur av alle sjangrar og mest mogleg frå alle deler av landet, og ikkje minst skulle det omfatta periodisk litteratur (både tidsskrift og aviser). Målet med denne delen av kartlegginga var å få ein så brei dokumenta-sjon som mogleg av det nynorske skriftspråket i bruk. Resultatet av dette arbeidet meinte Skard (1932: 26) ville utgjera det “fyrste samla utsynet over norsk kulturmaal i heilo”.

Arbeidet med desse tre dokumentasjonsoppgåvene skulle gjerast parallelt, og med ei arbeidsdeling mellom fagfolka og dei friviljuge medarbeidarane kringom i landet. Fagmedarbeidarane i redaksjonen utarbeidde grunnmanuskriptet, medan innsamlinga av folkemåls-materiale skulle gjerast av lokale uløna medar-beidarar. Desse siste stod under rettleiing av ordboksnemnda. Arbeidet med utskrivning av tekst frå skriftspråklege kjelder til nynorsk gjekk føre seg gjennom eit nært samarbeid mellom den vitskaplege ordboksredaksjonen og hjelparane ute i bygdelaga. Med dette framstod arbeidet med materialtilfanget for Norsk Ordbok heilt frå starten av som eit omfattande dugnadsarbeid. Utan den uløna innsatsen til svært mange medarbeidarar frå alle deler av landet ville aldri det store kartleggingsarbeidet vore mogleg å få i stand. Også i moderne tid er det slik at brorparten av tilveksten i folkemålsdelen av materialet bak Norsk Ordbok kjem som eit resultat av lokalt organiserte inn-samlingar og utgjevingar.

3.2 Koplinga mellom det norske folkemålet og det nynorske skriftspråket

Då Ivar Aasen planla eit eige norsk skriftspråk, skulle det utgjera eit einenskapleg system, samstundes som språket skulle vera ein felles-nemnar for dei norske dialektane. For å få til dette, gjorde Aasen

nokre prinsipielle og viktige språk-planleggingsval. Aasen laga eit eintydig, koherent system, og for å nå målet om ei felles skriftnorm med basis i det norske folkemåls mangfaldet, valde han ei etymologisk orientering, snarare enn ei ortofon. Likevel nytta han ikkje frå punkt til punkt gammalnorsk som utgangspunkt for det nynorske skriftspråket, men baserte i staden formverket på dei grunndraga han såg var sams for dialektane, og som dessutan også ofte fann støtte i nabospråka (jf Torp & Vikør 1993: 156). Til grunn for landsmålet som Ivar Aasen etablerte, låg det nettopp eit eksplisitt samband mellom skriftnormalen og det talte språket – dei norske dialektane, eller det norske folkemålet.

Norsk Ordbok representerer ei vidareføring av ideane og det omfattande vitenskaplege arbeidet pioneren Ivar Aasen la ned i andre halvdel av 1800-talet. Arbeidet med Norsk Ordbok går ut på å kartleggja folkemålet, og samstundes å dokumentera eit ungt skriftspråk. Gjennom dette framstår ordboka difor òg som ein viktig normeringsinstans for nytt ordtilfang. I Norsk Ordbok får store mengder norsk ordstoff ein normert skrivemåte for første gong. Artikkelmaterialet i Norsk Ordbok tydeleggjer dermed den ubrytelege samanhengen mellom dialektane og det nynorske skriftspråket. Mot denne bakgrunnen er det symboltungt at verket skal fullførast til grunnlovsjubileet i 2014.

3.3 Arbeidet fram til 2002

Utgjevingsarbeidet på 1900-talet var, som vist, omfattande og det tok lang tid. Utgangspunktet for arbeidet med Norsk Ordbok var at det skulle bli eit verk på 2–3 band, men etter kvart som materialinnsamlinga kom i gang og tilfanget auka, blei det heilt klart at ordboka måtte bli mykje større. Særleg andre bandet av ordboka viser den store auken i tilfang, og lengda på den einskilde artikkelen var også aukande (jf Bø 1989). Med store mengder tilfang på arkivkort, i protokollar og i bøker, med ein liten stab og med dei omfattande krava til opp-læring og tileigning av store mengder detaljkunnskap ei fullmannuell drift føreset, gjekk arbeidet med ordboka langsamt. Om dette skriv Tungesvik (2000: 4) at “status for N[orsk] O[rdbok] ved tusenårsskiftet er at ein held ei framdrift som gjer at det trengst 40-60 år til for å få verket ferdig.” Ferdigstilling i 2060 var også det perspektivet St.meld. 22 (1999-2000) la til grunn som eit realistisk slutt punkt ved ei vidareføring av utgjevingsarbeidet i same takt.

4. *Prosjektet Norsk Ordbok 2014*

I Stortingsmelding 22 (1999-2000) sette regjeringa ved Kulturdepartementet opp som målsetjing for Norsk Ordbok at arbeidet med verket skulle vera fullført til grunnlovsjubileet i 2014. I år 2000 blei det laga ein forprosjektrapport som kartla korleis det ville vera mogleg å nå denne målsetjinga med utgangspunkt i dei ressursane som låg føre i organisasjonen. I 2001 vedtok Stortinget å slutta seg til målet om at arbeidet med ordboksverket skulle sluttførast til grunnlovs-jubileet. Dette året laga Universitetet i Oslo (UiO) ved rektor Kaare Norum, og Kulturdepartementet ved statsråd Ellen Horn, ei avtale om fordeling av ansvaret og utgiftene som knytte seg til ferdigstillinga av ordboksverket. Meir spesifikt kom det til semje om at departementet skulle dekkja lønsutgiftene til arbeidet med å fullføra ordboka i 12 band og til fastsett tid, medan UiO skulle dekkja utgiftene til drift, samt syta for lokale.

Norsk Ordbok blei med dette organisert som eit sjølvstendig tidsavgrensa prosjekt med Institutt for nordistikk og litteraturvitskap (seinare Institutt for lingvistiske og nordiske studium) som eigar. Det blei bestemt at prosjektet skulle ha eit eige styre, og at det skulle leiast av ein prosjektdirektør med både fagleg og administrativt hovudansvar. Prosjektorganisasjonen Norsk Ordbok 2014 kom i gang frå 1. juni 2002, med Kristin Bakken som prosjektdirektør.

I perioden frå 2002 og fram til 2008 var prosjektorganisasjonen under utbygging, og frå og med 2008 har prosjektet vore i full drift. I første del av prosjektperioden blei det lagt mykje arbeid ned i å utvikla nye dataverktøy til å bearbeida dei store mengdene materiale som på 1990-talet hadde blitt digitalisert gjennom Dokumentasjonsprosjektet. Overgangen til ei digital plattform gjorde det også mogleg å målstyra produksjonen, og dette gjorde at utgjevingstakten kunne aukast. I 2005, bare tre år etter at redaksjonen hadde levert frå seg band 4, kom band 5 (*harm-jåttut*).

I 2011 kom det seinaste bandet av ordboka til nå (band 9, *ramost – skodda*), og med det er 3/4 av utgjevingsarbeidet fullført. Prosjektet har så langt levert i tråd med den oppsette og bindande produksjonsplanen. Staben omfattar nå (febr. 2011) 29 redaktørar, og tre av desse, Oddrun Grønvik, Lars S. Vikør og Dagfinn Worren, er hovudredaktørar. I tillegg til hovudredaksjonen i Oslo har prosjektet ei mindre redaksjonsgruppe i Trondheim. Desse har arbeidsplass ved Noregs Teknisk-Naturvitskaplege Universitet, NTNU. Vidare har

prosjektet ein administrativ stab på tre heiltidstilsette medarbeidarar, i tillegg til prosjekt-direktøren.

5. Meir om ordboka

5.1 Innhaldet i Norsk Ordbok

Norsk Ordbok har eit oppdrag og eit innhald som er svært komplekst. Ordboka dokumenterer bruken av ord i nynorske tekstar frå 1850 og fram til i dag. I tillegg dokumenterer ordboka det norske folkemålet, med blick attende til 1600-talet. Korleis ser tilhøvet mellom det deskriptive, preskriptive og normative ut i artikkelmaterialet vårt? Ordboksartiklane gjev normalt ikkje utfyllande fleksjonsopplysningar til oppslagsordet i toppdelen av artikkelen. I eit eige felt blir det gjort greie for dialektal fonologisk og morfologisk variasjon som går ut over det som er sjølvgeivne eller føreseielege former. Dette med sjølvgeivne former inneber at me i utgangspunktet ikkje fører opp målføredrag som er systematiske i dialektane, slik som lenisering $p, t, k > b, d, g$ i Agder-fylka og Rogaland, eller regelrett palatal uttale av l, n og til dels d og t etter kort vokal. Formvariasjon som skal gjerast greie for, er det som ikkje er heilt føreseieleg, t.d. bruk av /dd/ eller /dl/ for ll, som i *fjedd, fjødd* og *fjedl* for *fjell*.

I eksempeldelen av artiklane blir alle bruksdøme frå dialektkjelder i ordboka normerte etter 1938-normalen, slik også redaksjonelle døme blir. Sitat frå skriftlege kjelder blir sjølvstykke gjevne att i uendra form. Artiklane skal dekkja faktisk språkbruk, slik han kjem til uttrykk gjennom materialet og på ein representativ måte. Dermed blir resultatet at det nynorske ordtilfanget blir dekt i svært stor breidd.

5.2 Kven er mottakaren?

Noreg er eit land med sterk og levande dialektinteresse. Dette kjem til uttrykk gjennom at det over heile landet finst lokalt organiserte innsamlingar av dialektmateriale, gjennom utgjeving av lokale ordsamlingar og gjennom sterk oppslutnad om dialektspalter i lokalaviser, for å nemna noko. Det er mange uløna eldsjeler som arbeider med innsamling og kartlegging av lokale dialektar, og dei leverer verdfullt materiale til alle som forskar i dialektologiske emne. Erfaringa Norsk Ordbok 2014 har med formidling til ålmenta er at interessa for folkemålet er stor og levande, også blant dei yngre

mottakarane. Denne sterke og levande interessa for dialektane blei også understreka av Sigmund Skard (1932) i planen han sette opp for arbeidet med ordboka (jf kap. 3.1), og denne var utgangspunktet for at så mykje av innsamlingsarbeidet frå 1930-talet kunne organiserast som dugnadsprosjekt.

Norsk Ordbok 2014 arbeider på oppdrag frå Kulturdepartementet og det norske Stortinget. Dei løyver frå år til år midlar til arbeidet, og dei gjev uttrykk for at dei ønskjer denne dokumentasjonen av det norske folkemålet og det nynorske skriftspråket formidla attende til til ålmenta på ein breiast mogleg måte. Norsk Ordbok blir fortløpande gitt ut i prenta versjon. Fordi prosjektet er finansiert av offentlege midlar og med vesentlege bidrag frå ålmenta, ønskjer me å formidla ordboka attende til så mange som mogleg. Difor planlegg me i tillegg ein digital, nettilgjengeleg versjon av Norsk Ordbok, ved sida av den prenta. Så langt har me gjort kjeldene til Norsk Ordbok tilgjengelege på nettet (www.no2014.uio.no), men ein nettpresentasjon av sjølve ordboksartiklane er også under planlegging.

5.3 Den vitskaplege tilnærminga

Norsk Ordbok er, som nemnt på s. 51, ei vitskapleg dokumentasjonsordbok. Dette stiller strenge krav til kvalitetssikring av grunnlagsmaterialet, og det stiller krav til etterprøvbarheit. Alt kjeldematerialet som blir brukt i ordartiklane (både skriftspråklege kjelder og talemålskjelder) skal vera kvalitetssikra og vurdert etter vitskaplege kriterium. I tillegg har me gjort store delar av kjeldegrunnlaget vårt elektronisk tilgjengeleg, slik at ordboksbrukarane skal ha høve til på ein enkel måte å “sjå oss i korta”. På den måten får brukarane av ordboka moglegheit til å vurdere fortolkinga redaktørane har gjort av grunnlagsmaterialet.

Artiklane omfattar kartlegging av eldre kjelder, vurdering av ordetymologi, analyse av dialektal variasjon, semantisk analyse osv. Arbeidsforma i Norsk Ordbok 2014 inneber at redaktørane er generalistar og arbeider med alle delar av ordartiklane innanfor dei alfabetbolkane dei har ansvar for. Dette krev at heile redaktørstaben må vera grundig skolert både i tekstarbeid og tekstkritikk, dialektologi og diakron språkvitskap, og ikkje minst i semantisk analyse. Redaksjonen er samansett av forskarar med litt ulike spesialfelt, og både semantikarar, grammatikarar, dialektologar, edisjonsfilologar og etymologar er med i redaksjonen.

5.4 Norsk Ordbok på fullelektronisk plattform

Dei tre første banda av Norsk Ordbok er i all hovudsak handarbeid. Materialet som ligg under artiklane er, i tillegg til grunnmanuskriptet, ikkje minst store mengder ordsetlar. Desse blei sorterte og systematiserte på ordbokredaktørens skrivebord, og manuskriptet blei skrive for hand eller etter kvart inn i eit ordi-nært tekstbehandlingsprogram. Denne arbeidsforma kravde at redaktørane hadde skaffa fram fullstendig oversyn over alle deler av kjeldetilfanget, over ei stor mengd avstyttingar og redigeringskonvensjonar osv. Det seier seg sjølv at arbeidet måtte bli svært tidkrevjande.

Med den nye prosjektorganisasjonen kom det på plass ei digital plattform og ein struktur med eit sett av relasjonelle databasar der ulike delar av materialet er kopla saman og systematisert i ein elektronisk indeks. Denne måten å organisera materialet på gjer hovudmengda av materialet digitalt tilgjengeleg for redaktørane (og for ålmenta). Redigeringa skjer inn i ein fullelektronisk og skreddarsydd redigeringsapplikasjon som kommuniserer med den elektroniske indeksen og alle deldatabasane. Applikasjonen omfattar felt med spesifikke avgrensingar, og det inneheld dei mest brukte avstyttingane og konvensjonane i lister og rullemenyar redaktørane kan velja frå. Dessutan inneheld redigeringsapplikasjonen m.a. systematiske hierarki over heimfestingar der programmet kan utføra automatiske genereringar.

Hovudgrunnen til at dette er mogleg, er den massive digitaliseringa av materiale som gjekk føre seg i 1990-åra gjennom Dokumentasjonsprosjektet (Dokpro). Dette prosjektet utførte storskala-digitalisering av kulturhistoriske og språklege samlingsmateriale ved dei norske universiteta. Føremålet med Dokpro var å gjera kulturelatert informasjon tilgjengeleg både for forskarsamfunnet, for forvaltninga og for ålmenta. Dette skulle skje i ei form som kunne gjera det mogleg å henta inn og nytta informasjon frå fleire ulike kjelder og register samstundes.

Som ein del av dette arbeidet blei heile Norsk Ordboks setelsamling digitalisert med faksimile og oppslagsord, med ordklassemerking og med uttømmande kjeldeopplysningar. Grunnen til at det blei gjort på nett denne måten, er at setelmaterialet til Norsk Ordbok er svært heterogent. Materialet er samla inn over svært mange tiår (arbeidet strekte seg over ein tidsperiode på meir enn 60 år på det tidspunktet då Dokpro starta digitaliseringa). For ein svært stor del er

setlane handskrivne, noko som legg klåre avgrensingar på korleis det kan digitaliserast. Også grunnmanuskriptet til Norsk Ordbok er digitalisert, saman med fleire av dei andre kjeldene som ligg til grunn for ordboka.

I den moderne prosjektorganisasjonen ligg brorparten av grunnlagsmateriale som nemnt ovanfor systematisert og elektronisk kopla opp mot redigeringsprogrammet vårt i ein kjede av relasjonelle databasar. I ein eigen indeks, som me har gitt namnet *Metaordboka*, er materialet frå dei ulike datakjeldene sortert under normerte oppslagsord.

I redigeringsapplikasjonen, der ordboksartiklane blir utarbeidde, har kvar redaktør sin eigen signatur. Samtlege ferdige elektroniske artiklar er merkte med namnet til redaktøren som har skrive artikkelen. Redaktørane redigerer ordboks-artiklane rett inn i sjølve ordboksdatabasen, og programfunksjonen gjer det mogleg å generera opp einskildartiklar eller større alfabetbolkar direkte i trykk-sats, gjennom ein eigen utskriftsfunksjon.

5.5 Inn i ordbokslaboratoriet

For å forklåra nærare korleis ordutvalet og artikkelproduksjonen går føre seg, må me ta ein tur inn i sjølve ordbokslaboratoriet. Til grunn for utvalet av ordartiklar ligg den elektroniske indeksen (*Metaordboka*), som materialet frå databasane er kopla opp mot (jf kap. 5.4). I denne indeksen er det om lag 600 000 normerte oppslag. Ikkje meir enn halvparten av desse kan få plass i den prenta utgåva av Norsk Ordbok, så utvalskriteria må vera strenge.

Når ein redaktør går inn i redigeringsprogrammet og genererer opp artiklar i "sin" bolc, vil programmet automatisk gjera eit utval av artiklar redaktøren får tilråding om ikkje å redigera. Likevel er det redaktøren som har ansvar for å gjera den endelege vurderinga av dei einskilde orda i bolken, og han eller ho kan sjølvstøvt overstyra programfunksjonen på dette punktet. Dei forekomstane frå den elektroniske *Metaordboka* som blir utelukka i første utvalsrunde, er litterære hapax (einskildforekomstar av samansetningar og avleiingar), reine ordboksord som aldri har fått gjennomslag i språkb Bruken og einebelegg av dialektord som er oppførde med manglande eller ufullstendige opplysningar.

Ikkje bare utvalet, men også storleiken på artiklane blir elektronisk styrt. Ein logaritmefunksjon gjer ei utrekning av artikkel-lengd ut frå talet på registrerte forekomstar av eit ord i *Metaordboka*,

slik at redaktøren får ei tilrådd artikkel-lengd når han eller ho genererer opp sjølve ordartikkelen. Denne utrekninga baserer seg på forholdstal, og på dei absolutte grensene som er sette av talet på attståande ordboksband og talet på sider me har tilgjengeleg i produksjons-perioden. Redaktøren står fritt til å bryta den tilrådde artikkellengda, men dersom ordartikkelen blir på fleire liner enn det tilrådinga seier, må denne overproduksjonen takast inn att ein annan stad i alfabetbolken eller i bandet.

Ofte vil redaktøren kunna spara inn att det som blir brukt av ekstra plass på grunnordsartiklar når han eller ho startar redigeringa av fullstendig gjennom-siktige samansetningar som høyrer til det aktuelle grunnordet. Slike saman-setningar kan leggjast inn som døme i føreleddsartiklane i ordboka. Eit døme på ein føreledds-artikkel er artikkelen *krise*-. Denne omfattar eit sett fullstendig analyserbare samansetningar (som *krise-psykologi*, *krise-dempande* og *krise-situasjon*). Eit anna døme er føreleddsartikkelen *makt*-, som omfattar saman-setningar som *makt-demonstrasjon*, *makt-misbruk* og *makt-monopol*. Desse samansetningane, som er førde opp i føreleddsartikkelen, får då ikkje egne ord-artiklar i tillegg.

Norsk Ordbok omfattar ikkje proprium, og fordi det er ei allmennordbok, er ordboka restriktiv når det gjeld ord som bare høyrer heime i fagspesifikke spesialistordordforråd. Her er grenselandet mellom fagspesifikk terminologi og allmenn bruk ikkje alltid like klår. Redaktørane må vurdera frå førekomst til førekomst, og diakront ser me at ord kan flytta seg frå det spesialiserte og over i meir allmenn bruk. Dette har skjedd med eit ord som *tsunami* m., som etter kvart er så vanleg i bruk at ordet etter kvart er blitt ein rik metaforbase. Det kjem til uttrykk i bruksdøme frå grunnlagsmaterialet, slik som “[d]inest kom dissens-tsunamien under Borten, før det tilsynelatande roa seg med to dissensar i Brattelis fyrste regjering” (Søberg 2009: 45) og “Er ’Yngrebylgja’ ein tsunami eller ein brottsjø?” (Uniforum 2009). Det første dømet er henta frå Nynorsk-korpuset, medan det andre er registrert på ein ordsetel i det elektroniske setel-arkivet. Slike belegg dokumenterer at dette ordet har flytt seg over frå eit spesialistordforråd og over i allmennspråket.

Når me vel ut lemma, er me dessutan varsame med å ta med slangord. Grunnen til dette er at dei ofte viser seg å ha kort levetid, og i mange tilfelle er dei knytte til svært små sosiale grupper. Ord som skal ha ein plass i Norsk Ordbok må vera meir bestandige og ha ei anna type utbreiing enn det mange av desse slangorda viser seg å

ha. Av slangprega ord og utviklingar me likevel dokumenterer, er slike som slår sterkt og tydeleg inn i materialet vårt over tid, og som t.d. viser interessant grammatisk produktivitet, slik som med substantiva *kanon* og *krise*. Det første eksisterer i ordboka med ein eigen artikkel III *kanon* adj., med døme som “å drikka seg *kanon*”, “å få det heilt *kanon*”. Det andre er registrert i adjektivisk bruk som t.d. “ganske *krise*”.

Heller ikkje eksotika får utan vidare plass i Norsk Ordbok. Likevel er det slik her som i fleire av dei andre nemnde kategoriane at kva som fell inn i denne gruppa, kan endra seg over tid. Då band 4 av Norsk Ordbok blei redigert, fekk ikkje substantivet *hijab* m. plass i lemmautvalet. Dette har heilt klårt samanheng med at det på dette tidspunktet ikkje var noko materialgrunnlag i det nynorske kjeldetilfanget som tilsa at ordet var i utbreidd bruk i norsk. Dersom dette bandet hadde blitt redigert i dag, er det svært sannsynleg at *hijab* hadde fått ein ordartikkel, fordi me ser av kjeldematerialet vårt at ordet er i ferd med å bli godt dokumentert i nyare nynorsk tekstproduksjon.

5.6 Hovudkjeldene til Norsk Ordbok

5.6.1 Målførekjeldene

Svært mykje av målførematerialet som er tilgjengeleg for redigeringa av Norsk Ordbok, ligg i det elektroniske setelarkivet. Dette elektroniske arkivet omfattar dei gamle handskrivne og maskinskrivne ordbokssetlane i digitalisert form, og i tillegg inneheld databasen nye elektroniske setlar som er lagde inn i tida etter 2002. Dette materialet er i hovudsak godt kvalitetssikra, det omfattar stadig nye forekomstar, og denne deldatabasen inneheldt ved utgangen av 2009 i underkant av 3,5 mill. setlar. Denne delen av materialet aukar med om lag 15 000 setlar for kvart år.

Ei anna svært viktig kjelde til målføreopplysningar er målføresynopsien. Dette er eit systematisk oversyn over målførevariasjon (fonologisk, morfologisk og noko syntaktisk variasjon) som er samla inn herad for herad frå heile landet. Innsamlinga og systematiseringa av materialet i målføresynopsien er baserte på strukturen og eksempelorda i Storms ordliste, og materialet er samla inn i perioden 1950-1970.

Målføredelen av materialet som ligg til grunn for ordartiklane i Norsk Ordbok omfattar også ei større mengd lokale ordsamlingar

som kompletterer det innsamla materialet frå setelarkivet og synopsisen. I samråd med utgjevarane av slike samlingar forsøker me så langt me har rammer til det å tilgjengeleggjera desse elektronisk gjennom Ordbokshotellet. Dette virtuelle “hotellet” utgjer ein eigen deldatabase kopla opp mot Metaordboka.

5.6.2 Skriftspråklege kjelder

Norsk Ordboks setelarkiv omfattar i tillegg til målføredata også innsamla ord og brukseksempel frå nynorsk litteratur. I arbeidet med artiklane bruker redaktørane dessutan setelmaterialet frå den nynorske delen av Det norske litterære ordboks-verk. Dette siste delarkivet er dessverre ikkje elektronisk tilgjengeleg.

Ei stor og stadig veksande hovudkjelde til dokumentasjon av skriftleg nynorsk er Norsk Ordboks nynorske tekstkorpus (Nynorsk-korpuset). Dette er eit monitorkorpus som ved utgangen av 2009 omfatta over 41 mill. ord, og denne store tekstsamlinga er med på å gje eit meir komplett bilete av tekstproduksjonen på nynorsk. Korpuset inneheld tekst frå alle typar skriftspråklege sjangrar, og det omfattar tekster frå den tidlege nynorsken på 1860- og -70-talet og heilt fram til i dag. 75 % av materialet i korpuset er frå tida etter 1975.

Nynorskkorpuset er viktig for redigeringa av Norsk Ordbok av fleire grunnar. Det gjev innsyn i frekvente kollokasjonar frå andre typar språk enn det setelsamlingane dekkjer. Funksjonsord som er høgfrekvente i allmennspråket (preposisjonar, konjunksjonar osv.) er ofte ikkje så breitt representerte i inn-samla folkemålsmateriale som innhaldsord (substantiv, adjektiv, verb) er. Her vil eit representativt tekstkorpus vera til stor hjelp for redaktørane.

Det store omfanget av ordbelegg i korpuset vil også bidra til å gjera ordartiklane endå meir stringente. Tekstmengdene gjer det mogleg å få eit endå betre oversyn av t.d. kva for kollokasjonar som er høgfrekvente. Sidan korpuset inneheld så mykje nyare tekstproduksjon på nynorsk, vil det også bidra med både ord og ordtydingar som ikkje er like godt dekte i det eldre materialet.

Eit døme frå band 6 kan illustrera korleis det innsamla målførematerialet og Nynorskkorpuset kompletterer kvarandre og gjev oss eit endå betre bilete av nynorsken og folkemålet: Homografane I **krım** m. og II **krım** n. Artikkelen I **krım** m. (figur 1) inneheld tydingsbolken 3a) 'kriminallitteratur'. I tillegg er det denne definisjonen som ligg til grunn i førsteleddet i samansetningar som

krim-dronning og *krim-litteratur*. Både grunnordet I *krim* m. og samansetningane er hovudsakleg dokumenterte i Nynorskorpuset og det nyare kjeldematerialet.

I *krim* m [krim; avst for *kriminal-* og *kriminalitet*] 1 *kriminalitet*: *tabloidavisene skriv helst om krim, sport og kjendisar*. 2 *kvard.*, i bf, avdeling i politiet som arbeider med *kriminalitet*: *ho arbeider på krimmen* (jfr Tryfisleng 35 og 38). 3 a *kriminallitteratur*: *ho las romanar, lyrikk, fagbøker og også krim* (Firda 2000). b *roman* el *forteljing* skriven innanfor genren *kriminallitteratur*: *ein krim i den hardkokte stilen* (Strilen 1985); òg i sms som *påske-krim*. c *kriminalforteljing* som er dramatisert for radio el fjernsyn, ofte som *føljetong*: *radioteateret sender krim kvar laurdag*.

Figur 1: Artikkelen I *krim* m.

Motsett er det med artikkelen II *krim* n (figur 2 neste side). Han hentar hovudsakleg stoff frå tradisjonsdelen av materialet, og vesentlege deler av beleggmengda skriv seg frå setelarkivet og dei eldre kjeldene (jf også dei mange oppføringane av eldre belegg i hakeparentesen i artikkelen). Tydingsnummer 2 i artikkelen har definisjonsteksten 'forkjøling, krimsykje (1)', og denne tydinga finn me òg att som førsteleddet i samansetningar som *krim-full* ('sjuk av forkjøling') og *krim-farang* (der første del av definisjonen er 'farang 1, epidemi av forkjøling').

Kva typar språkendingar og samfunnsendingar vil verka inn på utval og redigeringsprinsipp i eit verk som blir til over så lang tid, og korleis skal ei ordbok som er produsert over ein tidsperiode på meir enn 60 år både ta omsyn til diakron utvikling og endringar i materialtilfanget, samstundes som ho tek omsyn til samanheng og konsistens? Artiklane i ordboka vil alltid til ein viss grad vera prega av tida dei er skrivne i, både når det gjeld utval (jf det nemnde *hijab* m. ovanfor) og innhald. Dersom ordartikkelen *neger* m. var blitt skriven 50 år attende i tid, ville han truleg sett noko annleis ut enn det han gjer i band 8 av Norsk Ordbok, som kom ut i 2009. I denne ordartikkelen er det lagt vesentleg vekt på det nedsetjande i bruken av ordet, men dét kjem ikkje til uttrykk i dei eldre delene av materialet.

II krim (*l*) n [*kri:m*; Fyresdal1698 og L 41 *kriim*, F 124, Lom1743 *kriim*, Va1743, Tinn1743 *krjm*, P *kriim*, Snn-1762, W og Ws *krim*, Seljord1786, VSO, H, NoOrds-1810-12 34, VTel1821 48 *kreem*, C, FayeD1840 u *snue*, BrgKøster, A; målf òg *kreim* (Finsland, SAudnedal, Eiken, Aurland); målf òg genus m (Hof i Sol, Tinn, Forsand, MøRo, Trl); f (Osen, Grong); jfr fær *krim* 'snue', jfr òg isl *augnakrim* 'augebetennelse'] **1 a** slim i nasen (serl ved forkjøling); snue (F; P; Ws; C; FayeD-1840; A; Hof i Sol, Frøn, Rog, Brg, Stryn, Hareid, Frøya, Meldal, Snåsa, NRana): *det fylgjer med krim når ein er forkjøla* (Forsand) / *ta krimet or nasen på deg* (Karm). **b** slim i luftvegane (Sodal, Stange, NGbr, Va, Ådal, Helleland, Vanylven, Astafjord; R.-Kj. VGTIV,84): *ha krim i bringa / krim er det ein hostar opp når ein er forkjøla* (ØSlidre) // i samb **krim og kov** slim og hoste (Tv. BP 64). **c** råke i munnen; spytt (Klinga). **2** forkjøling, krimsykje (1) (Fyresdal1698; L; W; Seljord1786; VSO; H; VTel1821; Dovre, Heidal, NFrøn, Gausdal, NLand, Hemsedal, Tel, Froland, Fjære, Set, SAudnedal, Eiken, Ryf, Brg, Breim, Vanylven, Surnadal, Stadsbygd, Oppdal, Grong; DølenOrdf): *så er det forkjølelsa da, eller krimen* (ÅrbMeldal 1968,42) / *nasen renn og halsen er sår når ein har krim* (Stadsbygd). **3 a** katarr el betennelse i luftvegane (A²; NoSold 269): *fyr hosten og krimen (katarrhen) gjev ein ? varm mjølk* (Hand. VS 52) / *eit av borni døydde av hardt krim (eller pleurit)* (Be. VR III,269); òg i sms som lunge-, nase-, øyre-krim. **b** strupehoste (Lauvd IV, 85).

Figur 2: Artikkelen *II krim m*.

Fordi me nå produserer ordboka på ei fullelektronisk plattform, blir alle artiklane lagra i ein elektronisk database. Det inneber at ein i framtidige revi-sjonar og i framtidig bruk av artikkelmaterialet i andre kontekstar vil ha høve til å arbeida med eller sortera ut artiklar innanfor t.d. semantisk avgrensa område. Dette er ei stor endring frå

tidlegare tiders leksikografi, der papirformatet gjer at ein må leggja til grunn éi (oftast alfabetisk ordna) sorteringsrekkefølgd.

6. Meir om arbeidsforma i den moderne prosjektorganisasjonen

I 2009 rådde prosjektet Norsk Ordbok 2014 over 27 ordboksredaktørar. Tre av desse er hovudredaktørar med kvart sitt særlege faglege ansvarsområde i redaksjonen (jf pkt. 4). I prosjektorganisasjonen er alle redaktørane generalistar, og det inneber at dei redigerer heile artikkelinnhaldet innanfor alfabetbolkane dei har ansvar for. Artikkelarbeidet er meritterande forskingsproduksjon for den einskilde redaktøren.

For prosjektorganisasjonen er det absolutte føremoner med å organisera det faglege arbeidet på denne måten. For det første vil redaksjonen bli mindre sårbar mot utskiftingar i staben dersom alle redaktørane er kompetente generalistar. For det andre blir det mindre dobbeltarbeid – som det fort kunne blitt om éin redaktør hadde ansvar for målføredelen av artikkelen, ein annan stod for den semantiske analysen og ein tredje hadde ansvar for etymologidelen. Dette ville ført til at minst to av redaktørane, og kanskje alle tre, måtte arbeida seg gjennom det same grunnlagsmaterialet, men på leit etter litt ulik informasjon. Eit tredje hovudargument mot spesialistredigering er at ord ofte dannar semantiske klynger. Dersom redaksjonen var organisert på den måten at nokre redaktørar bare arbeidde med funksjonsord, medan andre var verbspesialistar, og andre att arbeidde med substantiv og adjektiv, ville me fort mista samanhengane mellom orda i desse semantiske klyngene, og ikkje minst dei til dels glidande og fagleg interessante og utfordrande overgangane mellom ordklassekategoriane.

Eit eksempel frå band 6 kan tena til å illustrera dette poenget. I *kri*-bolken er det ei semantisk klynge av ord med ulik ordklasse-tilhøyrslé, men med heilt klåre samanhengar seg imellom. I denne klynga finst det både funksjonsord og innhaldsord, og eit utval er presentert nedanfor:

I <i>kring</i> m.	tyding 1. 'krins, ring'
II <i>kring</i> adj.	tyding 1. 'ledug, bøyeleg'
IV <i>kring</i> prep.	tyding 1. 'rundt om, om'
<i>kringa</i> v.	tyding 1. 'runda, gå i ring rundt'

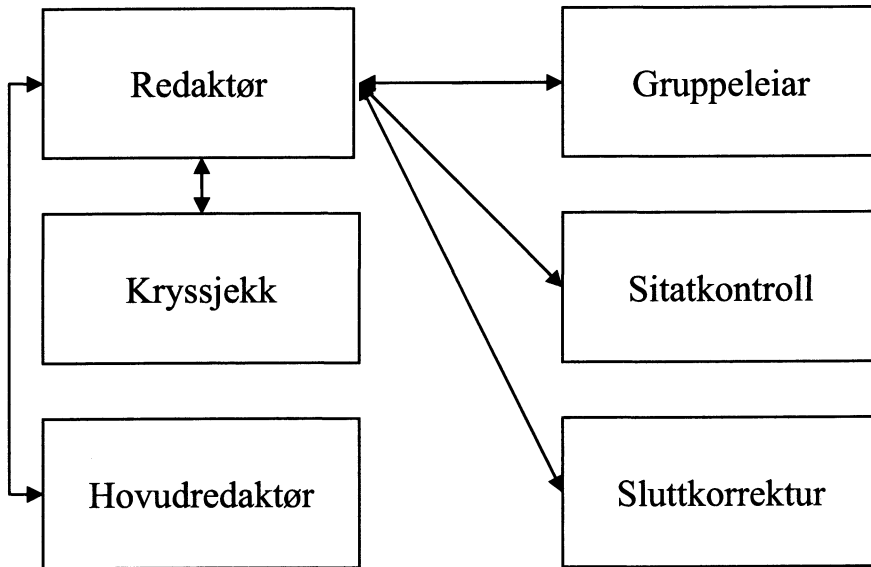
<i>kringla</i> v.	tyding 1. 'leggja i ring, kveila saman'
<i>kringle</i> f	tyding 1. 'ring, rund skive eller sirkel'
<i>krins</i> m.	tyding 1. 'ring, sirkel'

I tillegg til desse finst det fleire andre (m.a. *kring* adv., *kringe* m., *krings* m., *kringel* m., *kringgå* v. osv.), som også gjeld objekt som er runde eller rundvorne, som dannar ein sirkel eller som skildrar ei rørsle i ring, i sirklar om noko, eller som omsluttar osv. Dette store materialet med semantisk nærskylde ord frå mange ordklassar viser føremonene med at heile komplekset blir handsama under eitt, og av ein og same redaktøren.

For at me skal komma i mål med Norsk Ordbok i 12 band til grunnlovs-jubileet i 2014 er det ein føresetnad at me har oversikt over alt materialet og har ein plan for korleis det skal få plass innanfor det avgrensa prenteformatet. For å kunna vera sikre på å både få med dei siste bokstavane i alfabetet, samstundes som me ikkje sprengjer grensa på 12 band (kvart med 800 sider), må ordartiklane dimensjonast, slik at det er ei klår avgrensing av artikkellengd. Dette inneber at me kan måla progresjonen i arbeidet fortløpande, og dermed vita nøy-aktig kor mykje arbeid som står att i kvar bandproduksjon. I praksis ser me òg at det å ha avgrensingar i formatet ofte vil kunna gje strammare artiklar. Det er slett ikkje utenkeleg at stringensen i ordartiklane til og med kan bli endå høgare av det systematiske formatet redigeringsprogrammet representerer, og av den avgrensinga i linetal som er sett for kvar av artiklane.

Redaksjonen er organisert i redaksjonsgrupper med tre gruppe-medlemmar og éin gruppeleiar, og gruppeleiarane står for første runden med kollegalesing av manuskriptet redaktørane i gruppa produserer. På vegen fram mot den endelege og prenteklare versjonen av ordartiklane passerer manuskripta gjennom ei heil rekkje kvalitetssikringsinstansar (jf fig. 3 på neste side). Alle desse kvalitetskontrollane skal sikra at manuskriptet er mest mogleg heilskapleg og gjennom-arbeidd når me sender det til prenting. Ei eiga kryssjekkgruppe sjekkar at alle dei formelle krava til manuskriptbolken er oppfylte – at obligatoriske elektro-niske koplingspunkt mellom artiklar er lagde inn, at alle relevante eldre kjelder er med i hakeparentesdelen av artiklane, og at alle tilvisingar er korrekte, for å nemna noko. Sitatkontrollen syter for at alle litteraturtilvisingar og sitat er korrekte, og hovudredaktøren les alt materialet i ei siste fagleg kvalitetssikring av innhaldet. Mellom desse instansane er

manuskriptet til oppretting hos redaktøren, før det så endelig kan leggest til sluttkorrektur.



Figur 3: Kvalitetssikringa av manusarbeidet. Manuskriptet går fram og tilbake mellom redaktøren og kvalitetssikringsinstansane.

7. Avrundande kommentarar

Innhaldet i det daglege arbeidet i prosjektorganisasjonen Norsk Ordbok 2014 er produktet av eit nært samarbeid mellom IT-utviklarar og leksikografar. Ordboka blir produsert som ein elektronisk database, og hovudmengda av grunnlagsmateriale er digitalisert. Den digitale redigeringsplattforma syter for at arbeidet kan gjerast meir effektivt enn tidlegare, m.a. i form av at ulike typar avstytingar, kjeldetilvisingar, heimfestingar og andre typar faste entitetar må veljast frå lister. Alle redaktørane redigerer inn i ein og same databasen, og kan fortløpande laga elektroniske koplingspunkt mellom samtlege artiklar (og artikkeldelar) i denne basen. Redigeringsverktøyet, saman med ein detaljert redigeringsmanual (*redigeringshandboka*), sikrar utvalet av lemma, sikrar at rekkefølga av opplysningar i artiklane er standardisert, sikrar at same type opplysningar blir presenterte likt frå artikkel til artikkel, sikrar at tilvisingar frå ein artikkel til ein annan ikkje kan vera tomme og

sikrar at samansetningar ikkje kjem på prent både som del av rekkjer i føreleddsartiklar og samsundes i eigne ordartiklar, for å nemna nokre viktige punkt.

Ein meir langsiktig vinst av å redigera ordboka som ein full-elektronisk database er at dette gjer det mykje enklare å vidareutvikla og raffinera det endelege produktet, og å laga nye leksikografiske produkt med utgangspunkt i det ferdige arbeidet.

Litteratur

- Bø, Reidar. 1989. Arbeidet med Norsk Ordbok. I *Almenningen*, Olaf & Oddrun Grønvik (red.). 1989. *Ord og mål. Festskrift til Magne Rommetveit*, 80–94.
- Norsk Ordbok. Ordbok over det norske folkemålet og det nynorske skriftspråket*, bd. 1–8. Oslo: Det Norske Samlaget
- Ross, Hans. 1895. *Norsk Ordbog. Tillæg til "Norsk Ordbog" af Ivar Aasen*. Kristiania: Cammermeyer.
- Schjøtt, Steinar. 1914. *Norsk ordbok med ordtyding paa norsk-dansk*. Oslo: Det Norske Samlaget.
- Skard, Matias. 1954. *Nynorsk ordbok for rettskriving og litteraturlesnad*. 5. utgåve. Oslo: Aschehoug
- Skard, Sigmund. 1932. *Norsk Ordbok. Historie – plan – arbeidsskipnad*. Oslo: Det Norske Samlaget.
- St.meld. 22 (1999-2000). *Kjelder til kunnskap og oppleving*. Om arkiv, bibliotek og museum i ei IKT-tid og om bygningsmessige rammevilkår på kulturområdet. Tilråding frå Kulturdepartementet 17. desember 1999.
- St.meld. 35 (2007-2008). *Mål og meining*. Ein heilskapleg norsk språkpolitikk. Tilråding frå Kultur- og kyrkjedepartementet 27. juni 2008.
- Søberg, Morten. 2009. *Senatet som aldri vart*. Essay frå Stortinget. Oslo: Det Norske Samlaget
- Torp, Alf. 1919. *Nynorsk etymologisk ordbok*. Oslo.
- Torp, Arne & Lars S. Vikør. 1993. *Hovuddrag i norsk språkhistorie*. Oslo: Ad Notam Gyldendal.
- Vidsteen, Chr. 1900. *Ordbog over Bygdemaalene i Søndhordland*. Bergen.
- Aasen, Ivar. 1873. *Norsk Ordbog. Med dansk Forklaring*. Christiania.

Del B

Andre relevante dokument

NYNORSK FORUM

Samarbeidsorgan for nynorske institusjonar, organisasjonar, bedrifter

c/o Nynorsk kultursentrum, 6160 HOVDEBYGDA

Telefon 70 04 75 75, telefaks 70 04 75 71

ottar.grepstad@aasentunet.no

Politisk leiing
Kulturdepartementet
Boks 8030 Dep
0032 OSLO

Hovdebygda, 4. juli 2011

NYNORSKPRIORITERINGAR I STATSBUDDSJETTET 2012

Nynorsk Forum takkar hjarteleg for eit ope og konstruktiv møte med statssekretær Lubna Jaffery 9.6. I tråd med det som der blei avtalt, gjorde vi på møte i Nynorsk Forum 23.6. ei samla prioritering av nynorsktiltak i statsbudsjettet 2012.

Nynorsk Forum set stor pris på dei løfta fleire nynorskinstitusjonar har fått dei siste åra, og det har både vi og enkeltinstitusjonane uttrykt tydeleg mange gonger. Vi tek ikkje for gitt at ein slik vekst kan halde fram, men legg også vekt på at dei uløyste oppgåvene framleis er store. Det aukande politiske presset mot plassen for nynorsk i skulen kan ikkje åleine løysast gjennom løyvingar til institusjonar, men det er på det reine at handlekraftige institusjonar maktar å utvikle meir dokumentasjon, argumentasjon og aktivitet enn om rammene er tronge.

Vi er inneforstått med at den dokumentasjonen vi la fram i møtet 9.6., kan tolkast og bør tolkast, særleg av di fleire område kan ha kome inn under KUD i den aktuelle perioden. Det står likevel fast at justert for kpi auka budsjettet for KUD med 37 % frå 2006 til 2011 (frårekna Den norske Kyrkja). I same tidsrommet auka løyvinga til nynorskinstitusjonane med 22 %.

Korleis vi enn snur og vender på desse tala: Det trengst ei vidare sterk nynorsksatsing i tråd med framlegga våre nedanfor. Statsbudsjettet 2012 bør innehalde ein tydeleg og romsleg nynorskpakke på minst 25 mill. kr. Om lag 90 % av dei statlege løyvingane til nynorskinstitusjonar kjem frå KUD-budsjettet. Difor legg vi mest vekt på desse nedanfor.

1. Løyvingar til spesifikke tiltak

Nynorsk Forum legg heile tida til grunn at etablerte tiltak skal sikrast før nye kjem til. I prioritert rekkjefølgje rår Nynorsk Forum til slike påplussingar i 2012:

	<i>Løyst 2011</i>	<i>Behov 2012</i>	<i>Auke frå 2011</i>
1. Språkåret 2013	0,7	2,4	1,7
2. Innholdsprosjekt Språkåret 0		4,0	4,0
3. Det Norske Teatret			2,0 ekstraløyving
4. Norsk Ordbok 2014	15,1	16,3	1,2
5. Olav H. Hauge-senteret	1,0	2,2	1,2
6. Nynorsk Pressekontor	4,3	5,8	1,5

1. *Språkåret 2013* har eit høgt ambisjonsnivå: Det blir invitert breitt til å bli med på noko som ingen har gjort før og som ingen kjem til å gjere oss etter. Prosjektleiars Jens Kihl tok til i eit seks månaders engasjement 1.6., og etter avslutta arbeid med 2012-normalen for nynorsk har Grete Riise takka ja til å leie ei breitt samansett styringsgruppe for Språkåret. Alt dette og meir til

blir utført innanfor løyvinga frå KUD i 2010. Styret i Nynorsk kultursentrum har vedteke at prosjektplan og prosjektbudsjett skal liggje føre innan utgangen av januar 2012. Desse dokumenta blir grunnlaget for ein stor søknad om prosjektmidlar i både 2012 og 2013. Prosjekteringa av Språkåret må halde fram i 2012, og det trengst alt det året midlar til avtalar om større innhaldsprosjekt som krev lang førebuingstid.

2. *Det Norske Teatret* fyller 100 år i 2013. Dette er tenkt som ein integrert del av Språkåret 2013, men føreset også frie midlar som DNT disponerer. Skal teatret til dømes lage ein festival der ein inviterer nynorskproduksjonar frå andre teater rundt om i landet, lagar ikkje det seg sjølv og ikkje utan eit skikkeleg tilskot som dekkjer frakt, reiser opphald. Vi er kjende med at Nationaltheatret fekk ei liknande særloyving då dei jubilererte for 12 år sidan. Ei slik særloyving treng altså også DNT, og den treng dei i 2012.

3. *Norsk Ordbok 2014* er urolege for utviklinga i si løyving. Det er nødvendig alt no å planleggje eit post 2014-prosjekt for revisjon og fullintegrering av første delen av verket, til midt i band 5, av di oppslagsorda a–h er redigerte analogt. Dette vil gi ein heil database a–å som kan brukast til andre leksikografiske og språkteknologiske føremål. Eit slikt prosjekt vil også sikre arbeidsplassar og hindre personalflukt frå ein stab der berre to er fast tilsette, og dermed også sikre at band 12 faktisk blir ferdig i tide.

4. *Olav H. Hauge-senteret* i Ulvik fekk i 2011 ei samla løyving frå stat, fylke og kommune på 2,2 mill. kr. Viss arbeidet går etter planen, blir senteret offisielt opna sommaren 2013 eller våren 2014. Som for Ivar Aasen-tunet blir det føresett at Olav H. Hauge-senteret på få år når det planlagde nivået for tal tilsette og mengd aktivitet. Medarbeidar nr. 2 blei tilsett 1.7. Senteret har plass til fire fast tilsette i tillegg til funksjonar og kompetanse som blir dekt frå Ivar Aasen-tunet. Hausten 2011 blir fleire arbeidsgrupper oppnemnde for planlegging av innhald, utstilling og andre publikumstilbod.

5. *Nynorsk Pressekontor* treng 1,5 mill. kr ekstra til ekstra nyheitssatsing frå 2012. Tradisjonelle media produserer stadig meir innhald i stadig fleire kanalar, og fleire typar aktørar publiserer nyheiter. Det aller meste av dette kjem på bokmål, og sjølv aviser som på papiret er redigert på nynorsk, blir dermed i nettutgåvene tospråklege. Gjennom NTB kan NPK levere dei viktigaste nasjonale og internasjonale nyheitene på nynorsk, berre pengane og folka finst. Samstundes vil dei halde oppe og utvikle vidare sin eigen redaksjonelle identitet. NPK treng 1–2 redaksjonelle årsverk til å etablere ei teneste som kan vere attraktiv også for NRK og dei store tospråklege avisene. Eit prøveprosjekt blir sett i gang for eigne midlar hausten 2011.

2. Generell satsing på nynorskinstitusjonar

Nynorsk Forum rår til at basisløyvinga til nynorskinstitusjonane blir auka med 10 % i 2012. Dette gjeld særleg desse fem institusjonane:

- Det Norske Samlaget
- Det Norske Teatret
- Landssamanslutninga av nynorskkommunar
- Noregs Mållag
- Nynorsk kultursentrum

Ei slik satsing vil sikre grunnlaget for desse institusjonane i Språkåret 2013.

3. Nynorskscenen i Bergen

Nynorsk Forum ser positivt på planen om ein nynorskscene i regi av Hordaland Teater i Bergen. Nynorsk kultursentrum har tidlegare varsla KUD om at dei i søknaden for 2013 vil prioritere *Jon Fosse-arkivet*. Desse tiltaka kan det vere aktuelt å sjå i samanheng. Fosse er sjølv ikkje framand for at arkivet kan bli inkludert i nynorskscenen i Bergen.

4. Handlingsplan for nynorskopplæringa

Nynorsksenteret i Volda har vore ei positiv satsing. Senteret går i 2012 inn i siste året med ekstraløyvingar til den treårige tiltaksplanen for senteret. Store oppgåver må løysast både i opplæringa for elevar med nynorsk som hovudmål og nynorsk som sidemål. Nynorsk Forum strekar under at nynorskopplæringa må styrkjast gjennom ei mykje breiare satsing enn det Nynorsksenteret kan make åleine, både i lærarutdanninga og særleg i etter- og vidareutdanninga av lærarar. Vi meiner KUD og KD i fellesskap bør utvikle ein stor handlingsplan for å styrkje nynorskopplæringa for elevane og nynorskkompetansen til lærarane.

5. Tiltak utan budsjettkonsekvensar

Nynorsk Forum sorterte i 2009 alle dei nynorskrelevante tiltaka i språkmeldinga, og vi merka oss ei rekkje tiltak som ikkje har budsjettkonsekvensar. Vi viser til det dokumentet vi la fram på møtet 9.6. og ser fram til vidare dialog også om den sida av saka.

Vi utdjupar gjerne vurderingane av desse prioriteringane viss det er ønskjeleg, svarar når som helst på spørsmål og ønskjer lykke til!

Med venleg helsing

Nynorsk Forum



Ottar Grepstad, sekretær

Digitalisering og tilrettelegging for innlasting i ordboksbasen av Norsk Ordbok band 1-5 (a-h)

Digitalisering av teksten

Band 1-2 vart skrivne inn av E. Holm og korrekturlesne av AE ca 2001 – 2002 (før oppstart av prosjekt NO 2014). Band 3-4 og band 5 ut bokstaven H var skrivne inn på pc og tagga etter eller under innskriving i eit dataformat som var sett opp av Håvard Hjulstad (før 1987). Dette formatet har ei blanding av leksikografiske kategoriar og typografisk markering. Den som vil sjå på det, vil finna band 1-5-filene i V:\norsk-ordbok\NO\NO1-4. Teksten for NO a-h låg såleis føre i digital form før prosjektet NO2014 kom ordentleg i gang i 2002.

Føremål

Føremålet med å leggja inn NO a-h i ordboksdatatabasen var å gjera gamle NO digitalt tilgjengeleg for redaksjonen med tanke på kryssreferering, kjeldeføring mm. Dette var motivasjonen bak første innlegging, i 2005.

Eit par år synte det seg at Samlaget, forlaget til NO2014, mangla prenta eksemplar av band 2, og spørsmålet som kom opp var om ein skulle laga ei faksimile-utgåve av band 2, eller fullføra innlastinga av band 1-2 i ordboksbasen og generera opp satsen frå basen.

Det høyrdest både fornuftig og forlokkande ut å fininnlasta band 1 og 2 for å få teksten heilt på plass i NO-databasen, så det vart vedteki å prøve å få dette til. C, E, S. Ore og O. Grønvik fekk ansvar for dette tiltaket og arbeidde med det sommar og haust 2006.

Føremålet var at ordartiklane i band 1-2 skulle framstå som adekvate NO-artiklar både i ei prenta og ei digital utgåve av ordboka. Kvar artikkel skulle vera meiningsberande i seg sjølv, og ikkje føresetja at ordboksteksten rundt kunne tena til oppklaring. Dermed var det nødvendig

å flytta opplysningar

å fylla lakunar

å flytta innhaldstome artiklar over i samansetningsrekkjer

å stryka opplysningar som NO2014-formatet ikkje har plass til.

Opprettinga skulle ikkje vera ein innhaldsrevisjon, av di revisjon av redigert tekst i band 1-4 heile tida har vori rekna som eit eige (og etter måten resurskrevjande) separat prosjekt. Det var snakk om grader av innpassing i eit nytt og stringent databaseformat. Det viste seg likevel at detaljinnpassing av eit gammalt dokument i nye lagringsformat vart ei framfor alt tidkrevjande oppgåve,

Førelegg og første strukturanalyse

Band 1-4 vart grovanalyserte av EDD ved Christian-Emil Smith Ore med tanke på innlegging i basen, og var på plass i basen frå juni 2004, minus hakeparentesen. Då var oppslagsdel og artikkelstruktur i definisjonsdelen nokolunde på plass. Det vart gjort ein del manuell feilretting etterpå for å få inn alle tydingsbolkar i artiklar.

Det må strekast under at feilrettinga ikkje galdt feil i analysen av førelegget, men rett og slett at førelegget (den innskrivne teksten av band 1-5) ikkje har same strukturelle konsistens som ordboksmanuskriptet frå og med i-. Christian-Emil Smith Ore laga eit program til analyse av NO-teksten som tok prosessen eit stykke på veg, og mellom anna førte til at oppslagsord vart komplette (fylte ut førsteledd i samansette ord), og til opprydding i ordklassefeltet (reduserte talet på moglege ordklassenemne frå omkring 400 til den noverande lista). Kjelder, tilvisingar og sms-rekkjer hamna i felt for definisjon eller sitat (jf artiklar i band 3 og 4). Hakeparentesen var så variabel og inkonsistent i utforming at han ikkje vart teken med i første omgang. Derfor manglar hakeparentes-opplysningar i basen for band 3 og 4.

Gevinsten ved å ha band 1-4 tilgjengelege i basen er openberr. Men artiklane på plass er det til dømes mogleg å kryssreferera på tydingsnivå på tvers av heile verket, noko som ville vori umogleg om band 1-4 ikkje låg i basen.

Minus er at kjeldesystemet ikkje var skilt ut og tilgjengeleg i seg sjølv, tilvisingar og samansetningsrekkjer låg som sitat og var ukopla, hakeparentesen mangla, og det låg att ein del strukturknutar i manus (typen "UANALYSERBART DØME").

Finanalyse band 1-2

Desse manglane, pluss ein akutt mangel på prenta eksemplar av band 2, var bakgrunnen for at det vart sett i gang ein finanalyse av band 1-2 med tanke på feilfri innlasting i basen, noko som i sin tur skulle gjera det mogleg å generera satsen for band 1 og 2 direkte frå ordboksbasen.

Arbeidet gjekk føre seg i 2006 – 2007, ved Christian-Emil Smith Ore og Oddrun Grønvik. Føremålet var å få alle opplysningar inn på rett plass i ordboksbasen.

Sidan NO a-h ikkje innfridde konsistenskrava som gjeld for NO i-å, måtte det gjerast endringar i bandteksten for å få til ei fininnpassing.

Utgangspunktet for innpassinga var gjeldande konvensjonar for NO 2014. Dette innebar å tilpassa den tidlegare teksten til det NO-formatet som gjeld no, og som vart utarbeidd i hovudsak i 2002 - 2003.

Redaksjonelle premisser og konsekvensar

Ein premiss for innpassingsarbeidet (sett av dåverande prosjektdirektør Kristin Bakken etter tett konsultasjon med Laurits Killingbergtrø, som då var den eldste i redaksjonen og ein av hovudredaktørane, og med Oddrun Grønvik) var at alle opplysningar som fanst i

NO band 1-2 i utgangspunktet skulle takast vare på i redigeringdatabasen, med mindre det galdt opplysningar som eksplisitt skulle sløyfast i NO 2014 (sjå nedanfor).

Konsekvensar av dette var at artikkeltypar måtte skiljast (normal, føreledd, etterledd, tilvising; strengt alfabetisk ordning av artiklar i sms-rekkjer) ufullstendige normalartiklar måtte gjerast fullstendige (mange artiklar mangla definisjon). innførslar i sms-rekkjer med tilleggsopplysningar måtte takast ut av rekka og gjerast om til fullstendige artiklar

Ein vidare premiss var at rekkjefølgja av opplysningane i gamle NO skulle stå som før, så langt som råd var.

Konsekvensar av dette var at lange tydingsbolkar med repetisjon av serien [definisjon – sitat – sms-rekkje] måtte delast opp i fleire nummererte bolkar, sidan sms-rekkjer alltid kjem sist i tydingsbolken rekkjer av typen [ordsamband – forklaring – døme (somme gonger med fleire forklarings med døme til)] måtte skiljast ut i underoppslag. Om ordsamband med handsaming som underoppslag i NO2014 kom midt i ein tydingsbolke måtte tydingsbolken anten delast på det punktet, eller ein må godta at ordsambandet hamnar sist i bolken (framfor eventuelle sms-rekkjer).

Vidare: I NO 2014 kan ein artikkel berre ha eitt underoppslag pr ordsamband med ei viss form (t. d. ”føra til”). Dersom same ordsamband var omtala fleire stader i ein artikkel, ville alle omtalar bli samla og nummererte som ulike tydingar under same underoppslaget. I slike tilfelle vart ordsambanda ofte samla sist i artikkelen.

Fininnpassing - prosess i tre fasar

Arbeidet med fininnpassing av band 1 og 2 gjekk i tre omgangar.

Fase 1 var oppretting av filene for band 1 og 2 slik at det skulle vera mogleg å losa dei på plass i detalj i ordboksdatabasen. Christian-Emil Smith Ore stod for analysen og Oddrun Grønvik stod for den manuelle opprettinga av band 1-2 i 2006-2007.

Fyrst vart strukturen finanalysert av Christian-Emil Smith Ore, og send i konkordansform til Oddrun Grønvik med fokus på eit eller anna punkt, t.d. kjeldesignaturar. Oddrun Grønvik plukka ut feil og retta rett i bandfila. På denne måten vart alle kjelder isolerte og registrerte (ein del litterære kjelder mangla i bibliografibasen, ein del heimfestingar var ukorrekte i høve til målførehierarkiet), og krysstilvisingar, underoppslag og samansetningsrekkjer vart isolerte og fekk rett form og plassering.

Fase 2 var tilpassing av hakeparentesen i band 1-2 til eksisterande format. I ordboksdatabasen er hakeparentes-informasjonen kategorisert og ordna etter eit fast system med 5 hovudkategoriar som har fast rekkjefylgd. I band 1-2 var det frå fyrstinga

ein god del eksperimentering med presentasjonen av dette tilfanget, og kategoriseringa er mindre tydeleg enn det som er skikken i NO2014.

Proessen i fase 1 vart repetert, dvs at Lars Jørgen Tvedt stod for analysen og prøvekyringane og Oddrun Grønvik stod igjen for opprettinga. Lars Jørgen Tvedt gjorde deretter innpassinga i ordboksbasen. Samstundes vart alle koplingar gjorde, dvs kopling til Metaordboka, til bibliografibasen, og interne koplingar i ordboksbasen (implisitte og eksplisitte tilvisingar).

Etter fininnlastinga var det framleis stor skilnad mellom band 1-2 og NO2014 frå bokstaven i og utetter.

Fase 3 vart sett i gang vinter-vår 2007-2008 for å jamna ut dei skilnadene noko. Oppdraget var ”korrektur med minimale justeringar”, det vil seia å retta synlege feil etter gjeldande redigeringsreglar i dei oppgenererte artiklane i band 1 og 2. Oppdraget omfatta ikkje å fylla inn manglar i artiklane, eller vurdera artikkelinnhald i høve til grunnlagstilfanget slik det låg føre i 2007 – det ville blitt full revisjon.

Fire etter måten røynde redaktørar, Ålov Runde, Terje Svardal, Dagfinn Rødningen og Anne Engøe fekk kvar sin bolk av band 1-2 til oppretting, og Oddrun Grønvik tok ein bolk. Vi skulle arbeida med dette ein dag i veka i ½ år (20 veker).

Det synte seg snart at dette var ei svært lite tilfredsstillande arbeidsform for redaktørane, som har kvalitetskrava for NO 2014 i blodet og fingerspissane, og tykte det var mest uuthaldeleg å fira på krava til materialhandsaming og presentasjon, konsistens og kvalitetskontroll. Det synte seg også at prosessen ”korrektur med rettingar” tok mykje lenger tid enn venta, og greip inn i redigeringsprogresjonen for NO 2014. Tiltaket vart derfor stogga i juni 2008.

Tidsforbruk

I alt var det brukt 2/3 årsverk i fase 3. Med den tidlegare innsatsen til Oddrun Grønvik, Christian-Emil Smith Ore og Lars Jørgen Tvedt er det lagt minst eit årsverk i deldigitalisering av band 1-5 og findigitalisering av band 1-2.

Progresjon

Oddrun Grønvik vart ferdig med minimumskorrektur på 29 spaltar i løpet av fase 3. Om ein reknar arbeidsinnsatsen som 20 dagsverk, eit månadsverk, blir pårekna tidsforbruk per band a $800 \text{ sider} = 800 * 2 / 29 = 55$, dvs 5,5 årsverk (10 arbeidsmånader pr år). I tillegg kjem administrasjon, overordna redigering (kvalitetskontroll), IKT-innsats mm. I alt ville minimumstilpassing vorti minst 8 årsverk pr band.

Det må understrekast at dette ikkje var nokon innhaldsrevisjon, noko som er sterkt påkravt for dei fyrste banda av NO 2014.

Status per dato for band 1-4 og band 5 ut bokstaven h

Band 1 og 2 har artiklar med akribi på plass i basen, irekna entitetar, bruksmerkingar osv. Den kjeldeføringa som er der, er i samsvar med reglane for NO2014, men mykje er ufullstendig, særleg i høve til dei eldre kjeldene. Teksten er mangelfull i den forstand at han ikkje fyller kvalitetskrava for NO2014 (materialkontroll, konsistens, dimensjonering etter underliggjande tilfang), og at han ikkje er innhaldsrevidert sidan 60-70-talet.

Band 3-4 har artiklane på plass i basen, men akribi manglar eller er feilplassert (kjeldefestingar og heimfestingar ligg inne i felt for sitat og definisjon, hakeparentes manglar heilt).

Første del av band 5 – resten av bokstaven h – ligg føre som digitalisert tekst, men er aldri freista innlasta i basen.

Det må leggjast til at NO band 1-5 er skeivproporsjonerte. Band 1-2 har høg alfabetprogresjon og eit etter måten enkelt artikkelformat. Alfabetprogresjonen fall sterkt i band 3 og heldt seg låg i band 4 og ut bokstaven h i band 5. Samstundes voks kompleksiteten i artikkelstruktur og kjeldeføring. Det er derfor eit ope spørsmål om det er verdt det å freista å lasta inn første del av band 5 i basen, gjevi at teksten i alle fall må gjennom ein omfattande revisjon. Det kan henda at det mest tenlege er å la redaktørane som står for revisjonen arbeida utan hinder i basen, men med eit blick på den digitaliserte teksten til første del av band 5.

Oddrun Grønvik

Oslo 22.3.2012

Notat om leksikografi ved ILN etter 2015

Christian-Emil Ore, 12.08.2011

1 Om notatet

Dette notatet er ment som en analyse av en del av forutsetningene for et stort vitenskapelig/forskningsbasert leksikografisk dokumentasjonsprosjekt for moderne bokmål fra 2015. Basis for arbeidet er de konkrete erfaringene som er gjort gjennom arbeidet med mellomstore og (større) (prosjekt) ved ILN og andre steder.

Det har spesielt blitt undersøkt om det er sannsynlig at den elektroniske ordboken som NAOB-prosjektet skal levere i 2014, kan danne fundamentet for et slikt leksikografisk dokumentasjonsprosjekt. Analysen i dette notatet konkluderer med at den ikke er egnet som et fundament en kan bygge en vitenskapelig beskrivelse av det moderne bokmålet for f.eks. perioden 1950 til i dag. NAOB kan imidlertid brukes som en blant flere kilder til en beskrivelse av bokmålet. Det er flere årsaker til det. Hovedårsakene er at utgangspunktet for NAOB, Riksmålsordboken, er en gammel ordbok og at NAOB har begrensede ressurser og kanskje som en følge av denne ressursknappheten ikke har lagd eller brukt korpus eller annet systematisk kildemateriale¹. Den kan dermed ikke sies å være en vitenskapelig ordbok.

Resultatet av NAOB-prosjektet vil kunne bli en interessant og omfattende elektronisk ordbok. Så selv om NAOB ikke kan brukes som et fundament for en vitenskapelig beskrivelse av bokmål, bør NAOB i 2014 lenkes sømløst mot andre språkressurser for bokmål og også nynorsk, for på den måten å danne en helhetlig inngang til informasjon om norsk språk. Men det er ikke tale om noe forsøk på en samnorsk-ordbok, men bare en felles beskrivelse av norsk gjennom ulike kilder med sitt eget redaksjonsspråk.

Notatet er ikke ment å inneholde noen kritikk av hverken BRO-samarbeidet eller NAOB-prosjektet. BRO-samarbeidet er fornuftig fordi det kan bidra til å blegge en meningsløs språkstrid. NAOB-prosjektet er et privat forlagsprosjekt. Miljøene bak dette prosjektet, *Det norske akademiet for språk og litteratur* og *Kunnskapsforlaget*, står selvfølgelig fritt til å gjennomføre prosjektet slik de måtte finne det best. Notatet inneholder ingen vurdering av disse miljøenes egnethet og faglige kompetanse. Notatet drøfter heller ikke språkpolitiske spørsmål eller forholdet privat kontra offentlig kontroll med beskrivelsen av norsk. Det er viktige temaer, men bør drøftes andre steder.

For ILN er det viktig å sikre at BRO-samarbeidet utformes slik at ILNs leksikografimiljø og Norsk ordbank får best mulig nytte av resultatene. Det er også viktig at det kan brukes til å berede grunnen for et stort vitenskapelig dokumentasjonsprosjekt for bokmål i regi av ILN fra 2015 og utover. Et slikt prosjekt bør organiseres som et laboratorium for leksikografi/leksikologi og knytte til seg master- og phd-studenter. På den måten kan en sikre studie- og forskningsvirksomheten ved samlingsavdelingene ved ILN, jf. Forum for universitetshistorie ved UiO som har hatt en rekke studenter knyttet til seg. Det er derfor svært viktig at bokmålsprosjektet blir lagt til ILN.

¹ Se <http://ordnet.dk/ddo/fakta-om-ddo/metode-og-kilder> for en kort, men god gjennomgang av metoder for ordboksredigering

Notatet inneholder også en gjennomgang av ressursbruken knyttet til en revisjon av bindene 1 – 5 av Norsk Ordbok. Det kan være nyttig for planleggingen etter 2014. Men beskrivelsen er også tatt med for å gi bakgrunn til analysen av NAOB. De første to bindene av Norsk Ordbok er skrevet i den samme leksikografiske tradisjonen som den gamle riksmålsordbokens bind 1 – 4.

2 En beskrivelse av moderne norsk bokmål ved ILN 2015

De langsiktige planene ved ILN er at det i 2015 når NO2014 prosjektet er ferdig, skal settes i gang et prosjekt for å dokumentere moderne bokmål, se ILNs (INL) styrevedtak fra 2002 nevnt nedenfor under BRO-samarbeidet. Et slikt leksikografisk prosjekt er også nevnt i Stortingsmelding 35 (2007-2008) *Mål og mening* der en skriver "[etter 2014] bør det nemnde dokumentasjonsprosjektet for moderne bokmål koma inn med større tyngde". Dokumentasjonsprosjektet det vises til er prosjektskissen som ble presentert av ILN i 2006 – 2007 i KUD.

Det er altså et ønske ved ILN og en velvilje i Kulturdepartementet til å starte et bokmålsprosjekt i 2015. Akkurat hva prosjektet skal gå ut på er ennå ikke konkretisert. Det finnes bare en projektskisse som ble presentert for Kulturdepartementet i 2006-2007. I denne skissen opereres det med et omfang på 100 – 120 årsverk. Disse tallene har blitt brukt i Stortingsmelding 35.

Prosjektskissen omtaler en database eller en datamaskinell lesbar ordbok, altså noe som skal kunne ha språkteknologiske anvendelser. Slike datamaskinlesbare leksikalske databaser er ofte beregnet på å leses av mennesker. Men erfaringene fra Princeton Wordnet og det danske ordnettet viser at det er viktig å inkludere meningsbeskrivelser skrevet i vanlig fritekst beregnet på mennesker.. Det er derfor naturlig å tenke seg at dokumentasjonen av det moderne bokmålet skal være en tradisjonell elektronisk ordbok med både definisjoner lesbare for mennesker og en datamaskinlesbar komponent. Det er ingen motsetninger mellom komponentene i en slik kombiløsning.

Omfanget av en vitenskaplig basert dokumentasjonsordbok for moderne bokmål kan varieres både i tidsperiode den dekker og antallet ord som beskrives. *Den Danske Ordbok* er basert på et korpus fra 1980 – 1993 og har 96.000 artikler. Progresjonen var om lag 1.200 artikler pr årsverk. *NO 2014* har om lag 860 artikler pr årsverk, mens Bokmålsordboka og Nynorskordboka hadde om lag 2.000 artikler pr årsverk. Antallet årsverk er avhengig av kravene til artikkelen. Antall artikler er i prinsippet uavhengig av tidsperiodens lengde.

Anslaget på 100 – 120 årsverk som er brukt i Stortingsmeldingen, er som nevnt, tatt fra projektskissen. I skissen er det ikke noen detaljert grunnivning for tallene. Men slike tall har lett for å fange. Ut fra vurderingen over, kan det innenfor en slik ramme være mulig å lage en noe mer omfattende elektronisk ordbok enn *Den Danske Ordbok* og med språkteknologisk tillegg som for eksempel semantiske relasjoner, grammatisk informasjon og annet.

Det vil neppe være formålstjenelig å bygge en slik ny ordbok direkte på NAOBs elektroniske ordbok slik den vil være i 2015. NAOB er ikke basert på noe korpus eller andre systematiske kilder og vil selv etter revisjonsarbeidet ha mest fokus på eldre litterært norsk (før 1960). Men en bør sørge for å lenke NAOB inn i et metaordboksystem slik det finnes for Norsk Ordbok, og sørge for en portalløsning der en sømløst kan søke i både den nye bokmålsbeskrivelsen og i NAOB. Det vil hindre separate bokmålsbeskrivelser eller "søsterkirker" som Petter Henriksen kalte det. En slik portalløsning må også lenkes mot Norsk Ordbok-systemet og håndordbøkene, slik at en får en samlet

tilgang til en full beskrivelse av norsk i alle varianter. En forutsetning for å få dette til må være at NAOB blir en gratis og åpen ressurs slik de andre ressursene er det.

3 Ressurssitasjonen ved ILN

I dag (2011) har ILN et blomstrende leksikografisk miljø på over 30 heltidsansatte leksikografer og egen FoU-enhet for elektronisk leksikografi (EDD). Men de fleste ansatte er prosjektengasjerte til ut 2014. Den faste vitenskapelige staben teller i dag 5 personer. I 2015 vil det i verste fall bare være to vitenskapelig ansatte hvorav en leksikograf på 67 år og en humanistisk informatiker på 58. I tillegg vil det være to ingeniører. Så dersom det hverken kommer et bokmålsprosjekt, eller et revisjonsprosjekt for den første delen av Norsk Ordbok, vil det neppe finnes noe leksikografisk miljø ved ILN etter 2015.

Det er åpenbart at et stort bokmålsprosjekt må bemannes på annen måte enn ved den faste staben. Det er naturlig å tenke seg at deler av den nåværende staben ved NO2014 kan gå over i et bokmålsprosjekt. I staben finnes det flere med doktorgrad og også professorkompetanse. Bokmålsprosjektet vil dermed finne godt kvalifiserte redaktører, og UiO vil kunne løse utfordringen ved at 28 personer har opparbeidet seg utvidede rettigheter etter tjenestemannsloven til videre ansettelse ved UiO. Bokmålsprosjektet kan også dra nytte av FoU-enheten (EDD) og alle verktøyene som er utviklet for Norsk ordbok. Det ser ut til at Kulturdepartementet tenker på samme måten. I Stortingmelding 35 står det:

Departementet legg til grunn at eit slikt prosjekt i regi av Universitetet i Oslo vil ha stor nytte av kompetanse og ressursar som er bygde opp i samband med nynorskprosjektet Norsk Ordbok 2014.

4 Om ressursbruk ved ordboksredigering

Kartotek kort eller sedler med opplysninger om bruk av ord ble introdusert på slutten av 1800-tallet. Alle større ordboksredaksjoner bygde opp til dels svært omfattende seddelsamlinger. Norsk ordbok har 3,2 millioner, mens Oxford English Dictionary har flere titalls millioner. Sedlene brukes i redigeringen av ordartiklene. I dag er et godt og dekkende korpus overtatt for seddelsamlingene og ansees for å være en betingelse for å kunne redigere en ordbok av høy kvalitet. Seddelsamlinger gir informasjon om den partikulære, mens korpus om den generelle ordbruken. Når en arbeider diakront eller dialektoriente ordbøker vil det ofte være nødvendig med både et korpus og en stor seddelsamling siden de kompletterer hverandre. NO2014 bruker derfor begge deler, og prosjektet har et korpus på 100 millioner ord. NAOB har ikke opplyst om noe eget korpus, men har muligens en mindre seddelsamling. Tor Guttu, hovedredaktøren for NAOB, fortalte i sitt foredrag på sommerseminaret 17. juni at han noterte interessante ord han fant under lesning og fikk produsert elektroniske sedler. Se for øvrig <http://ordnet.dk/ddo/fakta-om-ddo/metode-og-kilder> for en kort, men god gjennomgang av metoder for ordboksredigering.

4.1 Norske håndordbøker

Nynorskordboka 1974 – 1985, 40 årsverk
Bokmålsordboka 1974 – 1985, ca. 30 årsverk

De to ordboksredaksjonene redigerte hver sin halvpart av alfabetet og byttet deretter manuskripter. Det ble dermed en klar rasjonaliseringsgevinst. Ressursforbruket var omlag 2200 artikler pr årsverk. I

dette overslaget er det ikke iberegnet årsverk til administrasjon. Det reelle tallet er derfor noe lavere. Forskjellen i antall årsverk skyldes at Nynorskordboka hadde ca 90.000 oppslag og Bokmålsordboka 65.000. Ingen av ordbøkene er korpusbasert, men de gir seg heller ikke ut for å gi en vitenskapelig beskrivelse av norsk.

4.2 Mellomstore ordbøker basert på korpus

Den danske ordbok 1994 – 2004, 80 årsverk + 10 årsverk korpusoppbygging

Den danske ordbog (DDO) har om lag 96.000 oppslagsord og har flere betydningsbeskrivelser (131.000) enn Nynorskordboka (112.000), og den har vesentlig flere brukseksempler med kildehenvisning. Den har i tillegg en mye luftigere sats og består av 6 trykte bind, i tillegg til en nettsversjon. I de originale planene var det forutsatt at en kunne gjenbruke materiale fra 22-bindsverket *Ordbog over det danske sprog* (ODS 1918-1956) gjennom tobindsversjonen *Nudansk Ordbog* (1956- >), men det viste seg lite hensiktsmessig, så DDO ble redigert på nytt fra bunnen av. Dette er det verd å merke seg for fremtidige prosjekter. Om en holder arbeidet med korpuset utenfor, så blir progresjonen ved DDO om lag 1.200 artikler pr årsverk.

4.3 Norsk ordbok, nyredigering i prosjektet NO2014

Prosjektet Norsk Ordbok 2014 ble startet i 2002. Prosjektets mål er å redigere og gi ut de 8 resterende bindene av 12-bindsverket *Norsk ordbok* til 2014. Hvert av disse bindene har omlag 30.000 artikler. Hele verket vil ha rundt 340.000 ordartikler. Det vil være det største ordboksverket utgitt i Norge noen sinne. Riksmålsordboken 1 – 4 og de to supplementsbindene har til sammenligning 220.000. Den planlagte NAOB skal ha 260.000 ordartikler.

Det er viktig å understreke at det er en vesentlig forskjell mellom de to ordboksverkene. Norsk Ordbok er en vitenskapelig dokumentasjonsordbok der ordartiklene er knyttet elektronisk og referansemessig opp mot grunnlagsmaterialet, og videre gir denne ordboken en deskriptiv språkbeskrivelse.

I NO2014-prosjektet er det bygd opp en effektiv prosjektorganisasjon med høy språkvitenskapelig kompetanse. Redigeringsprosessen er gjennomført digital. Redigeringsystemet er et velutviklet, rasjonelt og ikke minst velprøvd system. Etter omleggingen går det med 35 årsverk pr bind (1,2 årsverk pr 1000 artikkel) mot over 50 årsverk pr bind før NO2014-prosjektet tok til.

Redigeringsarbeidet, materialet fra Dokumentasjonsprosjektet og digitaliseringen etter 1998 har gjort at ILN har landets mest omfattende dokumentasjon av norsk språk. I 2014 vil det ha blitt brukt over 350 millioner kroner inkludert Dokumentasjonsprosjektets nynorskdel.

Det er også verd å merke seg at Kulturdepartementet i Stortingsmelding 35 skriver:

Det er likevel grunn til å streke 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 språkav som direkte vedkjem ein mykje større del av norske språkbrukarar enn dei som nyttar nynorsk som skriftmål.

Det gir uttrykk for en erkjennelse at det norske språket er et kontinuum og ikke to adskilte språk selv om de to skriftnormene representerer hvert sitt skriftspråk. Det er åpenbart slik, spesielt for moderne norsk, at ordtilfanget for bokmål og nynorsk har en svært stor andel av felles ord. Spesielt vil nærmest alle nyord være felles. Vitenskapelige dokumentasjonsbøker for bokmål og nynorsk vil dermed for en stor grad overlape hverandre, i den forstand at en stor del av artikkelinventaret vil være felles. En beskrivelse av bokmål bør derfor kunne trekke store vekslers på arbeidet som er gjort i NO2014-prosjektet. Men det er altså ikke tale om noe forsøk på en samnorsk-ordbok, men bare en felles beskrivelse av norsk.

4.4 Norsk ordbok, revisjon av de eldre bindene fra før NO2014-prosjektet

I diskusjonen rundt BRO-prosjektet, NAOB og hva en skal planlegge for etter 2014, vil det være nyttig å studere erfaringer fra et pilotprosjekt for revisjon av bindene 1–5. En revisjon av bindene 1 – 5 omfatter både det å tilpasse gammel ordbokstekst til nytt og stringent format og en full innholdsrevisjon. Ut fra pilotstudier er den første prosessen anslått til minimum 7 årsverk pr bind, eller 3 500 artikler pr årsverk. En full innholdsrevisjon med 10% nye artikler for å bringe de eldste bindene opp til samme nivå som de nye, vil kreve minimum 15 årsverk pr bind, eller 1800 ordartikler pr årsverk. Tallene passer godt med andre kjente ordboksverk og er nyttige fordi de gir en pekepinn om hva slags resultat en kan vente ut fra en gitt ressurstilgang.

4.4.1 Innpassing av gammel tekst i et nytt og stringent format

Det elektroniske redigeringsystemet ble tatt i bruk fra og med bokstaven *i*, som begynner 2/3 ute i bind 5. Bindene 1 og 2 ble redigert manuelt (skrivemaskin og tradisjonell sats). De ble skrevet inn elektronisk som tekstbehandlingsfiler rundt år 2000. Bindene 3, 4 og 2/3 av bind 5 ble som en del av I av redigeringsarbeidet skrevet inn i henhold til et enkelt felt- og kodesystem utviklet på 1980-tallet. Bind 3 og 4 ble satt elektronisk.

NO2014 og EDD har maskinelt analysert den elektroniske teksten (a-h) og gitt den en XML-koding. Resultatet ble lastet inn i redigeringsapplikasjonen for å gjøre det søkbart tilgjengelig for redaksjonen og muliggjøre kryssreferanser fra de nyredigerte delene (i-å).

Etter 2005 ble det et behov for å trykke opp igjen bind 2. Alternativene var å lage et fotografisk opptrykk eller å bruke settefunksjonen i det nye redigeringsystemet. En gikk derfor i gang med et mindre prosjekt for å rydde i og korrekturlese den innlastete teksten nevnt over.

”Føremålet var at ordartiklane i band 1-2 skulle framstå som adekvate NO-artiklar både i ei prenta og ei digital utgåve av ordboka. Kvar artikkel skulle vera meiningsberande i seg sjølv, og ikkje føresetja at ordboksteksten rundt kunne tena til å oppklaring. Dermed var det nødvendig å flytta opplysningar, å fylla lakunar, å stryka opplysningar som NO2014-formatet ikkje har plass til. Opprettinga skulle ikkje vera ein innholdsrevisjon av di revisjon av band 1-4 heile tida har vori rekna som eit eige (og etter måten resurskrevjande) separat prosjekt.”
[O.Grønvik: Digitalisering og tilrettelegging, Norsk Ordbok Band 1-5, 14.7.2011]

Fire erfarne redaktører ble satt på oppgaven. Det ble i alt brukt 2/3 årsverk. Ut fra redaktørens progresjon ville det gå med om lag 5,5 årsverk pr bind (25.000 artikler), bare til ren opprydning. Det viste seg at redaktørene opplevde det komplisert og frustrerende å ikke kunne rette innholdsmessig i

teksten. Men oppgaven skulle bare være innpassing av gammel tekst i et nytt og stringent format. Brutto ressursforbruk for hvert bind med tillegg av administrasjon, sluttkorrektur og IKT ble beregnet til om lag 7 – 8 årsverk pr bind, eller 32 – 37 årsverk totalt. Det gir mellom 2.800 og 3.100 artikler pr årsverk.

4.4.2 En full innholdsrevisjon av de første 5 bindene.

Første hefte av Norsk ordbok kom i 1950. Bindene 1 og 2 var ferdige i henholdsvis 1966 og 1978. Bindene 3 og 4 kom ut i 1994 og 2002. Bind 5 kom i 2005. For å oppnå en enhetlig språkbeskrivelse etter vår tids metoder er det åpenbart nødvendig med en grundig innholdsrevisjon av de første bindene. Det må gjøres mot det kildetilfanget som nå finnes, spesielt korpus. Som et ledd i dette arbeidet må innholdet i de eldre bindene tilpasses det stringente formatet som bindene 6 – 12 følger. Dette arbeidet kan enten gjøres ved å føre det opprinnelige manuskriptet inn i databaseformatet slik det er beskrevet i forrige avsnitt, og deretter oppgradere artiklene i henhold til det nå foreliggende materialet. Alternativt kan en gjøre begge steg artikkel for artikkel. Det vil erfaringsmessig være mer effektivt og redusere totalkostnadene. Her kreves det flere pilotstudier. Ressursbehovet vil nødvendigvis være større enn for tilpassingen beskrevet over (8 årsverk pr bind), men vesentlig mindre enn for dagens nyredigering (35 årsverk pr bind). Ut fra den generelle erfaringen med ordboksredigering anslår redaktørene i Norsk Ordbok revisjonsarbeidet til å kreve minimum 15 årsverk pr bind. For de 120.000 ordartiklene i bindene og et antatt tillegg på rundt ti prosent nyord vil det altså gå med om lag 70 årsverk. Progresjon blir da noe over 1700 artikler pr årsverk..

5 BRO-samarbeidet

I de siste 20 årene har det vært tatt minst to initiativer til samarbeid (1995 og 2005) mellom leksikografimiljøet ved UiO og *Riksmålsvernet/Det norske akademiet for språk og litteratur (NA)*. Men hovedintrykket når en ser tilbake, er uenighet og gjensidig manglende tillit. Det er det mange (språkpolitiske) grunner til. Striden toppet seg i 2002 da ILNs styre vedtok

” Styreren får i oppdrag å utforme et skriv til KKD der vi går imot at det bevilges midler til å etablere en stor ordbok ved en privat institusjon, utenfor fagmiljøene ved universitetene. Skrivet formes slik at vi ikke blokkerer for at et stort ordboksprosjekt i forlengelsen av NRO legges til UiO. Styreren konsulterer Ruth Vatvedt Fjeld og seksjonsleder ved SLM ved utarbeidelsen av dette skrivet.”

Vedtaket førte til en svært forsuret stemning og det ble publisert en artikkel i Aftenposten i klassisk språkstridsjanger.² Striden ble først bilagt ved opprettelsen av Bokmål-Riksmålsordbok (BRO) samarbeidet i 2009 gjennom et brev til Kulturdepartementet undertegnet av ILN, NA og Kunnskapsforlaget.

I stortingsmelding 35 står det (side 168):

Utgangspunktet vil difor vera at departementet i første omgang konsentrerer innsatsen på bokmålssida om det prosjektet som no har nemninga Det Norske Akademis store ordbok. Deretter bør det nemnde dokumentasjonsprosjektet [planene om en Norsk bokmålsordbok

² http://www.aftenposten.no/kul_und/article460336.ece

estimert til 100 – 120 årsverk, presentert for KD i 2006, seside 166 i St.meld 35)) for moderne bokmål koma inn med større tyngde. Departementet legg til grunn at eit slikt prosjekt i regi av Universitetet i Oslo vil ha stor nytte av kompetanse og ressursar som er bygde opp i samband med nynorskprosjektet Norsk Ordbok 2014.

På desse premissane har departementet for 2008 funne grunnlag for å auka tilskotet til Det Norske Akademi's store ordbok med i alt 1,5 mill. kroner, nær på ei fordobling av det låg inne i det opphavlege budsjettet for 2008. Målet er at prosjektet skal vera fullført i 2014, og det er viktig at rammevilkåra gjer det mogleg å gjennomføra arbeidet på ein fagleg tilfredsstillande måte. Departementet er difor innstilt på å vurdera tilskotsbehovet på nytt ved komande budsjettbehandlingar.

BRO-samarbeidets hovedhensikt ser ut til å være å hindre kiv mellom ILNs leksikografer som arbeider med bokmål og grupperingene rundt NAOB, dvs Kunnskapsforlaget, NA og Riksmålsforbundet. I følge Petter Henriksens presentasjon på Norsk Ordbok 2014s sommerseminar 17. juni ble det på møtet i Kulturdepartementet (Bente Christensen, Ruth Fjeld, Petter Henriksen og Helge Dyvik og representanter for embetsverket) ble det fra Kulturdepartementets side sagt at "Dersom partene inngår en avtale om teigdeling, er Kulturdepartementet altså villig til å finansiere arbeidet for å dokumentere norsk bokmål i ulike varianter", jamfør avsnittene fra Stortingsmelding 35. Hva som egentlig skal finansieres er noe uklart. Men i følge det som ble presentert av NAOB 17. juni ser det ut til å være snakk om to faser, fase 1 for perioden 2012 -2014 og fase 2 fra 2015.

I fase 1 skal følgende finansieres:

- en revidert utgave av NAOB publisert som nettordbok i løpet av 2014. Revisjonen ser ut til å skulle bestå i en omlegging av redaksjonsspråket til moderne, men konservativt bokmål, en utflating av artiklenes betydningsstruktur, en fullføring av den påbegynte revisjonen av Riksmålsordbokens artikler og en nyredigering for ord i norsk etter Norengs to supplementbind. Dette arbeidet skal utføres i sin helhet av en styrket NAOB-stab. Denne staben er organisatorisk en del av Kunnskapsforlaget (for tiden om lag 50% av staben der).
- en kollokasjonsordbok (KOB) basert på Leksikografisk bokmålskorpus utvidet til hundre millioner ord. Ordboken skal være elektronisk og rettet inn mot språkteknologiske anvendelser. Prosjektet vil i hovedsak være en kollokasjonsfokuseret analyse av korpuset. På seminaret ble det sagt at denne databasen skulle gjøres i samarbeid med NAOB-prosjektet, slik at det prosjektet kunne nyttiggjøre seg resultatene i sin revisjon og utarbeidelse av den elektroniske NAOB. KOB-prosjektet skal få tilgang til NAOBs manuskriptdatabase for å kunne analysere materialet for kollokasjoner og sammenligne mot korpus.

Fase 2 av BRO-samarbeidet skal strekke seg i et foreløpig ubestemt antall år fra og med 2015. Det er uklart hvem som skal delta i prosjektet. I følge Petter Henriksen (17. juni) drøftet en på møtet i Kulturdepartementet også bruken av NAOB i et bokmålsprosjekt post 2014. Det ble sagt at for å muliggjøre dette, må NAOB ha nok midler nå. Da vil KUD kunne sette i gang dokumentasjon for bokmål etter 2014. Petter Henriksen ga et inntrykk av at NAOB-miljøet skulle ha en ledende eller vesentlig rolle i dette arbeidet.

Dersom dette er innholdet av BRO-samarbeidet frem i tid, er det grunn til bekymring. Det er riktignok svært positivt at fase 1 skal sørge for en finansiering av en kollokasjonsordbok for bokmål. Men om en ser nøyer på fase 1, ser en raskt at det ikke er snakk om noen gjensidig nytte. Brukseksemplene og definisjonene i NAOB er av liten vitenskapelig verdi som et grunnlag for et studium av kollokasjoner i moderne norsk (jf Ruth Fjelds foredrag 17. juni 2011). Så samarbeidet vil her være at ILN vasker og forbedrer NAOB. Når det gjelder fase 2 så må en gå litt nøyer inn på hva NAOB er.

6 NAOB

NAOB er basert på NRO 1-4 redigert i perioden 1937 – 1957 og de to supplementsbindene til Noreng utgitt 1995. NRO er en solid ordbok av den tradisjonelle litterært orienterte typen som ble etablert av blant andre brødrene Grimm, OED og SAOB. NRO kan likevel ikke på noen måte sammenlignes med SAOB eller ODS. NRO er en adskillig mindre omfattende ordbok enn den det litterære ordboksverket som var planlagt på 1920-tallet, men som aldri har blitt realisert. Redigeringen av NRO er for en stor del basert på seddelmaterialet til Norsk Litterært Ordboksverk (del av Leksikografisk institutt og nå ILN, UiO). Det gjelder også supplementsbindene som kom i 1995. NRO hadde ikke i noen særlig grad noe egen kildesamling. Den er i så måte ingen vitenskapelig ordbok.

I annen halvpart vokste det frem en ide om å støpe sammen NRO og de to supplementsbindene til ett verk. Riksmålforbundet fikk tilbud om å delta i en felles dokumentasjonsdatabase for norsk i regi av Dokumentasjonsprosjektet (DP), men avsto fordi de mente at "bokmålsprofessorene" ville tvinge dem til å endre redaksjonsspråk til moderne bokmål (se også Langslet 1999) dersom de gikk inn på det. Et tilsvarende tilbud fra DPs etterfølger EDD ble det heller ikke noe av. Grunnen var at NA og KF så NRO som et kommersielt levedyktig prosjekt og ikke ville gi fri tilgang til materialet.

KF/NASL søkte og begynte å få en relativt beskjeden støtte fra Kulturdepartementet fra 2002 for å lage et felles ordboksmanuskript av NRO og supplementsbindene. I tillegg ønsket en å revidere ordboken til å omfatte mer moderne språk og publisere den som en trykt ordbok i 2014. Siden den gang har en erkjent at det ikke er mulig å få det til, og at en i stedet skal satse på en elektronisk ordbok. Det seneste vedtaket er at redaksjonsspråket skal endres til moderne konservativt bokmål og at det skal foretas en forenkling i artikkelstrukturen. Ut over dette er de konkrete redigeringsplanene og redigeringsreglene ukjente for allmennheten. Det er heller ikke kjent hvilket kildemateriale NAOB skal baseres på. Men dersom ordboken skal få en rimelig god kvalitet, må det foretas en fullstendig revisjon av manuskriptet (databasen). En styrke ved NRO er at den dokumenterer de "(ut)dannede klassers" verden i deres eget tidsriktige språk og tankeverden. Den er således et godt og viktig kulturmonument. Å endre redaksjonsspråket til moderne moderat bokmål vil muligens skape flere problemer enn det løser. Det vil oppstå en redaksjonell og metaspråklig distanse til det eksisterende NRO-materialet, og det vil føre med seg store endring både i lemmevalg og ordbeskrivelser.

NAOB-prosjektet er et ambisiøst prosjekt. Prosjektet presenteres som en (full) revisjon og modernisering av den gamle Riksmålsordboken fra 1930-1957 og supplementsbindene fra 1990-tallet. Ressursene står ikke i samsvar med ambisjonene. I henhold til analysen gjort i dette notatet, vil bare det å bringe Riksmålsordboken med supplementsbind inn i et stringent digitalt format og å revidere artiklenes struktur (jf analysen gjort for Norsk Ordbok) kreve mellom 60 og 70 årsverk. I tillegg kommer altså en gjennomgående revisjon av redaksjonsspråket, en tematisk innholdsrevisjon,

en omarbeidelse av artikkelstrukturen (utflating), en justering av oppslagsordene og en nyredigering av 40 000 artikler.

Med den nåværende bevilgningstakten vil prosjektet ha fått om lag 40 millioner kroner i 2014 (P. Henriksen 17. juni 2011). Bruttokostnadene til et redaktørårsverk er mellom 700.000 og en million kroner. Bevilgningene NAOB-prosjektet har fått til og med til 2014 er derfor langt fra tilstrekkelige til å gjennomføre de ambisiøse planene på en god måte. Staben må styrkes vesentlig for å få en revisjon på et jevnt nivå. Men de tre årene som gjenstår frem til utgangen av 2014 vil med all sannsynlighet være for knapp tid selv med en vesentlig økning av bevilgningene.

NAOB vil i 2014 altså være et delvis revidert verk og ikke åpenbart et godt valg som grunnlag for en videre vitenskapelig beskrivelse av norsk. Med vitenskapelig menes her et verk med en arbeidsform og en tilhørende elektronisk verktøykasse som sikrer etterprøvbar redigering, og der en faktisk kan dokumentere de kildene som ligger til grunn for ordutvalg og betydningsbeskrivelser, og der det finnes en konsistent redaksjonell praksis gjennom hele verket. Om en faktisk skulle ønske å gjøre NAOB om til en vitenskapelig ordbok, må en bygge opp et korpus for perioden ordboken dekker. Alle artiklene, også de som allerede vil ha blitt revidert frem til 2014, må gås gjennom på nytt og oppdateres i henhold til informasjonen i korpuset og eventuelt andre systematiske kilder.

Det å utvikle NAOB til en vitenskapelig beskrivelse av norsk (bokmål) vil altså kreve en fullstendig innholdsrevisjon. Som pilotprosjektet for Norsk Ordbok 1 – 5 viser, vil det en svært omfattende oppgave. Den vil være omtrent samme omfang som å redigere ordboken fra grunnen av, altså i størrelsesorden 200 årsverk.

Det er derfor svært mye som taler for at NAOB ikke er egnet som et fundament en kan bygge en vitenskapelig beskrivelse av det moderne bokmålet for f.eks. perioden 1950 til i dag. En bør heller satse på å skrive en ordbok fra grunnen av basert på eksisterende og velprøvde verktøy utviklet for NO2014, på eksisterende og nytt elektronisk tilgjengelig grunnlagsmateriale og den eksisterende dokumentasjonen i Norsk Ordbok-systemet.

Samarbeidsavtale for prosjektet Bokmålets og riksmålets ordbase (BRO)

Bakgrunn, grunnlag

1. Dette er en avtale som stadfester BRO-samarbeidet, etablert gjennom brev av 11/9-09 til kulturminister Trond Giske, og modifisert gjennom drøfting med Kulturdepartementet (KUD).
2. Partene i BRO-samarbeidet og i denne avtalen er *Det Norske Akademi for Sprog og Litteratur* (Akademiet) og *Institutt for lingvistiske og nordiske studier* (ILN) ved Universitetet i Oslo.
3. Formålet med BRO-samarbeidet er å koordinere ressursbruken og kompetansen innenfor bokmålsleksikografien gjennom å utvikle ett felles nasjonalordboksverk for bokmål/riksmål (kalt *BRO-ordboksverket*) i samsvar med KUDs ønsker. Med *nasjonalordboksverk* menes et ordboksverk som gir en vitenskapelig, korpusbasert ordboksdokumentasjon av målformen synkront og diakront, tilpasset brede brukergrupper (brukerorientert og funksjonelt), og med det omfang som tilkommer et nasjonalt referanseverk.
4. Ved avtalens inngåelse har Norge ett stort ordboksverk for bokmål/riksmål, som ikke oppfyller dagens krav til nasjonalordboksverk, *Norsk Riksmålsordbok* (NRO), skapt av Akademi-miljøet i tidsrommet 1937–95.
5. Ved avtalens inngåelse er en modernisering av NRO under arbeid, hovedsakelig for offentlige midler, under navnet *Det Norske Akademis store ordbok* (NAOB). NAOB har moderat bokmål som redigeringsspråk og norm for seleksjon av artikkelbærende lemmaer. NAOB vil oppfylle nasjonalordboksverkets krav til brukerorientering og funksjonalitet, men ikke til vitenskapelighet og omfang.
6. Ved avtalens inngåelse stiller ILN følgende til rådighet for BRO-samarbeidet: *Norsk ordbank* med en omfattende lemmaliste over norsk bokmål med hele formverket for den offisielle bokmålsnormen; definisjoner og betydningsbeskrivelser i *Bokmålsordboka*; det vektede *Leksikografisk bokmålskorpus* (LBK); tilhørende analyseprogrammer.
7. BRO-samarbeidets hensikt er at NAOB blir publisert i modernisert versjon innen utgangen av 2014, og deretter brukt som en konstituerende del i BRO-ordboksverket, sammen med den omfattende korpusbasering og vitenskapelige supplering som deretter må til for at kravene til nasjonalordboksverk skal oppfylles.

Spesifisering av innhold i BRO-ordboksverket

8. BRO-ordboksverket skal ha form som en digital ordbok.
9. BRO-ordboksverket skal gi tilgang til NAOB samt bl.a. a) formverk fra ILNs Norsk ordbank; b) kollokasjoner fra en vitenskapelig kollokasjonsdatabase utarbeidet ved ILN; c) eksempelmateriale fra LBK; d) alle former og bøyninger i bokmål/riksmål.
10. BRO-ordboksverket skal tilby publikum et grensesnitt i form av én artikkelrekke (lemmarekke) der NAOBs artikkelrekke inngår, der tilhørende innhold er innlenket/innarbeidet på relevant plass i artikkelen slik at ordboksverket fremstår som én integrert ordbok for brukerne.

Rettigheter

11. Alt innhold som skapes som ledd i BRO-samarbeidet, stilles til disposisjon for BRO-ordboksverket.
12. Partene beholder sin opphavsrett og krediteringsrett til det innhold de bidrar med i i BRO-samarbeidet.
13. BRO-ordboksverkets databasesystem skal sørge for at innholdets opphavsperson lar seg identifisere.
14. Partene har rett til å publisere eget vitenskapelig innhold separat i den grad det ikke underminerer eller dublerer BRO-ordboksverket.

Styring

15. Arbeidet med BRO-ordboksverket styres gjennom en styringsgruppe satt balansert sammen av representanter for Akademiet og ILN. Forslag til sammensetning: Akademiets preses, instituttleder for ILN, en faglig representant for ILN, en faglig representant for Akademiet, en nøytral styringsgruppeleder (f.eks. Språkrådets direktør).

Varighet

16. BRO-samarbeidet er tuftet på Departementets langsiktige engasjement i å gi bokmål/riksmål en fullgod, vitenskapelig dokumentasjon i ordboksform, og tenkes å bestå på ubestemt tid. Dersom statens finansielle engasjement på et tidspunkt opphører så samarbeidet ikke lenger kan oppfylle sitt formål, kan samarbeidet også opphøre.

Oslo,

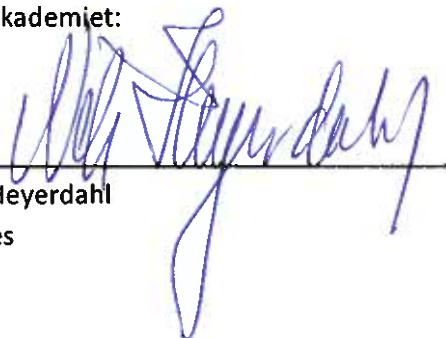
17/8/11

For ILN:



Bente Christensen
Instituttleder

For Akademiet:



Nils Heyerdahl
Preses



Du er her: [Kulturdepartementet](#) < [Dokument](#) < [Proposisjonar og meldingar](#) < [Meldingar til Stortinget](#) < [8 Språkdyrking – eit ...](#) < [8.3 Språkleg dokumentasjon](#)

St.meld. nr. 35 (2007-2008)

Mål og meining

Bla i dokumentet: < [8.2 Språkutvikling og...](#) | [8.4 Språknormering](#) > | [Framsida til dokumentet](#)

8.3 Språkleg dokumentasjon

8.3.1 Verdien av språkleg dokumentasjon

Alle språk som skal haldast i hevd og utvikla seg som levande kulturspråk, må kunna støtta seg på ein solid dokumentasjon av den språklege tradisjonen og av korleis språket blir brukt i samtida. Ikkje minst er dette eit nødvendig grunnlag for språkforskning og språkopplæring, men også for arbeidet med standardisering og normering av språket. Det har også mykje å seia for å gje språket status at det er vel dokumentert på ulike vis.

Verdien av språkleg dokumentasjon har også samanheng med at språket er ein så viktig del av kulturarven vår. Til dømes representerer gammalt ordtilfang eit uvurderleg kjeldemateriale for studiet av den indre språkhistoria og av tilhøvet mellom vårt eige språk og tilgrensande eller fjernare språk. I tillegg kan det fortelja mykje om den vidare kultur- og samfunnsutviklinga. I dag er dokumentasjon av språk i form av digitale språkressursar dessutan eit grunnlag for all slags språkteknologisk forskning og utvikling.

Å dokumentera korleis norsk språk har vore brukt i mellom anna skjønnlitteratur og sakprosa i fjern og nær fortid, og korleis det i dag blir brukt i ulike samanhengar, både skriftleg og munnleg, er i seg sjølv ei stor kulturoppgåve og ei viktig nasjonal oppgåve. Eit språk som skal vernast og styrkjast, som skal stå imot eit ytre press som i dag kanskje er sterkare enn på lenge, må ikkje minst kunna støtta seg på ein solid dokumentasjon. Her står ordbøker og elektroniske ordbasar heilt sentralt. I dag ligg Noreg på dette området etter mange andre land, både dei næraste nordiske grannelanda våre og fleire andre nasjonar.

8.3.2 Generelt om språksamlingar

8.3.2.1 Innhald og organisatorisk status

Ein viktig del av grunnlaget for dokumentasjonen av norsk språk, til dels også samisk og kvensk, er ulike typar språkmateriale som er samla gjennom åra, særleg knytt til universiteta, men også til dels utanfor desse sentrale forskingsmiljøa. Det dreier seg her om verdifulle forskings- og dokumentasjonsressursar som er bygde opp over fleire generasjonar.

Nokre av desse samlingane kan identifiserast gjennom namn som også indikerer ein eigen organisatorisk status på nasjonalt nivå, så som Norsk målførearkiv, Norsk stadnamnarkiv, Norsk folkeminnesamling og Runearkivet. Men dette er berre ein del av det heile.

Samlingane består først og fremst av papirbasert materiale i form av manuskript, katalogar, hefte, bøker og arkivkort, men også ein del talemålsmateriale på lydband og grammofonplater.

8.3.2.2 Evaluering av dei universitets- og høgskulebaserte samlingane

Situasjonen for dei universitets- og høgskulebaserte samlingane er omtalt i ein rapport frå mars 2007 **1** frå eit utval oppnemnt av Noregs forskingsråd for å følgja opp ein tidlegare evalueringsrapport frå 2005 **2** om forskning i norsk språk og litteratur ved universitet og høgskular. Her heiter det at tilhøva rundt desse samlingane gjev grunn til uro, men at digitalisering av papirbasert materiale også opnar for både nye utfordringar og nye måtar å utnytta materialet på.

Mykje namnemateriale vart digitalisert i Dokumentasjonsprosjektet ved universiteta på 1990-talet. Ein god del leksikografisk og språkhistorisk materiale er digitalisert som ledd i det store ordbokprosjektet Norsk Ordbok 2014.

Dei omtalte utfordringane er dels at det i dag ikkje skjer tilsvarende systematisk digitalisering av nyare materiale, at ikkje bokmålsmateriale og ordtilfang frå gammal- og mellomnorsk tid er vorte digitalisert tilsvarende dialekt- og nynorsktilfanget, og at det også er uklart kva som vil skje med det sistnemnde materialet når prosjektet Norsk Ordbok blir avslutta i 2014.

Rapporten reiser også spørsmål om vidare forvaltning av sikring av det materialet som enno ikkje er digitalisert, og generelt om rammevilkåra for arbeidet med oppbygging og forvaltning av samlingar, arkiv og digitale korpus.

I rapporten gjer oppfølgingsutvalet framlegg om at dei aktuelle institusjonane vurderer å ta initiativ til eit samarbeid om eit felles lyft for å modernisera det tradisjonelle samlingsarbeidet i eit nasjonalt digitalkorpus. Utvalet kastar fram tanken om ei nasjonal reorganisering av samlingsfunksjonane, ved at dei blir tekne ut av instituttstrukturen og flytta over i det dei kallar eit nasjonalt arkivverk.

8.3.2.3 Andre vurderingar av samlings- og digitaliseringssituasjonen

Departementet er kjent med at delar av det samlingsmaterialet som er digitalisert, ikkje er lagde til rette for forskning og formidling, og at det også her står att mykje arbeid å gjera.

Det synest å vera eit problem at dei digitaliseringstiltaka som er gjorde, ofte ikkje har vore samkøyrd, slik at resultatet er vorte fragmentering av dei digitale ressursane. Dette er ei av erfaringane som er gjorde av ei prosjektgruppe med basis ved Universitetet i Bergen som førebur eit stort forskingsprosjekt om norsk språkhistorie frå dei første kjeldene fram til i dag.

Ei førebuande kartlegging viser at språksamlingane ved universiteta er i svært ujamn tilstand, at dei ikkje alltid har hatt forsvarlege oppbevaringsforhold. Det er også døme på at innsamla materiale er øydelagt. Det synest å vera lang veg fram til fullførd digitalisering og tilgjengeleggjering av det som finst.

Departementet er kjent med at prosjektgruppa har teke kontakt med aktuelle faginstittutt ved dei andre universiteta for å få vurdert utsiktene til å kunna byggja opp eit felles nasjonalt digitalt dialektarkiv.

8.3.2.4 Oppsummerande vurdering

Dokumentasjon av norsk språk er ei permanent, nasjonal oppgåve. Gode, oppdaterte samlingar er ein føresetnad for forskinga.

Dei spørsmåla som er drøfta ovanfor, ligg i utgangspunktet utanfor det som er eit direkte sektoransvar for Kultur- og kyrkjedepartementet.

Som ledd i ein meir aktiv, sektorovergripande språkpolitikk vil departementet likevel ta opp desse spørsmåla med involverte aktørar med sikte på å få eit samla oversyn over situasjonen som grunnlag for ei overordna nasjonal løysing.

Det vil mellom anna vera naturleg å sjå spørsmålet om dokumentasjon av norsk språk og norske språksamlingar i samanheng med den satsinga på digitalisering av dokumentasjons- og kulturarvsmateriale som elles går føre seg med basis i institusjonar innanfor arkiv-, bibliotek- og museumssektoren.

Regjeringa vil presentera ein strategi for digital bevaring og tilgjengeleggjering av slikt materiale i ei eiga stortingsmelding som etter planen skal leggjast fram hausten 2008. Strategien vil omfatta institusjonar både i kultursektoren og i universitets- og høgskulesektoren. Dokumentasjon av norsk språk og norske språksamlingar vil inngå som eitt av mange faglege tema i ein slik overordna strategi og vil kunna behandlast nærare som ledd i oppfølging av digitaliseringsmeldinga.

8.3.3 Spesielt om innsamling av stadnamn

8.3.3.1 Å hindra eit omfattande kulturtap

Ei av dei samlingane som er nemnde ovanfor, er Norsk stadnamnarkiv, og det kan vera grunn til å sjå noko nærare på temaet stadnamn, både som eit døme på meir generelle problemstillingar, og fordi det her er eit meir spesielt spørsmål som kanskje blir sett ekstra på spissen.

Det spesielle er at store mengder gamle stadnamn som til no ikkje er systematisk samla inn og registrerte, representerer ein kunnskapsbase som kan gå i gløymeboka etter kvart som dei personane som har kjennskap til dei, fell ifrå. Dette vil i så fall bety eit uerstatteleg kulturtap. Her er det altså ikkje berre, og ikkje først og fremst, tale om å ta vare på og å digitalisera materiale som alt finst i samlingar og arkiv, men like mykje eit spørsmål om nødvendig nyinnsamling for å sikra ein kulturarv som elles står i fare for å gå tapt for alltid.

8.3.3.2 Verdien av bevarte stadnamn

Utviklinga frå det agrare til det stadig meir urbaniserte samfunnet og overgangen til maskinelle driftsformer innanfor primærnæringane gjer at det tette nettet av stadnamn til lands og til sjøs ikkje er i levande bruk slik som før.

Framvoksteren av ein allmenn skriftkultur og eit stadig meir skriftbasert talemål har dessutan redusert dei tradisjonelle talemåla si stilling som kommunikasjonsform. Vi har dermed fått andre standardar for munnleg overlevering og munnleg kommunikasjon. Sjølv om store delar av folket framleis nyttar dialekt, har terminologien endra seg.

Resultatet av alt dette er at overføringa av namn og termar frå ættledd til ættledd har stansa opp dei to-tre siste generasjonane, og ei mengd tradisjonsstoff har alt vorte borte saman med dei siste berarane.

I levande bruk er stadnamn først og fremst språklege uttrykk, munnlege eller skriftlege, som gjer oss i stand til å skilja mellom ulike lokalitetar. Slik har dei stor praktisk interesse på kart, skilt, vegvisarar og andre stader. For Statens kartverk og Stadnamntenesta i staten er eit godt utbygd og vitskapleg påliteleg stadnamnmateriale ein føresetnad. Det er ikkje sjeldan at namnekonulentane opplever at det manglar opplysningar om uttale og bakgrunn for namna. Også dette tilseier at det er viktig å byggja opp namnesamlingane vidare.

Stadnamn er dessutan historiske dokument som fortel om natur- og kulturtilhøve i det området dei er henta frå. Eit detaljert stadnamntilfang frå eit område gjer dermed at ein kan lesa den menneskelege bruken av området gjennom

ulike historiske lag.

I den reviderte lova om stadnamn som tok til å gjelda frå 1. august 2006, er det uttrykkjeleg slått fast at ein del av formålet med lova er å ta vare på stadnamn som kulturminne. Det som gjer stadnamn særleg viktige som språklege kulturminne, er at dei er knytte til visse punkt, linjer eller område i landskapet, og at dei inneheld informasjon om korleis namngjevarane opphavleg oppfatta lokaliteten på den tida namnet vart til. Dei representerer dermed viktige informasjonsbitar av fortida i det aktuelle området.

Mange stadnamn har dessutan ei alderdomleg språkform som til dels ikkje lenger er i levande bruk i talemålet. Stadnamn er såleis verdifullt språkleg tradisjonsstoff både med omsyn til innhald og form.

Stadnamn er også ein del av ein munnleg tradisjon som Noreg har teke på seg eit særskilt ansvar for gjennom ratifikasjon av UNESCO-konvensjonen om vern av den immaterielle kulturarven. Konvensjonen av 17. oktober 2003 som Noreg ratifiserte 17. januar 2007, omfattar etter artikkel 3 mellom anna munnlege tradisjonar og uttrykk, under dette språk som eit uttrykksmiddel for immateriell kulturarv. Som relevante vernetiltak nemner konvensjonen mellom anna kartlegging, dokumentasjon, forskning og bevaring.

8.3.3.3 Tidlegare innsamling av stadnamn

Heilt frå Norsk stadnamnarkiv vart grunnlagt i 1921 og fram til om lag 1990, har det gått føre seg ei systematisk innsamling av norske stadnamn. Fram til andre verdskrigen var innsamling gjennom skulebarn ein berebjele, men innsamlingsarbeid vart også utført av hovudfagsstudentar og vitskapleg tilsette ved Stadnamnarkivet. Dette sistnemnde innsamlingsarbeidet heldt fram i stor stil etter krigen, og mykje vart også samla inn av hovudfagsstudentar som valde namnegransking som emne for hovudoppgåva si.

Etter kvart som universiteta i Bergen, Trondheim og Tromsø vart skipa, har det også ved desse institusjonane skjedd vitskapleg basert innsamlingsarbeid. Likeins har det ved det som var distriktshøgskulane i Rogaland og Møre og Romsdal vore mykje innsamlingsarbeid, dels finansiert av institusjonane sjølve, dels av fylkeskommunane. I Nordland fylke har eit omfattande innsamlingsarbeid skjedd i regi av Arkiv Nordland, og ikkje minst i Sogn og Fjordane er det bygd opp omfattande samlingar i stor grad på grunnlag av lokalt innsamlingsarbeid.

Det er da også Vestlands-fylka og Nordland som i dag er best dekte med stadnamnsamlingar. På Austlandet og Sørlandet er det utført djupinnsamling i mange enkeltkommunar, men delar av Austlands-området og heile Trøndelag, Troms og Finnmark har store hol i samlingane. Dei samlingane som finst, har varierende kvalitet når det gjeld opplysningar om uttale og lokale tilhøve, også når det gjeld heimfesting og kartfesting.

Fordi dei namnefaglege miljøa ikkje har hatt kontroll med alt innsamla materiale, er det også verdifullt tilfang som er kome bort. Det gjeld mellom anna primærlistene til innsamlinga av namn på Økonomisk kartverk, og det gjeld ikkje minst tilfang på private hender.

Den siste store stadnamninnsamlinga gjekk føre seg frå midten av 1980-talet, utført av unge arbeidsledige og administrert av kulturkontora i kommunane med fagleg tilsyn frå dei aktuelle institusjonane.

8.3.3.4 Innsamlingssituasjonen i dag

Det har vore rekna ut at det finst, eller har funnest, fem–seks millionar stadnamn her i landet, og berre ein liten del av dei er samla og registrert ved dei sentrale og regionale stadnamnsamlingane.

Etter 1990 har det ikkje vore råd å finansiera nemnande innsamlingsarbeid, og det som likevel er samla, har skjedd i regi av interesserte amatørar under ei viss rettleiing frå vitskapeleg personale ved universitet og høgskular.

8.3.3.5 Oppsummerande vurdering

I kulturmeldinga frå Bondevik II-regjeringa heitte det at departementet ville ta initiativ til ei kartlegging av kva som er samla inn av gamle stadnamn her i landet, og vurdere behovet for eit meir systematisk innsamlingsarbeid. Det heitte også at departementet ville vurdere om det nye kompetansesenteret for norsk språk, dvs. Språkrådet, ville kunna ta på seg å koordinera desse oppgåvene.

Ettersom Språkrådet har vore inne i ein langvarig omdanningsprosess med fleire store utfordringar, har departementet ikkje funne grunnlag for å leggja endå ei oppgåve til denne institusjonen i denne fasen. Språkrådet har heller ikkje tilsette med namnefagleg kompetanse og erfaring med innsamling, arkivering og formidling av namn.

Dette inneber at det no hastar med å koma fram til ei løysing i spørsmålet om innsamling av stadnamn, og det hastar også med å få sentralt oversyn over materiale som er samla inn lokalt for å hindra at det skal gå tapt.

I første omgang kan det vera aktuelt å setja av midlar til eit prosjekt for å registrera og kvalitetsvurdere dei stadnamnsamlingane som finst rundt om i landet, både i ulike samlingar og arkiv og på private hender.

Departementet vil elles vurdere om det bør gjennomførast ei større utgreiing om den vidare innsamlingspolitikken og tilgrensande spørsmål. Det er i så fall aktuelt å sjå på om ei slik utgreiing bør femna vidare enn berre stadnamn. Heile situasjonen for dokumentasjon av norsk språk og for dei ulike språksamlingane som er behandla i førre avsnitt, vil kunna vera eit naturleg utgangspunkt for ei nærare avgrensing av mandatet for ei eventuell slik utgreiing.

8.3.4 Ordbøker og andre leksikalske oppslagsverk

8.3.4.1 Ordbøker og leksika

Alle former for ordbøker og leksikalske oppslagsverk utgjer ein viktig del av den totale skriftkulturelle infrastrukturen i eit språksamfunn. Alt dette er på sitt vis med på å styrkja skriftkulturen spesielt og nasjonalspråket generelt.

Rett nok er ikkje nemningane ordbøker og leksika nødvendigvis uttømande referansar for alle dei sjangervariantane og boktypane som med større eller mindre rett kan gjera krav på å tilhøyra det vi her har kalla den skriftkulturelle infrastrukturen.

Det går heller ikkje an å trekkja noko skarpt skilje mellom leksikonet og ordboka. Fellestrekket er at det dreier seg om kunnskap som er organisert på systematisk måte gjennom oppslagsord som følgjer etter kvarandre i alfabetisk orden. Medan ordboka inneheld meir eller mindre omfattande kunnskap om orda sjølve, er leksikonet ei kjelde til kunnskap om dei tinga, fenomenane eller saksforholda som orda refererer til.

8.3.4.2 Leksika

Det er ei kjend sak at vi i Noreg har ein særleg sterk tradisjon når det gjeld leksika. Leksikonverk, som regel i fleire band, har det funnest mange av, og dei har til vanleg vist seg å vera ei god inntektskjelde for forlaga. Interessa for leksikalsk kunnskap har på mange måtar vist seg å vera noko av eit norsk særkjønne; nordmenn har lege i europatoppen som leksikonopplåande folkeferd dei siste tretti åra. Marknaden for leksikalske verk har syntest uutømmeleg.

Slik er det ikkje utan vidare i dag. Éin ting er at mykje kunnskap i dag er flyktigare enn før; det som var uomtvisteleg fakta i går, er ikkje nødvendigvis gyldig kunnskap i dag. Kunnskapsutviklinga går jamt raskare framover. Éin annan ting er at særleg Internett har vist seg å fungera som ein konkurrent på kunnskapsmarknaden.

På Internett finst rett nok eit vell av informasjon om det meste, men all slik informasjon er ikkje nødvendigvis det same som redigert informasjon. Det som finst på nettet, er kvalitetssikra i ulik grad. Difor treng vi også i dag og for framtida nasjonale kunnskapsbasar som er kvalitetssikra på ein måte som gjer at vi kan stola meir på den informasjonen som der finst.

Kunnskapsutviklinga i landet vårt er avhengig av slike kunnskapsbasar, men dei har også ein viktig funksjon i å halda hevd eit norsk språk som er fullverdig oppdatert i takt med kunnskapsutviklinga. Eit livskraftig nasjonalleksikon, som i dag må finnast i digital grunnform, er ein nødvendig del av ein språkleg infrastruktur for framtida.

Eit nasjonalleksikon er i seg sjølv ein kulturinstitusjon. Han består av handplukka ekspertar og fagforfattarar frå alle tenkjelege fagmiljø i landet, og det er dermed desse fagmiljøa som går god for heile spekteret av fagleg spesialisert informasjon som på same tid er framstilt på norsk, i eit språk som alle skal kunna forstå, utan at dei på nokon måte er ekspertar sjølve.

Departementet legg til grunn at eit livskraftig digitalt nasjonalleksikon er eit viktig instrument for å sikra norsk språk som formidlingsspråk for fagfolk, fagforfattarar, skuleelevar og studentar. Eit kjenneteikn ved eit slikt nasjonalleksikon er at apparatet av omgrep og vokabular i dei ulike faga blir røkta av spesielt utpeikte fagansvarlege ekspertar med sans også for dei språklege krav.

8.3.4.3 Tospråklege ordbøker

Eit hovudskilje når det gjeld ordbøker, går mellom dei eittspråklege på den eine sida og dei to- eller fleispråklege på den andre sida.

Tospråklege eller fleispråklege ordbøker er ordbøker som forklarar og definerer ordtilfanget mellom ulike språk. For alle som har norsk som morsmål og skal læra seg eit framandspråk, er ei ordbok mellom norsk og vedkomande språk nærast eit uunnverleg hjelpemiddel.

Til skulebruk finst det stort sett eit tilfredsstillande tilbod av tospråklege ordbøker, i alle fall for dei største og mest skulerelevante framandspråka. Derimot manglar vi oppdaterte, større ordbøker mellom norsk og sentrale framandspråk som fransk, italiensk og spansk. Likeins skortar det på gode ordbøker mellom dei skandinaviske språka. Ordbøker mellom norsk og smalare europeiske språk er eit endå større mangelområde, for ikkje å nemna mellom norsk og ikkje-europeiske språk.

Her ligg store utfordringar for framtida. I utgangspunktet er dette ei forlagsoppgåve, men mange ordbokprosjekt av dette slaget har ikkje noko stort innteningspotensial, og finansieringa er ofte eit hovudproblem.

Tospråklege ordbøker er utan tvil ei rikdomskjelde for norsk språk ved at det sentrale ordforrådet i språket vårt på den måten blir definert i høve til så mange andre språk som råd er. Det kastar nytt lys over norske ord og uttrykk når vi kan relatera dei til ord og uttrykk i andre språk, for relasjonane mellom ord og ordtydingar i ulike språk er noko langt meir komplekst enn eit ein-til-ein-forhold.

Tospråklege ordbøker mellom norsk og flest moglege framandspråk vil gjera norsk språk lettare tilgjengeleg for ein større del av omverda, for ein større kulturkrins. Det vil føra til større merksemd om norsk språk og dermed vera med på å gje norsk språk høgare internasjonal status.

Tospråklege ordbøker etablerer eit direkte samband mellom norsk og vedkomande framandspråk. Dei legg grunnlaget for språkkontakt og kulturutveksling utan at ein treng å gå omvegen om engelsk eller eit anna felles framandspråk. Tospråklege ordbøker vil gjera det lettare å omsetja mellom norsk og vedkomande framandspråk. Ikkje minst vil dei stimulera til at nordmenn lærer seg eit større spekter av framande språk. Tilsvarende vil dei vera eit uunnverleg hjelpemiddel for opplæring i norsk språk i utlandet.

8.3.4.4 Eittspråklege ordbøker

Av eittspråklege ordbøker, dei som for vårt vedkomande konsentrerer seg om norsk språk åleine, finst det mange slag. Eit hovudskilje kan seiast å gå mellom allmennordbøker og spesialordbøker.

Blant spesialordbøker kan vi rekna terminologiske ordbøker, dei som definerer spesialtermar innanfor eit fagspråk. Men det kan også tenkjast andre typar spesialordbøker, til dømes etymologiske ordbøker, stilordbøker, uttaleordbøker og framandordbøker.

I eit framtidig arbeid med å dokumentera norsk språk og norsk skriftkultur er det behov for å utvikla fleire slike spesialordbøker. Her står alt i alt mykje att å gjera dersom vi samanliknar situasjonen i Noreg med det tilbodet av slik ordboklitteratur i fleire større kulturland.

Dersom vi held oss til dei meir tradisjonelle sjangrane innanfor det vi kan kalla allmennordbøker, kan vi elles grovt sett operera med ei tredeling, utan at det går skarpe grenser mellom dei ulike kategoriane.

Det første og enklaste nivået er ordbøker med hovudvekt på rettskriving, mest typisk ordlister til skulebruk. Her kjem talet på oppslagsord til vanleg ikkje over 30 000.

Det andre nivået er eittbands definisjonsordbøker, gjerne omtalte som handordbøker. I tillegg til å visa korleis orda skal stavast og bøyast, gjer dei også informasjon om kva orda tyder, supplert med praktiske bruksdøme, og elles opplysningar om ordhistorie og faste ordsamband med forklaringar.

Desse to kategoriane kjem vi nærare inn på under kap. 8.4.4 nedanfor i samband med behandlinga av rettskrivingsspørsmålet.

8.3.4.5 Generelt om dokumentasjonsordbøker

Også rettskrivings- og dei enkle definisjonsordbøkene har til ein viss grad ein dokumentasjonsfunksjon.

Men det er dei store, vitenskaplege ordbøkene som her er grunnleggjande, og som ofte fungerer som kjelde- og referanseverk for dei mindre ordbøkene. Dokumentasjonsordbøkene er ikkje primært rettskrivingsordbøker, men har først og fremst til formål å forklara og eksemplifisera kva orda tyder, og korleis dei blir brukte.

Dokumentasjonsordbøkene er eit særleg viktig og allment tilgjengeleg resultat av språkleg dokumentasjonsarbeid. Dei representerer den mest avanserte kategorien av ordbøker. Allmenne dokumentasjonsordbøker er vitenskaplege ordbokverk i fleire band, som regel med fleire hundre tusen ord og fleire og fylldigare opplysningar om dei enkelte orda enn det som er vanleg i andre slags ordbøker.

8.3.4.6 Langsiktige prosjekt av høg kulturell verdi

Bak dei store vitenskaplege ordbokverka ligg det som regel eit omfattande arbeid av høgt kvalifiserte fagfolk. Mykje slikt språkleg dokumentasjonsarbeid har vore drive fram av eit fåtal fagpersonar, til dels ved sida av anna fagleg arbeid, delvis også som ubetalt arbeid på fritida.

Uavhengig av korleis arbeidet blir finansiert, er det her tale om investering i eit produkt som har stor kulturell og vitenskapleg verdi, ein verdi som ikkje minkar ved bruk, men som tvert imot vil stå seg langt fram i tid, og som seinare dokumentasjon av den framtidige språkutviklinga kan byggja vidare på. For eit språk og ein skriftkultur har slikt ordbokarbeid ein uvurderleg verdi.

Like fullt har det vist seg svært krevjande å finansiera slike ordbokprosjekt. Det dreier seg om ei form for grunnforskning som til vanleg ikkje kan rekna med å gje kortsiktig gevinst rekna i kroner og øre. Arbeidet er langsiktig, og salspotensialet for dei ferdige bøkene er heller lite.

Både prosjektstorleiken, til dels mangel på kvalifiserte fagfolk og problemet med å finansiera arbeid som dette har tradisjonelt ført til at slike dokumentasjonsprosjekt har pågått i generasjonar. Å få fullført slike vitenskaplege ordbokverk har vore samanlikna med å byggja gotiske katedralar.

Til illustrasjon kan nemnast at Svenska Akademiens ordbok, som kan skriva historia si attende til 1787, enno ikkje er mål. Arbeidet kom i gang tidleg på 1800-talet, men det første heftet kom ikkje ut før i 1893, og det første heile bokbandet låg føre fem år seinare. I 2007 var ein kome så langt i alfabetet som til midt i bokstaven T.

Svenska Akademiens ordbok er eit historisk ordbokverk som dokumenterer all skriven svensk frå 1521 til våre dagar. Det finst berre tre andre ordbokverk over germanske språk som i storleik kan jamførast med Svenska Akademiens ordbok. Det er *Deutsches Wörterbuch*, *Woordenboek der Nederlandsche Taal* og *The Oxford English Dictionary*.

I dag kan slikt ordbokarbeid effektiviserast monaleg ved å leggja tekstmateriale og ordstoff inn i elektroniske databasar og behandla og redigera det ved hjelp av språkteknologiske metodar. På midten av 1980-talet rekna ein såleis med at Svenska Akademiens ordbok kunne fullførast i 2045, men etter at ein i 1993 gjekk over til elektronisk redigering, reknar ein no med at sluttstrek kan setjast i 2017.

Ordboksmaterialet kan også gjerast elektronisk tilgjengeleg over nettet, og difor er det ikkje sikkert at resultatet av dokumentasjonsarbeidet i framtida nødvendigvis vil bli spreidd i bokform, og i alle fall ikkje berre i bokform.

Oppbygging av elektroniske orddatabasar gjer også at materialet lettare kan gjerast tilgjengeleg etter kvart som arbeider skrir fram, i staden for at ein må venta til heile verket er fullført, eller i alle fall til eit nytt band kan gjevast ut.

8.3.4.7 Ordboksituasjonen i Noreg jamført med andre land

At eit språk eller ei språkform blir grundig utforska og dokumentert, må seiast å vera noko av det som er med på å gje det prestisje og status. Men her er det nok også ein veksilverknad. Eit språk som alt har stor prestisje og høg status, kan slik sett vera lettare å få dokumentert gjennom slike ordbokprosjekt enn språk med lågare status.

Både når vi ser på ordboksituasjonen generelt og sjangeren dokumentasjonsordbøker spesielt, kan vi ikkje seia at

Noreg høyrer med blant dei mest velutvikla ordboklanda. Til dels må nok dette sjåast i samanheng med at norsk har ein kortare samanhengande skriftspråkstradisjon enn både dansk og svensk, for ikkje å snakka om store europeiske kulturspråk som engelsk, tysk og fransk.

Når det gjeld dokumentasjonsordbøker, kan her nemnast at til dømes Danmark har to fullførte ordbokverk. Det største, *Ordbog over det danske Sprog*, dekkjer dansk språk frå ca. 1700 til 1950, kom ut i 28 band i perioden 1918–56 og er i dag supplert med fem nye band som vart utarbeidde i perioden 1956–2004. Og så har Danmark i tillegg eit nyare verk som dokumenterer moderne dansk språk. Det er seksbandsverket *Den danske ordbog*, som kom ut i perioden 2003–2005.

Men når vi samanliknar med grannespråka våre, svensk og dansk, må vi også sjå i auga at vi her i landet har ein meir innfløkt språksituasjon å dokumentera, med to offisielle skriftmål, ofte med fleire alternative skrivemåtar og bøyingsformer, i tillegg ein uoffisiell riksmålsvariant, dessutan eit stort spekter av ulike målføre.

Når det nettopp gjeld målføra, dialektane, står likevel Noreg på mange måtar i ei særstilling i dokumentasjonssamheng. Det har samanheng med at ein innanfor den nynorske tradisjonen reknar talemålet som ei like god kjelde for nytt ordtilfang som skrift. Dei store ordbokverka for svensk og dansk inneheld såleis stort sett ikkje opplysningar om talemål. Dette representerer ein markert skilnad jamført med det største ordbokprosjektet som vi har gåande her i landet for tida, kalla Norsk Ordbok 2014.

8.3.5 Spesielt om Norsk Ordbok

8.3.5.1 Stutt historisk riss

Dei første planane for dette ordbokverket går attende til slutten av 1920-åra. Ambisjonen var å gje ei uttømmende framstilling av ordtilfanget i norske dialektar og nynorsk skriftspråk. Prosjektet fekk difor undertittelen *Ordbok over det norske folkemålet og det nynorske skriftmålet*.

I 1930- og 40-åra vart det arbeidd fram eit grunnmanus samanredigert av eldre nynorske ordsamlingar og ordbøker, og det er i tillegg bygd opp eit omfattande ordarkiv basert på både munnlege og skriftlege kjelder. Dette omfattar i dag om lag 3,2 millionar setlar.

Den første fullførte delen av *Norsk Ordbok* var eit hefte som kom ut i 1950, medan det første heile bokbandet, på 750 sider, tilsvarande fem hefte, låg føre i 1966 og det andre i 1978. Då var ein ikkje komen lenger enn halvvegs i bokstaven f.

Ambisjonane for verket hadde vakse undervegs, og ressursane var små. Då tredje bandet kom ut i 1994, gjekk dei vidare planane ut på at verket i alt skulle omfatta tolv band, men det var små utsikter til å få eit så stort arbeid fullført i ei overskodelig framtid.

Dette var bakgrunnen for at Kulturdepartementet i St.meld. nr. 22 (1999–2000), meldinga om arkiv, bibliotek og museum, tok til orde for ei monaleg opptrapping av ressursinnsatsen til Norsk Ordbok og sette som mål at verket skulle fullførast til 200-årsmarkeringa for Grunnlova i 2014.

Norsk Ordbok hadde fått diverse statstilskot heilt sidan 1929, men frå 1987 vart dette auka gjennom eit øyremerkt tilskot over kulturbudsjettet, som første året var på 0,8 mill. kroner, og som i 2001 var kome opp i nærare 1,7 mill. kroner.

8.3.5.2 Reorganisering av prosjektet og opptrapping av den statlege innsatsen

Omtalen av Norsk Ordbok i meldinga frå 1999 var basert på nye initiativ og nye planar med sikte på monaleg effektivisering og påskunding av arbeidet. Dette vart seinare følgt opp gjennom eit særskilt forprosjekt og etter kvart konkretisert i form av detaljerte framdrifts- og gjennomføringsplanar innanfor ramma av ein særskild prosjektorganisasjon under leiing av ein prosjektdirektør.

Dette la grunnlaget for at departementet kunne trappa opp statstilskotet, først med 5 mill. kroner frå 2002 og seinare med nesten like mykje fordelt over åra 2005 til 2007. For 2008 er det lagt på endå 1,5 mill. kroner. Inkludert prisomrekning over åra er det samla tilskotet over kulturbudsjettet for 2008 kome opp i 13,6 mill. kroner. Tilskotet over kulturbudsjettet går uavkorta til å dekkja lønnskostnader.

I tillegg har Norsk Ordbok finansiell støtte frå Universitetet i Oslo, der verksemda i hovudsak er lokalisert. Norsk Ordbok disponerer såleis fleire årsverk ved Institutt for lingvistiske og nordiske studium, i tillegg til at universitetet dekkjer driftsmidlar og yter administrative tenester. Norsk Ordbok disponerer også eitt årsverk ved Eining for digital dokumentasjon.

Denne eininga stod i spissen for det store Dokumentasjonsprosjektet ved Universitetet i Oslo i 1990-åra, der også det viktigaste kjeldematerialet for Norsk Ordbok vart digitalisert. Det digitaliserte materialet er no tilgjengeleg som søkje-teneste over Internett. Digitaliseringa var elles ein av føresetnadene som gjorde det mogleg å effektivisera utgjevingsarbeidet frå 2002.

Verksemda vart då organisert som eit sjølvstendig prosjekt – Norsk Ordbok 2014 – lausrive frå instituttstrukturen ved universitetet, og det er oppnemnt eksternt styre i tillegg til den nemnde prosjektdirektøren. Prosjektet er elles basert på ei gammal samarbeidsavtale mellom Universitetet i Oslo og Det Norske Samlaget.

Digitaliseringa, ressursauken og prosjektorganiseringa frå 2002 har ført til at det alt er kome ut tre nye band i tillegg til dei tre som var komne fram til 1994. Det er i tillegg laga ein bindande utgjevingsplan for dei resterande seks banda. Dei skal koma ut med om lag eitt års mellomrom fram til planlagd fullføring i 2014, i samsvar med føresetnaden i

meldinga frå 1999. Kvart band skal vera på ca. 800 tospalta sider og innehalda om lag 25 000 ordartiklar, og heile verket vil dermed bli på nærare 9 600 sider og omfatta 300 000 ordartiklar.

8.3.5.3 Oppsummerande vurdering

Norsk Ordbok er på fleire måtar eit eksepsjonelt prosjekt i norsk samanheng. Den uvanleg store auken i det årlege statstilskotet dei seinare åra representerer ein sterk konsentrasjon av den statlege innsatsen. Alternativet til ei opptrapping av ressursinnsatsen slik det såg ut i 1999, var at verket ikkje ville vera fullført før om lag i 2060.

Med det noverande tilskotsnivået over kulturbudsjettet saman med medverknaden frå Universitetet i Oslo kan vi vera tilnærma sikre på at verket er i mål i 2014. Det inneber at allereie nolevande generasjonar kan dra full nytte av den kulturelle og vitskaplege verdien som verket representerer.

Departementet legg til grunn at prosjektet no er fullt ut finansiert så sant tilskotsnivået for 2008 blir vidareført fram til og med 2014.

Departementet vil følgja opp dette i budsjettsamanheng i dei åra som står att fram til og med 2014.

Departementet legg også til grunn at 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 språkarv som direkte vedkjem ein mykje større del av norske språkbrukarar enn dei som nyttar nynorsk som skriftmål.

Det at Norsk Ordbok framstiller skriftspråket og talemålet som éi eining, gjer også dette til eit sjeldan verk i europeisk samanheng.

8.3.6 Nærare om dokumentasjonen av bokmål

8.3.6.1 Utgangspunkt og planar

I tillegg til å fullføra arbeidet med Norsk Ordbok er det overlag viktig å få dokumentert også den dominerande målforma på ein fullgod måte. Rett nok har vi frå før eit framifrå ordbokverk i *Norsk Riksmålsordbok*. Dette er også eit fullført verk i den meininga at det dekkjer heile alfabetet frå A til Å.

Det er altså Norsk Riksmålsordbok som i dag er det einaste komplette, vitskaplege ordboksverket for norsk språk, og det har det vore i femti år. Norsk Riksmålsordbok er dermed grunnleggjande også som kjelde og referanseverk for andre norske ordbøker, også for Norsk Ordbok, og det er fagleg godt omtykt. Men det er sett i gang eit arbeid med å vidareføra det eksisterande verket med sikte på ei revidert, oppdatert og utvida utgåve, og finansieringa av dette prosjektet har ikkje vore avklart.

I tillegg finst planar for eit anna stort leksikalsk dokumentasjonsprosjekt for bokmål, særleg for å dokumentera det moderne bokmålet. Institutt for lingvistiske og nordiske studium ved Universitetet i Oslo har såleis dei siste åtte-ni åra bygd opp eit stort elektronisk tekstkorpus gjennom systematisk utval av alle slags teksttypar på bokmål henta frå perioden 1985 til i dag.

Korpuset omfattar ca. 40 millionar ord og er komponert slik at det samsvarer med det ein gjennomsnittsperson les og skriv i vår tid, og gjev mellom anna grunnlag for å finna ut kva bøyingsformer og variantformer som er mest brukte av ulike språkbrukarar og i ulike teksttypar. Korpuset er laga først og fremst for leksikalske formål, og tanken er at det skal danna grunnlag for å etablere ein database som skal gje ein vitskapleg basert dokumentasjon av ordtilfanget i moderne bokmål.

Departementet vart første gongen presentert for desse planane gjennom brev frå Universitetet i Oslo i mai 2004 og mars 2005. Prosjektet vart då omtalt som *Norsk bokmålsordbok*. Departementet har også registrert den offentlege debatten som gjekk føre seg i 2006 om behovet for eit stort, standard ordboksverk for moderne bokmål. På nyåret 2008 vart dei oppdaterte planane lagde fram for departementet gjennom ei informativ projektskisse.

I projektskissa er det totale omfanget av prosjektet estimert til 120–125 årsverk. Institutt for lingvistiske og nordiske studium har difor lagt til grunn at prosjektet må tilførast eksterne midlar og har vendt seg til departementet med tanke på å kunna koma i gang med prosjektet i større eller mindre skala.

Her er det altså tale om eit prosjekt som førebels ligg i startgropa, medan vidareføringa av det eksisterande verket Norsk Riksmålsordbok er eit prosjekt som for lengst er i gang. Prosjektet vart omtalt i den same stortingsmeldinga frå 1999 som la grunnlaget for påskundinga av arbeidet med Norsk Ordbok.

Rettane til Norsk Riksmålsordbok ligg hos Det Norske Akademi for Sprog og Litteratur (Akademiet). Akademiet har teke opp i seg det tidlegare Riksmålsvernet, ein språkpolitisk organisasjon som vart stifta i 1919, og som stod som utgjevar av den opphavlege utgåva av verket. Dette kom ut i fire band frå 1937 til 1957.

Seinare er verket supplert med to tilleggsband, som kom ut i 1995. Desse vart utarbeidde ved hjelp av elektronisk redigering. Til saman har dei seks eksisterande banda av Norsk Riksmålsordbok vel 5000 tospalta sider og omfattar ca. 220 000 oppslagsord.

Arbeidet med supplementsbanda til Norsk Riksmålsordbok fekk særskild støtte over kulturbudsjettet i åra 1979–1993, det første året på 34 000 kroner, auka til 180 000 kroner det sistnemnde året, og det vart gjeve eit eingongstilskot på

¹ Andersen, Per Thomas et al. 2007

² Gunnarsson, Britt-Louise et al. 2005

³ Brev frå Akademiet av 12.3.1998.

Strategi 2020

UiO • Universitetet i Oslo

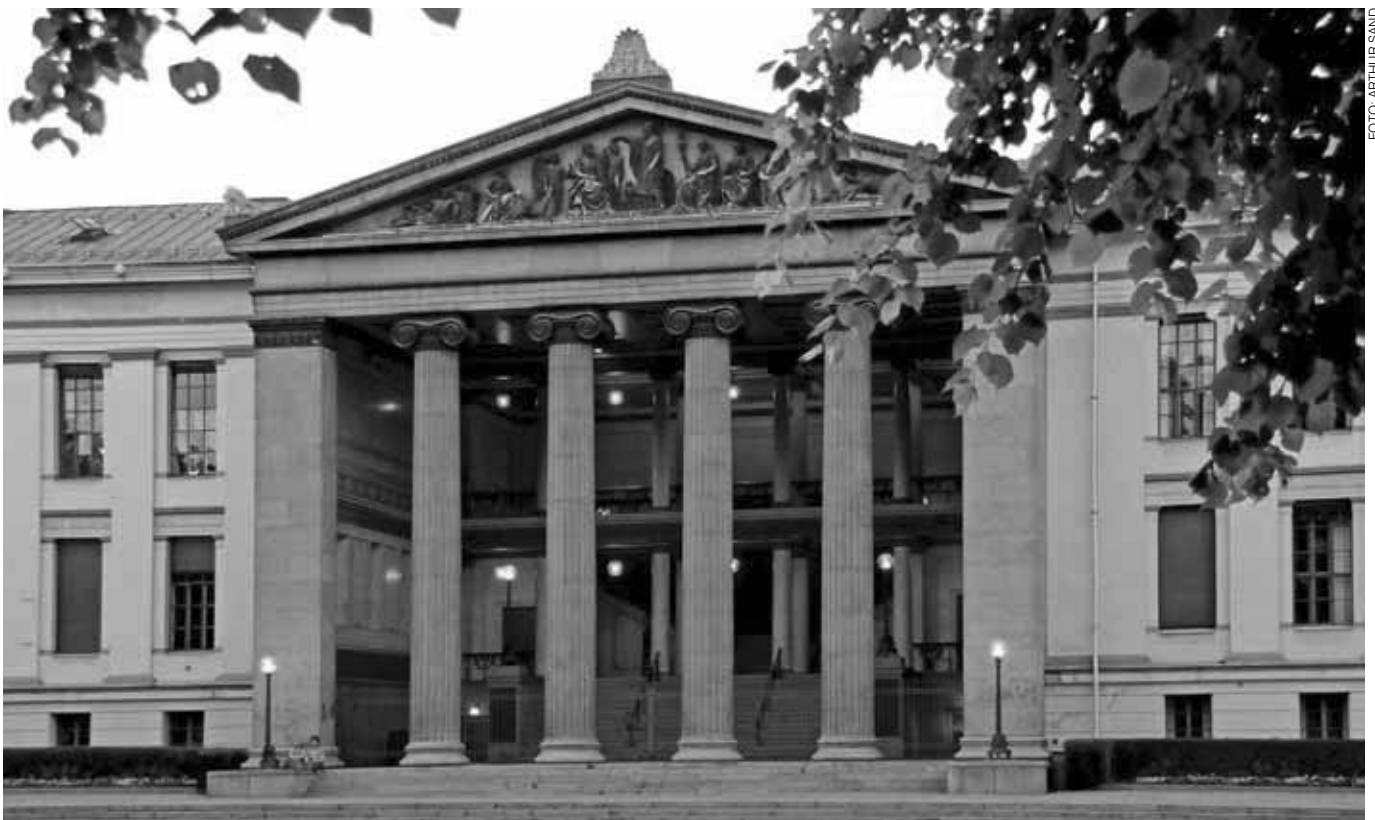


FOTO: ARTHUR SAND



Strategisk mål for UiO 2020:

UiO skal styrke sin internasjonale posisjon som et ledende forskningsuniversitet, gjennom et nært samspill mellom forskning, utdanning, formidling og innovasjon.

● Et grensesprengende universitet

MÅL 1: Universitetet i Oslo skal fremme grensesprengende forskning, utdanning og formidling og være en etterspurt internasjonal samarbeidspartner.

STRATEGIER:

1. Styrken i den faglige bredden skal utnyttes enda bedre gjennom tverrfaglig forskning og utdanning. Det skal utvikles gode finansieringsmekanismer for tverrfaglige aktiviteter og tidsavgrensede satsinger.
2. Det internasjonale engasjementet skal bli større, blant annet gjennom økt deltakelse i selektive, strategiske partnerskap og Det europeiske forskningsområdet.
3. Alle utdanningsprogrammer skal gis en internasjonal profil, og samarbeidet med utenlandske institusjoner skal økes for å gi større relevans og høyere kvalitet.
4. En målrettet rekrutteringspolitikk skal øke omfanget av internasjonal rekruttering.
5. Forskning i internasjonal front skal understøttes av strategiske investeringer og samarbeid om forskningsinfrastruktur.
6. Universitetet i Oslo skal i større grad profilere seg gjennom universitetsmuseene.

Fremragende forskning er naturlig grensesprengende ved at eksisterende kunnskap og dogmer utfordres, og ved at nye ideer prøves ut. Eksisterende grenser mellom fag overskrides, man samarbeider på tvers av landegrenser og mellom universiteter. Flere av det 21. århundrets globale utfordringer knyttet til temaer som klima, energi, helse, fattigdom og brudd på menneskerettigheter, kjenner verken disiplingrenser eller geografiske grenser. Her har et breddeuniversitet som Universitetet i Oslo et klart fortrinn. For å utnytte dette skal universitetet i større grad identifisere og utvikle tverrfaglighet i forskning og utdanning.

UiO vil i strategiperioden prioritere en tverrfaglig satsing på livsvitenskap (Life Science) i utvidet forstand, gjennom fagutvikling og nybygg. Her har UiO et særlig potensial som vil bidra til å realisere sentrale forskningspolitiske målsettinger og møte samfunnets behov for ny kunnskap innen helse, miljø og bærekraftig energi, samt effekter som globale klimaendringer har på liv og helse.

Den faglige aktiviteten skal være styrende for hvordan Universitetet i Oslo organiserer sin virksomhet. Dette forutsetter at institusjonen er fleksibel. For å styrke fleksibiliteten skal det etableres bedre mekanismer for å håndtere tidsavgrensede satsinger, slik at kompetanse ivaretas og rom skapes for nye satsinger.

Et breddeuniversitet har et særlig ansvar for å frembringe nye idéer og problemstillinger som utfordrer eksisterende kunnskap. Dette i erkjennelse av at

nye og radikale løsninger ofte springer ut av den frie grunnforskningen. Samtidig er det en viktig oppgave for universitetet å utvikle ny kunnskap for kunnskapens egenverdi. Det er derfor nødvendig at det er rom for forskning på områder som ikke er en del av nåværende kjerneområder.

Universitetet i Oslo skal intensivere internasjonaliseringen gjennom å samarbeide mer med de mest attraktive miljøene, og ved å konkurrere mer om de beste prosjektene, forskerne og studentene.

Forskningsmiljøene skal bidra til den internasjonale forskningsfronten og delta i gode internasjonale nettverk. UiO skal ta aktivt del i nordisk samarbeid og i Det europeiske forskningsområdet (ERA). Det skal bli lagt særlig vekt på å øke bevilgningene fra Det europeiske forskningsrådet (ERC), samt deltakelse i store europeiske prosjekter for forskningsinfrastruktur.

UiO skal være mer selektivt og målrettet i det institusjonelle samarbeidet, og prioritere langsiktig samarbeid med noen av de beste internasjonale forsknings- og utdanningsinstitusjonene. Samarbeidet med prioriterte land i andre verdensdeler skal styrkes, også med utvalgte institusjoner i sør. EU-samarbeidet vil være en plattform for internasjonalt samarbeid med resten av verden. Universitetet skal bidra til å styrke institusjoner og fagmiljøer i utviklingsland gjennom samarbeid basert på faglig kvalitet og gjensidighet. UiO vil særlig prioritere internasjonalt samarbeid om de store globale utfordringene, i tråd med egne faglige og tverrfaglige prioriteringer.

Internasjonalisering av studiene skal styrkes gjennom å satse på strategiske partnere og sikre avtaler med gode universiteter. Sammen med disse skal UiO etablere flere og

bedre internasjonale programmer, rekruttere flere utenlandske studenter og bruke flere internasjonale forelesere i studieprogrammene.

Universitetet opererer i et globalt arbeidsmarked for vitenskapelig ansatte hvor mobiliteten øker. Det innebærer bedre muligheter for rekruttering internasjonalt, men også at noen av de beste forskerne forlater UiO. Internasjonal mobilitet er en viktig faktor for kvalitet. For å oppnå dette skal fagmiljøene og enhetene utvikle sin rekrutteringspraksis, bl.a. gjennom aktive søk i internasjonale fagmiljøer. Alle utlysninger av vitenskapelige stillinger skal skje internasjonalt.

En sterkere satsing på internasjonalisering krever investeringer i Norge i form av økt språkkompetanse i forskning, undervisning og administrasjon. En internasjonal campus, et internasjonalt preg på UiOs nett og et profesjonelt mottaksapparat for internasjonale studenter og ansatte, er viktige elementer i internasjonaliseringen.

Tilgang til fremtidsrettet infrastruktur, særlig i form av IKT- og bibliotekjenester, er en forutsetning for høy kvalitet. Noen investeringer er for store for en institusjon å bære alene og krever derfor samarbeid mellom institusjoner og med myndigheter. UiO må bruke sin posisjon til å påvirke beslutninger om nasjonal deltakelse og ressursbruk i overnasjonale infrastrukturprosjekter.

Museenes omfattende samlinger skal utnyttes bedre både i forskning og undervisning. Ett virkemiddel er et styrket forskningsfokus ved museene. Som brobygger mellom samfunnet og universitetet, har museene en unik rolle gjennom sin utadrettede formidlingsvirksomhet. Her vil oppgradering og nybygg være sentrale virkemidler.