

LexicoNordica-symposium 32

# Nordiske ordbøger – opdatering, udvikling og tilgængeliggørelse

23.-25. januar 2025

Voksenåsen, Oslo – [www.voksenaasen.no](http://www.voksenaasen.no)

1

Temabeskrivelse og Abstracts

## *Nordiske ordbøger – opdatering, udvikling og tilgængeliggørelse*

### Temabeskrivelse

#### **LexicoNordica-redaktionen**

**Henrik Hovmark** (Ømålsordbogen, Københavns Universitet, [hovmark@hum.ku.dk](mailto:hovmark@hum.ku.dk))

**Terje Svardal** (Språksamlingane, Universitetet i Bergen, [terje.svardal@uib.no](mailto:terje.svardal@uib.no))

Nordiske ordbøger spiller en vigtig rolle i sikringen af nordisk sprogforståelse og dokumentationen af fælles nordisk kulturarv. En forudsætning for at ordbøgerne kan udfylde denne funktion er imidlertid at de er kvalitetssikrede og indholdsmæssigt opdaterede. Derudover må de være tilgængelige i former som er i overensstemmelse med brugernes behov og virkelighed.

Nordiske ordbogsressurser oplever aktuelt en række udfordringer i denne forbindelse: Samtidsordbøger er udsat for stor konkurrence fra ikke-kvalitetssikrede webordbøger og fra applikationer der baserer sig på og leverer maskingenereret tekst og indhold. Samtidig forventes det at ordbøgerne ikke kun leverer korrekt information til fx skolesystemet og andre sektorer og brugergrupper, men også løbende opdateres med hensyn til repræsentation og fremstilling af samfundenes kulturer og minoriteter. Historiske ordbøger med omfattende kulturarvsoplysninger er især udfordret med hensyn til omstillingen til moderne teknologier og tilgængeliggørelse for en bredere brugerkreds.

2

Nordisk Forening for Leksikografi og tidsskriftet LexicoNordica indkalder hermed forslag til foredrag ved 32. LexicoNordica-symposium som kan belyse denne problematik fra forskellige synsvinkler. Vi er interesserede i foredrag inden for følgende delemner:

- Udfordringer og løsninger med hensyn til fortsat anerkendelse, sikring og finansiering af samtidsordbøgerne som vidensbaserede og kvalitetssikrede sprogresurser – i en verden med øget konkurrence fra ikke-kvalitetssikrede webordbøger og fra applikationer der baserer sig på og leverer maskingenereret tekst og indhold.
- Hvilke – nye – krav stilles der til samtidsordbøgernes indhold og brugergrænseflade? Hvilke – nye – samfundsrelaterede forventninger mødes ordbøgerne med med hensyn til repræsentation af bestemte samfundsdomæner, samfundsgrupper eller minoriteter?
- Tilgængeliggørelse og revitalisering af ældre, (kultur)historiske ordbøger med uvurderlig nordisk kulturarv. Der ønskes eksempler på overordnede, gennemgribende planer og forløb, men også eksempler på konkrete, detaljerede ændringer og justeringer, fx med input fra eller i dialog med forskellige brugergrupper.
- Store, historiske dokumentationsordbøger har traditionelt henvendt sig til forskersamfundet og særligt interesserede, men med digitalisering og webpublicering har ordbøgerne fået et nyt og markant bredere publikum. Hvordan tænkes disse nye brugere ind i revision og publicering samtidig med at det specialiserede publikums behov

fortsat tilgodeses? Hvordan udnyttes mere tidssvarende former for formidling uden at faglighed og indholdsmæssige nuancer går tabt?

Symposiet gennemføres i samarbejde med henholdsvis Språkrådet, Oslo (samtidsordbøger) og Svenska Akademiens ordbok (SAOB, historiske ordbøger) og med støtte fra Nordplus Nordens Sprog.

## Språkrådet



SVENSKA AKADEMIENS  
ORDBOK



\*\*\*

## Revisjon av Bokmålsordboka og Nynorskordboka

**Bente Selback** (Universitetet i Bergen, [Bente.Selback@uib.no](mailto:Bente.Selback@uib.no)),

**Klara Sjø** (Universitetet i Bergen, [Klara.Sjo@uib.no](mailto:Klara.Sjo@uib.no)),

**Sturla Berg-Olsen** (Språkrådet, Noreg, [sturla.berg-olsen@sprakradet.no](mailto:sturla.berg-olsen@sprakradet.no))

**Knut E. Karlsen** (Språkrådet, Noreg, [knut.karlsen@sprakradet.no](mailto:knut.karlsen@sprakradet.no))

*Bokmålsordboka* og *Nynorskordboka* viser offisiell skrive- og bøyingsmåte for de norske skriftspråkene og kom ut for første gang på papir i 1986 og på nett i 1994. Kulturdepartementet innvilget i 2017 en søknad fra Språkrådet om midler til å gjennomføre en gjennomgående revisjon av *Bokmålsordboka* og *Nynorskordboka* i samarbeid med Universitetet i Bergen. Søknaden inneholdt en relativt inngående beskrivelse av hva revisjonen skulle omfatte, kildegrunnlag, tidsestimat, ressurser, kostnader med mer. Beskrivelsen i prosjektsøknaden har rimeligvis vært styrende for gjennomføringen av revisjonen, men i et så omfattende og langvarig prosjekt har det vært nødvendig å gjøre tilpasninger og justeringer underveis. Overordnede mål for revisjonen har vært å oppdatere innholdet i ordbøkene og legge til rette for nye brukergrupper med tanke på alder, utdanningsbakgrunn og språkbakgrunn.

I innlegget gjør vi rede for hva som har gått som planlagt, og hvilke uforutsette faktorer som har dukket opp. På grunnlag av erfaringene kan vi si noe om hvilke sider av et revisjonsprosjekt det er vanskelig å planlegge for i detalj. Vi vil presentere hvilken tilstand ordbøkene var i da prosjektet ble satt i gang, og hvordan redaksjonen i tett samarbeid med Språkrådet har arbeidet for å oppdatere og videreutvikle ordbøkene og gjøre dem bedre tilpasset en bredere brukergruppe.

Revisjonen startet i 2018 og gikk frem til 2024. I revisjonen har redaksjonen vært i dialog med mange brukere, både i form av skriftlige henvendelser, svar på flere brukerundersøkelser og møter med ulike interessegrupper. I innlegget viser vi eksempler på hvordan brukerkontakten har påvirket selve innholdet i ordbøkene. Vi har gjort endringer for å møte forventninger i samtiden til hvordan minoriteter skal omtales, og for å bedre representativiteten og kjønnsbalansen i brukseksempelene.

I løpet av prosjektperioden ble det klart at det var nødvendig å bygge opp en helt ny dataarkitektur, utvikle en ny visningsside og gjøre endringer i innhold og presentasjon tilpasset den nye og utvidede målgruppen. Vi har også brukt tilbakemeldingene i arbeidet med et brukergrensesnitt som fyller kravene til universell utforming.

Fordi ordbøkene er normative, er bøyingsinformasjon viktig og mye etterspurt, og mye plass er satt av til å vise slik informasjon. Dette er også et ledd i tilpassingen av ordbøkene til brukere som ikke er vant til å bruke tradisjonelle oppslagsverk. Derfor har vi også valgt å bruke færre grammatiske termer i definisjonene og løse opp forkortelser. Det har dessuten vært viktig å ta med flere og bedre utbygde brukseksempler, helst med bøyde former. Hensynet til de nye brukergruppene har gjort det nødvendig å ta med artikler som tidligere ble ansett som

overflødige, som verbalsubstantiv og adjektiv avledet av verb. Vi har også systematisk lagt til flerordsuttrykk, inkludert fraseverb, refleksiver og faste uttrykk.

Ordbøkene har fått bedre søkefunksjonalitet, med treff i bøyde former og forslag til lignende oppslagsord dersom det ikke finnes treff. Det er også lagt inn forslag til oversettelse mellom nynorsk og bokmål, noe som mange brukere lenge har etterspurt.

\*\*\*

## Revideringen av SAOB – från planering till praktiskt genomförande

**Pär Nilsson** (Svenska Akademiens ordbok, SAOB, [par.nilsson@svenskaakademien.se](mailto:par.nilsson@svenskaakademien.se))

I slutet av år 2023 publicerades band 39 av Svenska Akademiens ordbok (SAOB) i tryckt och digital form, och med det är första upplagan av Nordens mest omfattande ordboksverk färdigställt.

Den första delen av ordbokens allra första band, innehållande ord från *A-afbild*, gick i tryck 1893, och sedan dess har redaktionen (eller rättare sagt, de genom åren olika redaktionerna) arbetat sig alfabetiskt framåt utan att revidera eller uppdatera tidigare publicerade beskrivningar i ordboken. Det innebär att de äldsta artiklarna i SAOB i skrivande stund är mer än 100 år gamla, och att mycket av ordbokens innehåll är i stort behov av en uppdatering; metaspråket har i flera avseenden blivit föråldrat och i lemmalistan saknas ett flertal ord eftersom de etablerades i språket efter att det aktuella alfabetiska spann av ord till vilket de hör redan hade publicerats i ordboken. Men inte bara lemmalistan och olika lexikografiska formuleringar vittnar om att tiden har gått, verket innehåller också exempel på stereotypa och icke-inkluderande beskrivningar som i en modern kontext inte är lämpliga. Dessutom har tiden som gått i många fall medfört att artiklarnas sakinhåll inte längre stämmer eftersom samhället, den tekniska utvecklingen osv. har förändrats sedan definitionerna formulerades. Behovet av revidering är således uppenbart, och redan flera år innan den första upplagan färdigställdes inleddes ett omfattande planeringsarbete för att komma till rätta med vad som behöver åtgärdas och hur.

6

I januari 2024 inleddes ett revideringsprojekt på SAOB-redaktionen i Lund, och planeringsarbetets resultat kan därmed börja omsättas i praktiken. Arbetet är organiserat i tre olika övergripande uppgifter: 1) lemmalistan ska kompletteras med nya ord, 2) systematiska förändringar ska utföras i flera eller samtliga artiklar samtidigt, och 3) existerande artiklar behöver ses över var och en och moderniseras manuellt. Dessutom har ett nytt redigeringsverktyg köpts in som anpassas och justeras efter ordbokens struktur och lägger grunden för ett nytt användargränssnitt och möjliga nya sökfunktioner. Under föredraget kommer detta pågående revideringsarbete att beskrivas närmare, och de utmaningar kommer att diskuteras som har att göra med anpassningen av ett så omfattande verk från en helt annan tid till en modern kontext.

SAOB beskrivs ofta med hjälp av olika dikotomier; det är en ordbok med en tvådelad mottagargrupp – den riktar sig både till allmänt ord- och språkintresserade användare och till forskare och experter. Den innehåller både allmänspråkliga ord och betydelser, men samtidigt många (framför allt äldre) specifika facktermer. Ordboken har vidare både ett historiskt perspektiv och ett modernt samtidsperspektiv, och relationen mellan dessa båda perspektiv

behöver ofta skrivas fram och förklaras, vilket kan innebära ganska långa och utförliga resonemang och beskrivningar i ordbokens definitioner.

Ett av skälen till den utdragna publiceringsperioden för ordbokens första upplaga var den vetenskapliga akribin. Hela det material som fanns till förfogande skulle analyseras, och varje form, betydelse och användning av orden som identifierades och beskrevs belades med språkprov. Det beskrivna materialet samlades också in med stor noggrannhet, med målet att nå en så god språklig representativitet som möjligt. Dessutom etablerades en omfattande granskningsprocedur av de producerade ordboksartiklarna, med kollegiala läsningar i flera omgångar.

Det aktuella revideringsprojektet har finansiering för ca 7 år – något helt annat än den första upplagens tidshorisont på mer än 100 år. Utgångspunkten för projektet kan därför inte vara att revidera allt som går eller allt som är i behov av förbättring, utan istället att prioritera, göra så mycket som möjligt och att ta itu med de mest akuta bristerna först. Upplägget kring dessa prioriteringar kommer också att utvecklas under föredraget, med fokus på balansen mellan vetenskaplig noggrannhet och ett effektivt arbete med anständig publiceringstakt.

\*\*\*

## *Fremtidens ordbøger: udfordringer og tilpasninger*

Lars Trap-Jensen (Det Danske Sprog- og Litteraturselskab, DSL, [ltj@dsl.dk](mailto:ltj@dsl.dk))

Ordbogen som genre er i disse år udfordret af den teknologiske udvikling, og det er tvivlsomt om den i fremtiden overhovedet vil eksistere som andet end et nicheprodukt. Historisk set er ordbøger blevet opfattet og brugt både som praktiske redskaber til at løse sproglige problemer med og som værker til dokumentation af sprogets tilstand og udvikling. Hvis man især fokuserer på ordbogens funktion som sprogligt værktøj, kan man godt forestille sig andre og mere effektive måder at formidle oplysninger om eksempelvis et ords betydning i en konkret sætning end ved at lade brugere slå op i en ekstern ordbogsressource. Den integrerede ordbog i e-bogslæsere er et eksempel på dette, og brugerne bliver især tilfredse hvis de kan nøjes med at få vist den forklaring der er relevant i den aktuelle kontekst.

Man kan dog også forestille sig situationer hvor en samlet fremstilling af et ords betydninger og anvendelser fortsat vil være nyttig, for eksempel for fremmedsprogs elever eller for sprogforskere og -studerende. Her vil der dog være tale om mere perifere brugssituationer, og ordbogsredaktionerne bliver nødt til at organisere deres data på nye måder så de kan anvendes på andre måder end i den traditionelle ordbogsform.

I mit oplæg vil jeg på baggrund af en analyse af forskellige brugssituationer og brugerbehov diskutere hvilke tilpasninger af data der er nødvendige for at ordbogsoplysninger kan klare sig i konkurrencen med andre teknologier. Herunder indgår ikke mindst hvad fremkomsten af store sprogmodeller og chatbotter vil betyde for ordbøgerne, og hvordan de bør forholde sig til denne udvikling. Selvom teknologien er forholdsvis ny, findes der allerede en del erfaringer med og teoretiske overvejelser om, samt ikke mindst holdninger til både perspektiver, udfordringer og potentielle bekymringer. Spørgsmålet er om den kunstige intelligens skal opfattes som en konkurrent til den menneskelige redaktør, eller om teknologien kan integreres i det redaktionelle arbejde og føre til bedre sprogbeskrivelser.

8

de Schryver, G.-M. (2023). Generative AI and Lexicography: The Current State of the Art Using ChatGPT. In *International Journal of Lexicography*, 36(4), pp. 355-387.

Trap-Jensen, L. (2024). The Best of Two Worlds: Exploring the Synergy between Human Expertise and AI in Lexicography. In Tinatin Margalitadze (ed.), *Proceedings of the 1 International Conference Lexicography in the XXI Century*. Centre for Lexicography and Language Technologies Iliia State University.  
<[https://lexicography21.iliauni.edu.ge/wp-content/uploads/2024/06/03\\_Lars-Trap-Jensen.pdf](https://lexicography21.iliauni.edu.ge/wp-content/uploads/2024/06/03_Lars-Trap-Jensen.pdf)>

Trap-Jensen, L & Lorentzen, H. (in press). Back to Basics: Meaning Description for Human Users and for Computers. In *Proceedings of the 21st EURALEX International Congress. Lexicography and Semantics*. 8–12 October 2024, Cavtat/Dubrovnik, Croatia.

\*\*\*



## *Kildegrunnlag og presentasjon i Det Norske Akademis ordbok (NAOB)*

**Hanne Lauvstad** (Det Norske Akademis ordbok, NAOB, [hanne.lauvstad@naob.no](mailto:hanne.lauvstad@naob.no))

Autentiske eksempler, sitateksempler, spiller en sentral rolle i en diakron dokumentasjonsordbok over skriftspråket. Denne ordboktypen byr på særlige utfordringer når det gjelder utvalg, behandling og presentasjon av kildematerialet. Det er derfor relevant å se nærmere på forholdet mellom en slik ordbok, i dette tilfellet *Det Norske Akademis ordbok* (NAOB), og dens litterære kilder. Ikke minst er spørsmålet aktuelt i en moderne, digital dokumentasjonsordbok, som har et bredere publikum enn samme type tradisjonelle, papirbaserte ordbøker.

Ett tema er hvilken leksikografisk rolle sitateksemplene spiller, om de f.eks. viser ord og uttrykk i typisk bruk, om de reflekterer forskjeller i betydninger og betydningsutvikling og om de viser variasjoner i det språklige formverket (jf. Svensén 2004). Et annet tema er kildeutvalget, dvs. hvilke forfatterskap, sjangre og litterære perioder som er representert, hvilken fordeling man har mellom mannlige og kvinnelige forfattere, hvordan fordelingen er mellom skjønnlitteratur og sakprosa, inkludert tekster fra aviser og tidsskrifter (jf. Bergenholtz 1996) og i hvilken grad barne- og ungdomslitteratur benyttes som kilder. Også metoder for innsamling av relevant sitatmateriale (ekserpering, digital sanking) og deres fordeler og ulemper vil bli omtalt. Et ytterligere aspekt som vil bli nevnt, er tekstkritiske problemstillinger knyttet til valget av litterært kildegrunnlag og hvilke konsekvenser slike valg kan få (jf. Kraft 1990 og Kondrup 2011).

9

NAOB bygger på *Norsk riksmålsordbok* (NRO), men er blitt en så kraftig utvidelse og modernisering av forelegget at det nå er tale om to forskjellige dokumentasjonsordbøker. Det er derfor nærliggende å sammenligne NAOBs tilnæringsmåte med den som beskrives i innledningene til NRO, både til bind I–IV (Knudsen et al. 1983[1937]) og til supplementsbind V og VI (Noreng 1995). Hensikten er å belyse hvordan utvalget av sitateksempler impliserer endrede holdninger til hva slags litteratur en dokumentasjonsordbok bør bygge på, med andre ord et skiftende litteratursyn. Litteratursynet er påvirket av hvilket publikum ordboken er beregnet for, og dette vil også utdypes.

Til slutt kommer en omtale av noen andre funksjoner sitateksemplene i ordboken kan ha, hvordan de viser frem bruddstykker fra den litterære arven, bl.a. forskjellige forfatteres språkbruk. Alt blir belyst med eksempler fra NAOB. Dessuten vil det bli redegjort for hvordan den leksikografiske presentasjonsmåten har endret seg fra NRO til NAOB.

Bergenholtz, Henning: «Korpusbasert leksikografi». I: *LexicoNordica* 3, 5–18.

Kondrup, Johnny 2011: *Editionsfilologi*. København: Museum Tusulanums Forlag.

Kraft, Herbert 1990: *Editionsphilologie. Mit Beiträgen von Jürgen Gregolin, Wilhelm Ott und Gert Vonhoff. Unter Mitarbeit von Michael Billmann*. Darmstadt: Wissenschaftliche Buchgesellschaft.

Knudsen, Trygve et al. 1983 [1937]: «Innledning». I: *Norsk riksmålsordbok*, bind I, III–IX. Oslo:

Kunnskapsforlaget/Det Norske Akademi for Sprog og Litteratur.

NAOB = *Det Norske Akademis ordbok*. Det Norske Akademi for Språk og Litteratur. <https://www.naob.no> (august 2023).

Noreng, Harald 1995: «Innledning». I: *Norsk riksmålsordbok*, bind V, VII–XII Oslo: Kunnskapsforlaget/Det Norske Akademi for Sprog og Litteratur.

NRO = Knudsen, Trygve et al. 1983 [1937] og Noreng, Harald 1995: *Norsk riksmålsordbok*, bind I–VI. Oslo: Kunnskapsforlaget/Det Norske Akademi for Sprog og Litteratur.

Svensén, Bo 2004: *Handbok i leksikografi. Ordböcker och ordboksarbete i teori och praktik*. 2 utgave, 50–95. Stockholm: Norstedts Akademiska Förlag.

\*\*\*

## Kjønnsrepresentasjon i Bokmålsordboka (BOB) og Nynorskordbokas (NOB) brukseksempel

**Emma Josefin Ölander Aadland** (Universitetsbiblioteket, Universitetet i Bergen, [EmmaAadland@uib.no](mailto:EmmaAadland@uib.no))

Jeg ønsker å presentere en leksikografisk pilotstudie med et kjønnskritisk og komparativt perspektiv på to norske samtidsordbøker. Pilotstudien utgjør en forløper til en større studie, og undersøker hvordan kjønn fremstilles og er representert i ordbokartiklers brukseksempel i *Nynorskordboka* (NOB) og *Bokmålsordboka* (BOB) før og etter den pågående revisjonen av ordbøkene som utføres i perioden 2018–2024 i *Revisjonsprosjektet* («Revisjonsprosjektet», u.å.). Pilotstudien gir mulighet for å teste den metodiske fremgangsmåten, og er derfor avgrenset til et snevrere lemmautvalg for begge ordbøkene. Den tar utgangspunkt i å besvare følgende forskningsspørsmål:

1. Hvordan representeres og fremstilles kjønn i brukseksempel i ordbokartikler i *Bokmålsordboka* (BOB) og *Nynorskordboka* (NOB)?
2. I hvilken grad har revideringen av ordbøkene medført en mer balansert og representativ fremstilling av kjønn i brukseksempelene?

Den metodiske fremgangsmåten er følgende: Brukseksempel før og etter revisjonsarbeidet for det avgrensede lemmautvalget hentes ut fra ordbokdatabasen. Deretter gjennomføres en syntaktisk analyse ved bruk av den automatiske grammatikkanalysatoren for bokmål og nynorsk, NorGram (Dyvik mfl. 2016), som er innlemmet i det korpuslingvistiske verktøyet INESS (<https://clarino.uib.no/iness/>; Rosén mfl. 2012). Av interesse er å undersøke syntaktiske relasjoner med tanke på kjønn, hvem som gjør hva i brukseksempelene, bruken av aktiv og passiv fordelt på kjønn, adjektivbruk knyttet til kjønn, samt i hvilken grad kjønn brukes, er utelatt eller nøytralt i brukseksempelene.

Tidligere leksikografiske studier av ordbøker med et kjønnskritisk perspektiv har undersøkt ulike skandinaviske ordbøker både enkeltvis og komparativt (se f.eks. Hageberg 1990; Nikula 1997; Fjeld 2001; Hult 2010; Larsson 2017; Sköldberg 2020; 2022). Disse påviser blant annet en subtil kjønnsdiskriminering som gjenspeiles i at lemmautvalget for mer tradisjonelle kvinnedominerte interesser er mangelfulle, og videre ved at brukseksempler i ordbøkene ofte gir en stereotypisk fremstilling av kjønn (Sköldberg 2020, 11–13). Min studie kan frembringe ny kunnskap om språkbruk og kjønn i to norske samtidsordbøker, og kan dermed være bevisstgjørende for det videre redaksjonelle ordboksarbeidet i årene som kommer.

Dyvik, Helge, Paul Meurer, Victoria Rosén, Koenraad De Smedt, Petter Haugereid, Gyri Smørødal Losnegaard, Gunn Inger Lyse, og Martha Thunes. 2016. «NorGramBank: A 'Deep' Treebank for Norwegian». I *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, redigert av Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Sara Goggi, Marko Grobelnik, Bente Maegaard, Joseph Mariani, mfl., 3555–62. Portorož, Slovenia: European Language Resources Association (ELRA). <https://aclanthology.org/L16-1565>.

Fjeld, Ruth Vatvedt. 2001. «Ingenting å snakke om - kvinneord og ordbøker». *Språknytt* 29 (4): 10–13.

Hageberg, Arnbjørg. 1990. «Kvinner og menn og ordbøker». *Norsk Lingvistisk Tidsskrift* 8:48–67.

- Hult, Ann-Kristin. 2010. «Kort och gott. Om idiomens språkprov i Svensk ordbok utgiven av Svenska Akademien (2009)». I *Nordiska studier i lexikografi 10. Rapport från Konferensen om lexikografi i Norden, Tammerfors 3-5 juni 2009*, 209–222.
- Larsson, Lennart. 2017. «Vilken plats hade kvinnan?: Om grammatiskt och sexuellt genus i SAOL:s sjunde upplaga 1900». *Språk och stil (Uppsala, Sweden : 1991)* NF 27:49–66.
- Nikula, Kristina. 1997. «Språkprov i könsperspektiv». Redigerat av Rune mfl. Ingo. *Fackspråk och översättningsteori. VAKKI-symposium XVII*, 194–206.
- «Revisjonsprosjektet». u.å. Universitetet i Bergen. Åpnet 30. august 2024.  
<https://www.uib.no/le/121561/revisjonsprosjektet-bokmålsordboka-og-nynorskordboka>.
- Rosén, Victoria, Koenraad De Smedt, Paul Meurer, og Helge Dyvik. 2012. «An Open Infrastructure for Advanced Treebanking». I *META-RESEARCH Workshop on Advanced Treebanking at LREC2012*, redigert av Jan Hajič, Koenraad De Smedt, Marko Tadić, og António Branco, 22–29. Istanbul, Turkey.
- Sköldberg, Emma. 2020. «En Mer Inkluderande Lexikografi – Bortom Frågan Om Kvinnor Och Män». *Oslo Studies in Language* 11 (1): 7-. <https://doi.org/10.5617/osla.8165>.
- . 2022. «Andra upplagan av Svensk ordbok: förutsättningar och redaktionella val». *LexicoNordica*, nr. 29 (november). <https://tidsskrift.dk/lexn/article/view/134779>.

\*\*\*

## Framtidssäkrade samhällsord

**Magnus Ahltop** (Språkrådet, Institutet för språk och folkminnen (Isof), [magnus.ahltop@isof.se](mailto:magnus.ahltop@isof.se))

**Laila Höglund** (Språkrådet, Institutet för språk och folkminnen (Isof), [Laila.Hoglund@isof.se](mailto:Laila.Hoglund@isof.se))

Världen har förändrats sedan Lexinseriens tillkomst i början av 1980-talet och med den användarnas behov. Dagens informationssamhälle ställer krav på avancerad litteracitet och längre utbildning krävs för att komma in på arbetsmarknaden.

Samhällets utveckling driver också fram nyare teknik. Bruket av ordböcker har förändrats avsevärt sedan smarta mobiltelefoner introducerades. I dagens digitala värld går ordboks försäljningen stadigt nedåt – vi slår hellre upp ord på datorn, läsplattan eller i mobilen.

Även språkbruket förändras. Genom resande, fri rörlighet och utbyte av varor och tjänster i Europa samt via sociala medier tar vi del av intryck, varor och nya vanor från hela världen och med dessa följer nya ord och uttryck i snabb takt. I dagens Sverige finns också en större medvetenhet kring social rättvisa, vilket påverkar språket.

Åtskilligt har gjorts för att bibehålla Lexinlexikonens relevans, aktualitet och användaranpassning, senast genom den uppdatering och genomgripande metalexikografiska förbättring av den svenska orddatabasen som genomfördes vid Göteborgs universitet 2008–2010 under Sven-Göran Malmgrens ledning.

De s.k. *samhällsorden* i Lexin har också uppdaterats sedan de först sammanställdes, men någon genomgripande översyn och revidering har inte gjorts på senare tid. Samhällsorden rör huvudsakligen offentlig sektor, som *förskola*, *kommunalval*, *moms*, samt kulturspecifika företeelser som *fredagsmys* och *allemansrätt*. Dessa är särskilt relevanta att lära sig tidigt för att lära känna och bli en del av det nya samhället (Loenheim och Hult, 2020).

Vi ämnar revidera och uppdatera samhällsorden i Lexin. Dessa är inte uppmärskade i Lexins orddatabas eller angivna i de xml-format som används. Därför bör vi först identifiera vilka samhällsorden är i dag, därefter gallra ut inaktuella och tillföra aktuella nya ord. Samhällsinstitutioner, kulturföreteelser och lagar i urvalet bör vara aktuella och relevanta för Lexin användarna.

Vi avser även kategorisera och märka samhällsorden för att möjliggöra automatiserade ändringar, och därmed smidigare och effektivare uppdateringar i framtiden.

Samhällsförändringar driver teknikutvecklingen, men skapar även nya förväntningar och krav. Exempelvis fungerar metalexikografiska anpassningar för datorskärmar inte optimalt i mobila gränssnitt, och somliga krav på användaranpassning står i motsatsförhållande till andra. Vi måste hantera och förhålla oss till det, samt även dra nytta av det. Ett annat viktigt gränssnitt är öppna data, inte mot slutanvändarna, men mot mellanleverantörer. Öppna data möjliggör en större spridning och flexibilitet, men de måste vara uppdaterade för att kvalitetssäkra de applikationer som görs av andra.

Av särskild vikt är att den kategori av lemman som kallas samhällsord är aktuell och relevant för att underlätta för andraspråksinlärare att orientera sig i det svenska samhället. Deras specifika lexicografiska behov inkluderar ett urval och förklaringar av ord som rör (relevanta) lagar, samhällsinstitutioner och kulturfenomen (Tarp, 2008). Argument har också framförts för en stramare definition av samhällsord i Lexin (bl.a. Dahl, 1999).

Dahl, Adolf (1999). Lexinlexikonen och översättaren. *LexicoNordica* 6, 1999, s. 47-53.

Loenheim, Lisa och Hult, Ann-Kristin (2020). *Framtidens Lexin – forskningsöversikt*. Institutet för språk och folkminnen.

Tarp, Sven (2008). *Lexicography in the Borderland between Knowledge and Non-Knowledge. General Lexicographical Theory with Particular Focus on Learner's Lexicography*. Lexicographica Series Maior 34. Tübingen: Max Niemeyer Verlag.

\*\*\*

## *Den Danske Begrebsordbog – opdatering, udvikling og tilgængeliggørelse*

**Sanni Nimb** (Det Danske Sprog- og Litteraturselskab, København, [sn@dsl.dk](mailto:sn@dsl.dk))

**Agnes Aggergaard Mikkelsen** (Det Danske Sprog- og Litteraturselskab, København, [aam@dsl.dk](mailto:aam@dsl.dk))

*Den Danske Begrebsordbog* (Nimb et al. 2014) er hidtil kun udkommet i sin helhed i trykt udgave. I 2025 udkommer det reviderede og opdaterede manuskript digitalt i to forskellige versioner. Den ene er en netversion målrettet forskere, som lanceres i januar 2025. Den giver adgang til udvælgelse af større eller mindre grupper af semantisk beslægtede ord der efter eget valg kan suppleres med oplysninger fra bl.a. *Den Danske Ordbog* (DDO). Data kan downloades og anvendes til videre undersøgelser og eksperimenter. Den anden udgave er en offentligt tilgængelig online begrebsordbog, planlagt til at udkomme sidst i 2025.

*Den Danske Begrebsordbog* indeholder lemmaer, flerordsudtryk og betydninger fra DDO, der dækker dansk sprog fra ca. 1950 til i dag. Ordforrådet er i *Begrebsordbogen* inddelt i 888 nummererede og navngivne afsnit (fx 10.31 *Vrede* og 17.20 *Vintersport*) under 22 navngivne kapitler (fx 10. *Følelser* og 17. *Sport og fritid*). I de enkelte afsnit er ordene ikke listet alfabetisk inden for ordklasserne, men derimod ud fra deres betydning. Samtidig er de koblet til betydningsnummeret i DDO og dermed indirekte også til en række formelle semantiske ressourcer for dansk, herunder *COR.SEM* ([ordregister.dk](http://ordregister.dk)). Sammenkoblingen og den semantiske rækkefølge har fx muliggjort automatisk visning af grupper af nærsynonymer ved de fleste betydninger i online DDO (Nimb et al. 2018) samt udvikling af benchmark-datasæt til måling af kvaliteten af sprogmodellens håndtering af dansk (Pedersen et al. 2024).

15

I oplægget vil jeg beskrive de generelle revideringer der er foretaget i *Begrebsordbogens* leksikalske indhold gennem de sidste 10 år i forhold til den første trykte udgave fra 2014. Jeg vil komme ind på behovet for at ændre visse afsnitsbetegnelser og fremhævede nøgleord der i dag forekommer umoderne. Så godt som alle afsnit er suppleret med en lang række lemmaer og betydninger fra DDO, i alt ca. 25.000 tilføjelser, inkl. ord og betydninger der løbende er integreret i DDO i den samme 10-års periode. Også en række lemmaer og udtryk der endnu ikke er beskrevet i DDO, er tilføjet, vel at mærke ud fra samme krav til korpusfrekvens som anvendes i DDO-redigeringen. Mange af de oprindelige ord er samtidig opdateret med nye sprogbrugsmarkører; oplysninger der efterfølgende vil kunne indgå som materiale i DDO-redigeringen. *Begrebsordbogen* danner således på mange måder et godt udgangspunkt for undersøgelser af sprogudviklingen i dansk i de senere år.

De enkelte afsnit i den digitale forskningsudgave suppleres derudover med helt nye emnemarkeringer der nogle gange samler afsnit på tværs af kapitlerne i *Begrebsordbogen*. Oplysningerne er udarbejdet med inspiration i de overordnede emner som man har opstillet i det engelske projekt *Historical Thesaurus* (<https://ht.ac.uk/>), herunder også i projektets opnåede viden om hvilken del af ordforrådet der ikke er velegnet til emneidentifikation fordi det er repræsenteret ligeligt i alle tekster. Formålet med de nye emneoplysninger er dels at forbedre grundlaget for at arbejde emnebaseret med ordforrådet i DSL's historiske danske ordbøger, dels at undersøge om de kan anvendes til automatisk identificering af emner i et

nyetableret tekstkorpus på DSL (som udgør en del af korpusset *Bakspejlet*) bestående af DSL-udgaven af forfatteren Peter Seebergs samlede værker, der alle er udgivet inden for DDO's periode.

Bakspejlet = DDO's interne korpus, Det Danske Sprog og Litteraturselskab (2024).

*Den Danske Begrebsordbog* (2014). Sanni Nimb, Henrik Lorentzen, Liisa Theilgaard & Thomas Troelsgård. København/Odense: Det Danske Sprog- og Litteraturselskab og Syddansk Universitetsforlag.

DDO (2003-05) = *Den Danske Ordbog*. København: Det Danske Sprog- og Litteraturselskab og Gyldendal.

DDO online = *Den Danske Ordbog*. Det Danske Sprog- og Litteraturselskab. [www.ordnet.dk/ddo](http://www.ordnet.dk/ddo) (april 2024).

Nimb, Sanni, Nikolai Hartvig Sørensen & Thomas Troelsgård (2018): From Standalone Thesaurus to Integrated Related Words in The Danish Dictionary. I: Jaka Čibej, Vojko Gorjanc, Iztok Kosem & Simon Krek (eds.): *Proceedings of the XVIII EURALEX International Congress: Lexicography in Global Contexts*. Ljubljana: Ljubljana University Press, Faculty of Arts. 916-923. <[euralex.org/publications/from-standalone-thesaurus-to-integrated-related-words-in-the-danish-dictionary](http://euralex.org/publications/from-standalone-thesaurus-to-integrated-related-words-in-the-danish-dictionary)>

Pedersen, Bolette Sandford, Nathalie C. Hau Sørensen, Sussi Olsen, Sanni Nimb (2024): Evaluering af sprogforståelsen i danske sprogmodeller - med udgangspunkt i semantiske ordbøger. I: *NyS 65* (<https://www.nys.dk/issue/view/11038>)

\*\*\*



## Revitalisering af Íslensk orðsifjabók (islandsk etymologisk ordbog)

**Trausti Dagsson** (Árni Magnússon-institutet för isländska studier, [trausti.dagsson@arnastofnun.is](mailto:trausti.dagsson@arnastofnun.is))

**Ellert Þór Jóhannsson** (Árni Magnússon-institutet för isländska studier, [etj@hi.is](mailto:etj@hi.is))

*Íslensk orðsifjabók (IO)*, kompileret af Ásgeir Blöndal Magnússon og udgivet på tryk i 1989, er en etymologisk ordbog, der udforsker det islandske sprogs rødder og udvikling. Det er den eneste eksisterende etymologiske ordbog med fokus på moderne islandsk. Ordbogen giver detaljerede forklaringer på ordenes oprindelse og har informationer om semantisk udvikling af enkelte ord og orddele samt tilsvarende orddannelser eller beslægtede ord i andre sprog, især de nordiske og germanske sprog. Denne ordbog har vist sig være et vigtigt redskab for sprogforskere, studerende og sprogentusiaster der ønsker at forstå det islandske sprogs historie og dets kulturelle kontekst.

Digitaliseringen af *Íslensk orðsifjabók*, der blev lanceret online for omkring et år siden, markerer en væsentlig forbedring i tilgængeligheden og anvendelsen af denne vigtige ressource. Gennem digitaliseringen er det ikke kun blevet lettere at søge i ordbogens indhold, men også muligt at præsentere de oprindelige data på nye måder. For eksempel har brugeren adgang til visuelle hjælpemidler hvor der opstilles grafiske diagrammer der viser ords indbyrdes relationer. Desuden gør hyperlinks mellem relaterede ord det muligt at navigere hurtigt mellem forskellige opslag og få et overblik over sproglige forbindelser. Brugergænsefladen understreger de forskellige funktioner og gør det bl.a. nemmere for brugeren at få hurtigt overblik over diverse forkortelser.

I vores præsentation diskuterer vi hvordan nye funktioner forbedrer brugervenligheden af ÍO markant og åbner op for at ordbogen kan anvendes i undervisnings- og forskningssammenhænge på nye og innovative måder. Vi ser på hvordan brugeren nu har mulighed for at navigere på en mere dynamisk og effektiv måde i ordbogens indhold. Vi diskuterer også potentielt videreudviklingen af ordbogsmaterialet. En digital udgave kan lettere opdateres og udvides i takt med ny forskning, hvilket gør den til en potentiel dynamisk kilde der kan tilpasses efter behov. Hvad angår ÍO, er videreudviklingen dog hæmmet af problematiske regler om ophavsret der begrænser muligheden for at udvide ordbogen på den optimale måde. Digitaliseringen af *Íslensk orðsifjabók* understreger dog vigtigheden af at bevare og udvikle specialiseret leksikografisk viden i en moderne kontekst, og viser hvordan et statisk ordbogsværk kan revitaliseres med tilføjelsen af forskellige digitale funktioner og innovativ fremstilling af materialet.

\*\*\*

*Gamle historiske sprogarkiver og deres værdi i nutiden*  
– *Anvendelse og formidling af materialet*

**Ásta Svavarsdóttir** (Árni Magnússon-institutet för isländska studier, [asta.svavarsdottir@arnastofnun.is](mailto:asta.svavarsdottir@arnastofnun.is))  
**Helga Hilmisdóttir** (Árni Magnússon-institutet för isländska studier, [helga.hilmisdottir@arnastofnun.is](mailto:helga.hilmisdottir@arnastofnun.is))

På vores institut, Árni Magnússon-institutet for islandske studier, bevares nu omfangsrigt sprogmateriale der blev indsamlet for at skabe grundlag for en historisk ordbog over islandsk. Kernen i materialet er to traditionelle seddelsamlinger. (1) Det såkaldte skriftsprogsarkiv (*Ritmálssafn Orðabókar Háskólans*; ROH) er langt det største med over 2 millioner excerperede belæg fra trykte bøger og håndskrifter i perioden fra det 16. århundrede til nutiden. Indsamlingen begyndte i 1940'erne og fortsatte indtil århundredeskiftet da nye metoder var blevet introduceret ved leksikografisk arbejde. (2) Talesprogsarkivet (*Talmálssafn Orðabókar Háskólans*; TOH) består af udsagn og kommentarer om ord og ordforbindelser, hovedsagelig fra almindelige sprogbrugere. Størstedelen af disse er kommet som reaktion på forespørgsler i et radioprogram om islandsk sprog, der udsendtes siden 1950'erne indtil ca 2000, og lytternes svar gælder fx ordenes semantiske og strukturelle egenskaber, deres brug, udbredelse osv.

Digitalisering af de islandske sprogarkiver startede allerede i 1980'erne, og den har været udført i større eller mindre grad i begge samlingerne. En del af det digitaliserede materiale, især fra ROH, er senere blevet gjort tilgængeligt på webben. I begyndelsen så man computeren og den elektroniske databehandling først og fremmest som et vigtigt værktøj ved registrering, klassificering og redigering af materialet som forberedelse til en trykt ordbog. Det viste sig dog med tiden at ny teknik påvirkede alle trin af det leksikografiske arbejde, fra indsamling af kildemateriale til publicering af ordbøger. I 1990'erne var der gjort forsøg med redigering af materialet i samlingerne og de viste at det ville være et enormt arbejde at redigere det hele og at en tilfredsstillende beskrivelse af mange lemmaer desuden ville kræve mere materiale. Med tilkomsten af webben blev det desuden muligt at offentliggøre digitaliseret leksikografisk materiale selv om det ikke (endnu?) var en færdiggjort ordbog i nogen forstand og det valgte man at satse på, i det mindste foreløbigt.

Tilkomsten af store og konstant voksende sprogkorporer samt udviklingen af sprogteknologiske metoder og redskaber til automatisk analyse af materialet stiller vigtige spørgsmål om de gamle sprogarkivers værdi i nutiden: Har udviklingen gjort dem forældede eller har de stadigvæk vægt som sprogkilder – og i så fald, hvilken? Hvordan skal digitalisering af samlingerne videreudvikles og i hvilken grad skal materialet gøres (yderligere) tilgængeligt? I hvilken form kan materialet i hver samling bedst præsenteres mht. almene brugere? Har de gamle samlinger og moderne digitale korpusser muligvis forskellige fordele og ulemper som gør begge to vigtige på hver sin måde? Hvilken forbindelse er der (eller skal være) mellem de to? I vores indlæg vil vi fundere over sådanne spørgsmål mht. de islandske sprogarkiver.

*Ritmálssafn Orðabókar Háskólans*. Digital udgave. <https://ritmalssafn.arnastofnun.is/>.