

LEXICONORDICA

LEXICONORDICA

21 · 2014

STORA ORDBÖCKER I NORDEN

NORDISKA FÖRENINGEN FÖR LEXIKOGRAFI

LexicoNordica 21 • 2014
Stora ordböcker i Norden

Huvudredaktörer

Henrik Lorentzen
Emma Sköldberg

Redaktionskommitté

Christian Becker-Christensen
Sturla Berg-Olsen
Annika Karlholm
Mariann Skog-Södersved
Ásta Svavarsdóttir

© 2014 LexicoNordica och författarna

Omslag och sättning: Laurids Kristian Fahl
Tryckt hos: Tarm Bogtryk a-s, Danmark

LexicoNordica trycks med ekonomiskt stöd av
Nordplus Nordiska språk



ISSN 0805-2735

Innehåll

<i>Henrik Lorentzen & Emma Sköldberg</i> Stora ordböcker i Norden	9
--	---

Tematiska bidrag

<i>Øystein Eek & Boye Wangenstein</i> <i>Tanums store rettskrivningsordbok</i> – det norske bokmålets fasit	19
---	----

<i>Jón Hilmar Jónsson</i> Oppbyggingen av en relasjonsbasert islandsk ordboksbase.....	39
--	----

<i>Lennart Larsson</i> En ”mer l. mindre stor” stor ordbok – om variationerna i SAOB:s omfång och ambitionsnivå	61
---	----

<i>Sven-Göran Malmgren</i> <i>Svenska Akademiens ordlista</i> genom 140 år: mot fjortonde upplagan.....	81
---	----

<i>Marita Akhøj Nielsen</i> Moths ordbog og dens virkningshistorie	99
---	----

<i>Christian-Emil S. Ore & Åse Wetås</i> <i>Norsk Ordbok</i> i den digitale tidsalderen	121
--	-----

Klaas Ruppel & Caroline Sandström
Stora finska ordböcker i ett historiskt perspektiv.....141

Lene Schøsler
Blinkenberg og Høybyes *Dansk-Fransk og Fransk-Dansk*
Ordbog – fra seddelsamling til trykte ordbøger og
til onlineudgave.....161

Arnfinn Muruvik Vonen
Behovet for store ordbøker i en samlet norsk
ordbokpolitikk181

Icke-tematiska bidrag

Margrethe Heidemann Andersen, Jørgen Nørby Jensen
& Anita Ågerup Jervelund
Kommentarer til Sven Tarp:
Retskrivningsordbog på godt og ondt 195

Jan Engh
Små ordbøker – mindre og annerledes.....215

Louise Holmer & Emma Sköldberg
Appifiering till allas lycka? Om danska ordboksappar
med särskilt fokus på DDO 235

Recensioner

Anna Helga Hannesdóttir
Bloomsburys vägledning till lexikografin 257

Jon Helgason

Almqvist – en (meta)lexikografisk pionjär 275

Henrik Hovmark

Norsk Ordbok på nettet 289

Hans Landqvist

Pionjärprojekt om svenska översättare 311

Caroline Sandström

Skånsk dialektordbok i behändigt format 331

Rune Svarverud

Nye kvalitetsverktøy for brukere av kinesisk
i Skandinavia 341

Lars S. Vikør

Etymologi på ein ny måte 357

Meddelanden

Henrik Hovmark

Nyt fra bestyrelsen for Nordisk Forening for
Leksikografi 375

Redaktionellt 381

Stora ordböcker i Norden

Henrik Lorentzen & Emma Sköldberg

Härmed presenterar Nordiska föreningen för lexikografi (NFL) det 21:a numret av tidskriften *LexicoNordica*. Som vanligt innehåller tidskriften en tematisk del. De tematiska bidragen bygger på föredrag som hölls vid *LexicoNordica*-symposiet i januari 2014 på Lysebu konferenshotell i Oslo. Vidare innehåller numret en icke-tematisk del, en del med recensioner samt ett avsnitt, författat av föreningens nuvarande ordförande Henrik Hovmark, med information om nästa NFL-konferens, den 13:e i ordningen. Konferensen äger rum i Köpenhamn i slutet av maj 2015. Så till sist, längst bak i boken, tillhandahålls upplysningar om tidskriften och anvisningar till blivande skribenter.

Som framgår av rubriken är årets tema stora ordböcker i Norden. Ordet ”stor” kan i det givna sammanhanget tolkas på olika sätt. En ordbok kan givetvis vara grandios till sitt yttre. Vidare kan en ordbok vara stor genom att den är inflytelserik hos användarna, har stark folklig ställning eller betraktas stå för den allmänt vedertagna normen när det gäller ordförrådet. Men en stor ordbok kan också vara banbrytande i något avseende – antingen i sin samtid eller för senare arbeten. Som redan antytts kan det röra sig om en modern ordbok – såväl en tryckt som en elektronisk – men likaväl om ett äldre verk som haft stor betydelse för, just i detta fall, nordiska länders nationsbygge.

De tematiska bidragen är nio till antalet. Ett av dessa, som författats av Lennart Larsson, rör en riktigt omfattande ordbok, *Svenska Akademiens ordbok* (SAOB 1893–), som hittills kommit ut i 36 volymer. I sin artikel visar Larsson hur ordboken varierat i omfång och ambitionsnivå under de mer än 100 år som den har

givits ut. Ett annat omfångsrikt verk som lyfts fram är *Norsk Ordbok. Ordbok over det norske folkemålet og det nynorske skriftmålet* (1966–). Närmare bestämt diskuterar Christian-Emil S. Ore och Åse Wetås hur denna i första hand tryckta dokumentationsordbok kan vidareutvecklas som digital resurs och hur man kan öka dess bruksvärde bland nya användare i framtiden. Vidare presenterar Jón Hilmar Jónsson en omfattande relationsbaserad isländsk ordboksbas. Författaren ger bl.a. en översikt över utvecklingen av detta betydande arbete, hur enskilda komponenter i ordboksbasen är relaterade till varandra och hur de samspelar i en nätbaserad ordbok.

Ännu en välkänd ordbok som belyses i tidskriftens tematiska del är *Tanums store rettskrivningsordbok*. Det pågående arbetet med en ny utgåva av denna ordbok, som kan betraktas som standardreferensen – eller ”facit” – för bokmål, presenteras av Øystein Eek och Boye Wangensteen. Den svenska motsvarigheten till detta norska verk är *Svenska Akademiens ordlista*, SAOL. I sitt bidrag ger Sven-Göran Malmgren en kort historik över SAOL och hur ordlistan har sett ut under de 140 år som den har kommit ut. Han redogör också för de principiella nyheter som kommer att införas i en kommande fjortonde upplagan av boken.

Samtliga verk som nämnts är enspråkiga ordböcker men givetvis finns det också tongivande tvåspråkiga ordböcker i Norden. Bland dessa återfinns Blinkenberg och Høybyes *Dansk-Fransk* och *Fransk-Dansk Ordbog*. I sitt bidrag kastar Lene Schøsler bl.a. ljus på de lexikografiska principer som kännetecknar verken och diskuterar några av de svårigheter som dykt upp i samband med överföringen av innehållet i de tryckta ordböckerna till en elektronisk plattform.

Klaas Ruppel och Caroline Sandström tar i sin artikel ett större grepp och granskar ett flertal större finska ordböcker i ett historiskt perspektiv. Tidsmässigt tecknar de en linje från 1600–1700-talen då de första finska ordböckerna uppstod till den utgivningssituation

som råder i Finland idag. Ännu ett bidrag av historisk karaktär är författat av Marita Akhøj Nielsen som i sin artikel uppmärksammar Mathias Moths 1600-talsordbok. Tidigare förelåg den bara som manuskript, men sedan 2013 är den också tillgänglig online. I sitt bidrag visar Akhøj Nielsen konkret på det stora inflytande som Moths ordbok haft på senare danska ordböcker.

Men lika väl som man ser bakåt är det, ur ett nordiskt lexicografiskt perspektiv, viktigt att fundera över behovet av stora ordböcker i framtiden. Detta gör Arnfinn Muruvik Vonen i sitt ordbokspolitiska bidrag som tar utgångspunkt i norska förhållanden. Framför allt berör han två stora ordboksprojekt, dels det nästan avslutade *Norsk Ordbok 2014*, dels *Bokmålets og riksmålets ordbokverk*, som fortfarande befinner sig på planeringsstadiet.

Som redan antytts innehåller *LexicoNordica* 21 också tre icke-tematiska bidrag. Inledningsvis är det Margrethe Heidemann Andersens, Jørgen Nørby Jensens och Anita Ågerup Jervelunds kommentar till Sven Tarps recension av den senaste utgåvan av *Retskrivningsordbogen*. Tarps kritiska granskning av ordboken publicerades i förra numret av tidskriften. Vidare har Jan Engh skrivit en artikel som, i kontrast till de verk som behandlas i temadelen, rör små ordböcker. Som utgångspunkt för resonemangen tjänar en portugisisk-norsk-portugisisk fickordbok. Avslutningsvis återfinns en artikel av Louise Holmer och Emma Sköldberg som gäller ordboksappar, i synnerhet fyra danska sådana som utarbetats vid Det Danske Sprog- og Litteraturselskab.

Precis som i tidigare nummer har vetenskapliga recensioner en viktig plats i tidskriften. I detta nummer granskar Anna Helga Hannesdóttir en internationell samlingsvolym som rör såväl metalexikografiska frågeställningar som olika verk och projekt. Jon Helgason ser närmare på en nyutkommen textkritisk utgåva av C.J.L. Almqvists ordboksfragment från 1840-talet. Fragmentet, som har betraktats som banbrytande inom svensk lexikografi, innehåller, bortsett från en läsvärd inledning, standardsvenska ord,

dialektala ord och ortnamn. Vidare tittar Henrik Hovmark närmare på nätutgåvan av ovan nämnda *Norsk Ordbok*, som i skrivande stund täcker avsnittet *i-tåvrrut*. Bland annat diskuterar han ordbokssajtens uppbyggnad och design och vilka sökmöjligheter som den erbjuder. Hans Landqvist undersöker i sin tur en i lexikografiska sammanhang förhållandevis okänd digital resurs, *Svenskt översättarlexikon*, som är ett gott stöd för den som söker kunskap om översättare till svenska från äldsta tid till idag. Caroline Sandström granskar en nyskriven svensk dialektordbok i tryckt format. Närmare bestämt redogör den för det talspråkliga ordförrådet hos skånsk landsbygdsbefolkning i slutet av 1800-talet och första hälften av 1900-talet. I denna del recenserar också Rune Svarverud den första upplagan av en kinesisk-svensk ordbok, den första av sitt slag. Ordboken finns i såväl tryckt format som i form av en app. Avslutningsvis synar Lars S. Vikør en omfattningsrik norsk etymologisk ordbok som bryter på traditionen på så sätt att den faktiskt är tematiskt ordnad. Sammanfattningsvis är det både intressant och inspirerande att få en inblick i mångfalden av lexikografiska verk som presenterats på senare tid.

Den avslutande delen av boken rör som sagt själva tidskriften och dess redaktion, som liksom tidigare består av två huvudredaktörer, Henrik Lorentzen (Danmark) och Emma Sköldberg (Sverige) och fem landsredaktörer från de olika nordiska länderna: Christian Becker-Christensen (dansk redaktör), Sturla Berg-Olsen (norsk redaktör), Annika Karlholm (svensk redaktör), Mariann Skog-Södersved (finsk redaktör) samt Ásta Svavarsdóttir (isländsk redaktör). Redaktionen har beslutat att de närmaste symposierna ska behandla följande teman:

2015: Språkliga varieteter i nordiska ordboksresurser

2016: Nordiska ordböcker och pedagogik

Förslag på möjliga föredragshållare – och framtida *LexicoNordica*-teman – tas tacksamt emot av redaktionsmedlemmarna.

Avslutningsvis vill vi rikta ett stort tack till landsredaktörerna för goda insatser och ett väl fungerande lagarbete under det gångna året. Vidare tackar vi Rikke E. Hauge för ovärderlig hjälp i samband med symposiet på Lysebu i januari 2014. Vi är också mycket tacksamma för det noggranna arbete som Laurids Kristian Fahl har lagt ner vid sättningen av *LexicoNordica* och hans insatser i samband med distributionen av detta nummer. Vi vill även tacka Fondet for dansk-norsk samarbeid, som ännu en gång gjort det möjligt att genomföra symposiet till ett förmånligt pris, och Nordplus Nordiska språk, som beviljat ekonomiskt stöd till såväl symposium som till tryckning av tidskriften. Givetvis vill vi också tacka NFL:s styrelse för vårt goda samarbete under det år som gått.

Litteratur

- Blinkenberg, Andreas & Poul Høybye (1991): *Dansk-fransk ordbog I-II*, 4. reviderede udgave. Ved Jens Rasmussen et al. København: Nyt Nordisk Forlag Arnold Busck.
- Blinkenberg, Andreas & Poul Høybye (1997): *Fransk-dansk ordbog*. 3. reviderede udgave. Ved Jens Rasmussen et al. København: Nyt Nordisk Forlag Arnold Busck.
- NO = *Norsk Ordbok. Ordbok over det norske folkemålet og det nynorske skriftmålet*. (1966–). Oslo: Det Norske Samlaget.
- SAOB = *Ordbok över svenska språket utgiven av Svenska Akademien (Svenska Akademiens ordbok)*. (1893–). Lund: Gleerups förlag.
- SAOL = *Svenska Akademiens ordlista*. (2006). 13:e upplagan. Stockholm: Norstedts.

Tanums store rettskrivningsordbok (2005) 9. utgave. Boye Wangensteen (red.). Oslo: Kunnskapsforlaget.

Henrik Lorentzen
seniorredaktør
Det Danske Sprog- og
Litteraturselskab
Christians Brygge 1
DK-1219 København K
hl@dsl.dk

Emma Sköldberg
universitetslektor, docent
Institutionen för svenska språket
Göteborgs universitet
Box 200
SE-405 30 Göteborg
emma.skoeldberg@svenska.gu.se

TEMATISKA BIDRAG

Tanums store rettskrivningsordbok – det norske bokmålets fasit

Øystein Eek & Boye Wangenstein

The authoritative spelling dictionary of Norwegian Bokmål has for 75 years been *Tanums store rettskrivningsordbok*. The manuscript has recently been converted into a simple, but well tagged database. Redesigned digital and printed versions are planned for 2014. The flexible spelling system of Norwegian Bokmål will be clearly visible in the new user interface, with full documentation of the variant form and inflection systems, and with numerous cross-references. The publication might inspire the Language Council of Norway to reconsider some peculiarities of the standardized Bokmål. The editors of the Tanum dictionary also invite interested parties to a discussion about the selection of lemmas in a modern and comprehensive spelling dictionary of a language where new compounds emerge every day.

1. Innledning

Den autoritative rettskrivningsordboken for norsk bokmål har gjennom 75 år vært *Tanums store rettskrivningsordbok* (heretter kalt Tanum). I denne artikkelen vil vi presentere konverteringen av ordbokbasen og tilretteleggingen for ny utgivelse. Konverteringen og datarettingen er det mest omfattende databaseprosjektet som har vært fullført i Kunnskapsforlaget i senere år, i et vellykket samarbeid mellom den danske dataleksikografen Jens Erlandsen, tittelansvarlig ordbokredaktør Boye Wangenstein, Kunnskapsforlagets porteføljeansvarlige redaktør Øystein Eek, samt forlagets IT-utviklere Per Halvor Tryggeseth og Jonas Knutsen.

Prosjektet har en eldre database som grunnmateriale, så det har

ikke vært nødvendig å starte med et bokmanus, men denne databasen har en komplisert elementstruktur og er ikke fullt oppdatert i samsvar med rettskrivningsendringene i 2005. Siste bokutgave har derfor vært en viktig referanse. Hensikten med prosjektet har vært å gi det meget omfattende innholdet i Tanum en så enkel struktur som mulig, likebehandle alle valgfrie former i bokmål, og knytte ordleddene i sammensatte ord til riktig grunnord gjennom lenking. Ved å publisere denne databasen i digital og trykt form vil vi både gi det norske språksamfunnet et forbedret leksikografisk hjelpemiddel og bidra til å rette søkelyset mot en del spesielle forhold ved den norske bokmålsrettskrivningen.

2. Historikk

Tanum utkom første gang i krigsåret 1940. Arbeidet hadde da pågått i ca. fire år, og med et imponerende resultat, når man tenker på at dataverktøy ikke fantes, og at det ikke var en redaksjonell stab til rådighet for redaktørene. Den opprinnelige redaktør var professor Jacob Sverdrup (1881–1938), men han døde før arbeidet var fullført og ble etterfulgt av lektor Marius Sandvei (1905–1993). Rettskrivningsordboken var dermed påbegynt før 1938-rettskrivningen var på plass, men boken ble utgitt som en «bibel» for det nye bokmålet kort tid etter reformen. Den imøtekom tydelig et stort behov, og fikk ikke noen virkelig konkurrent før *Bokmålsordboka* (1986).

Det var utkommet rettskrivningsordbøker i Norge også tidligere, men disse var av langt mindre omfang. Historien kan føres tilbake til 1885, da Aars' *Retskrivningsregler* (som i tillegg inneholdt en ordliste) ble autorisert av departementet. Senere kom flere nye utgaver, men dette var små trykksaker der korte ordlister tjente til eksemplifisering av rettskrivningsreglene, jf. Aars & Hofgaard (1907). Etter språkreformen i 1907 kom de første egentlige ordlister

til skolebruk med offentlig godkjenning. I 1910 utkom Ivar Alnæs' *Norsk uttale-ordbok*, som ikke inneholdt bøyningsformer, men hadde hele 14 000 oppslagsord. Den første litt større ordboken med bøyningsformer og korte forklaringer var S. Juell Tønnessens, som kom i to utgaver, 1910 og 1926, altså før og etter 1917-reformen. Den mest utbredte rettskrivningsordboken med offisielt stempel etter 1917 var trolig Krogsrud & Seip (1924).

Omfanget av Tanum, som opprinnelig het *Norsk rettskrivningsordbok* og ble utgitt av Johan Grundt Tanum, innebar et kvantesprang. Her skulle ikke bare selve rettskrivningen dokumenteres, men også mest mulig av det ordforrådet som kunne registreres gjennom bruk av mange ulike kilder, som redaktør Sandvei gjør rede for i forordet. En rekke ordbøker og andre oppslagsverk er benyttet,

[m]en det var ikke nok med ordbøker. Også alle slags spesialhandbøker er brukt [...]. En rekke store firmaer i ulike bransjer har velvilligst overlatt redaksjonen varekataloger, og et stormagasin som Steen & Strøm A/S har tilmed utarbeidd fortegnelse over alle sine varer for ordboka. Videre har foreninger og institusjoner hjulpet med å skaffe oppgaver over ord som brukes i arbeidslivet i Norge i dag. Dette store tilfang av moderne bruksord, som ellers er vanskelig tilgjengelig, har redaksjonen ekserpert, for at ordboka i særlig grad også skal bli høvelig for ervervslivet. (Sverdrup & Sandvei 1940:III–IV)

Tanum blir dermed også et interessant samfunnshistorisk dokument.

Norsk rettskrivningsordbok fikk påskriften «Godkjent av Kirke- og Undervisningsdepartementet til offentlig bruk og som hjelpebok i skolen». Språkrådets og dermed myndighetenes medvirkning har i senere år begrenset seg til å «se igjennom» reviderte

utgaver for å forvise seg om at normen er dekket på en fullgod måte.

Marius Sandvei samarbeidet fra 1974 med Bernt Fossetøl (1928–88) om redaktøransvaret. Fra 1983 overtok Kunnskapsforlaget ansvaret for utgivelsen, og i 1994 ble Boye Wangensteen (f. 1942) engasjert som redaktør.

Selve konseptet Tanum har bestått uendret i alle årene etter 1940, med bare mindre justeringer av artikkelstrukturen. Mange nye ord er kommet til, mens svært få ord er tatt ut, selv om de etter hvert er blitt lite aktuelle. Ordboken har dermed et noe mer diakront preg enn den opprinnelig hadde. Ord fra jord- og skogbruk, håndverk, industri, sjøfart og samferdsel slik disse næringene fremstod ved utbruddet av den annen verdenskrig, er fortsatt godt representert i 9. utgave fra 2005. I senere år har avis-korpuset ved Universitetet i Bergen (avis.uib.no) vært en hovedkilde for supplering av lemmautvalget. Nyordmaterialet som er utgitt i *Nyord i norsk* (Guttu & Wangensteen 2012), er blitt gjennomgått og brukt som ressurs. Allmennspråket er i prinsippet holdt oppdatert, og fagord som er brukt i allmennspråket likeså.

Den første digitale utgaven av Tanum kom på disketter i 1991, og en CD-ROM-utgave i 1997, begge på Kunnskapsforlaget. Disse digitale utgavene hadde komplettvisning av bøyningsformer, i samsvar med daværende rettskrivning. Da 9. bokutgave skulle utgis så raskt som mulig etter rettskrivningsendringene 2005, var det ikke ressurser til å oppdatere hele databasen med de nye bøyningsformene, så det ble valgt en enklere løsning der de formene som vises i bokutgaven, ble oppdatert, mens resten måtte ligge.

Tanum er omtalt i regjeringens Språkmelding fra 2008:

På bokmålssida representerer *Tanums store rettskrivningsordbok*, også utgjeven av Kunnskapsforlaget, eit godt supplement til Bokmålsordboka. [...] Med i alt rundt 300 000 ord er dette den største og mest omfattande

rettskrivingsordboka for bokmål. Ho er gjennomgått av Språkrådet i tråd med gjeldande rettskriving, har fullstendige opplysningar om alle tillatne former og varianter og er dermed ei sikker kjelde til den offisielle bokmålsrettskrivinga. (Stortingsmelding nr. 35 (2007–2008), kap. 8.4.4.7)

Ordboken har gjennom mange årtier etablert seg som standardreferansen for bokmål for offentlig sektor, for forfattere og journalister, for studenter etc. Den står på utallige kontorer landet over (om enn til dels i gamle utgaver), og boken har også av ordlisteforfattere vært brukt som kilde i tillegg til *Bokmålsordboka*.

3. Offisielle rettskrivningsordbøker i Sverige og Danmark

At Tanum er et forlagsprodukt, men samtidig har en offentlig anerkjent status som referanse for bokmålet, setter utgivelsen i en særstilling i Skandinavia. Dansk Sprognævn har sin egen *Retskrivningsordbog*, som utgis digitalt på deres eget nettsted og som applikasjon (app), som utlisensieres til andre publiseringsportaler etc. mot betaling, og som utgis i bokform av et dansk forlag (for tiden Alinea A/S) mot royalti til Sprognævnet etter anbud. Sprognævnet har en egen redaksjon for rettskrivningsordboken, med redaktører som bruker ca. 50 % av sin arbeidstid på dette prosjektet, med et redaksjonsråd og med kontakt med Sprognævnets representantskap for å få føringer i prinsipielle normerings- og redigerings spørsmål. Siste bokutgave kom i 2012. Knyttet til utgivelsen er en nyordtjeneste på Sprognævnets nettsted.

I Sverige eier Svenska Akademien rettighetene til *Svenska Akademiens ordlista* (SAOL), som er utkommet siden 1874 i 13 utgaver. Den har en redaksjon ved Göteborgs universitet med vide

fullmakter når det gjelder redigeringen. Bokutgaven publiseres av forlaget Norstedts, som har Sveriges eneste gjenværende ordbok-redaksjon i forlagsverdenen, i et langsiktig samarbeid. Det finnes både en enkel nettutgave og en mer sofistikert nedlastingsutgave, samt en app, som utgis i regi av Svenska Akademien selv. Både i Sverige og Danmark er de digitale utgavene gratis for brukerne.

4. Den nye Tanum-basen

Kunnskapsforlaget gjorde i 2011 en avtale med Jens Erlandsens firma EMP om å konvertere den eksisterende basen i programmet GestorLex over til programmet iLex, med en betydelig endret elementstruktur. Planen var så å gjøre datarydding i iLex, og dernest konvertere den nye basen videre til Kunnskapsforlagets ordbok-program Dict, før endelig redigering og utgivelse. Boye Wangenstein og forlagets seniorredaktør Øystein Eek begynte sin del av dataryddingen høsten 2012, og fullførte denne vinteren 2014.

I utgangspunktet synes artiklene i Tanum å ha en enkel struktur, med oppslagsord, uttale ved avvik fra hovedreglene, ordklasse-opplysninger, korte bøyingsformer, stikkordsmessige ordforklaringer, samt utvalgte uttrykk. S sammensatte ord er samlet i såkalte reder under grunnordet, men et rede kan bestå av flere semantisk beslektede homografer, f.eks. substantivet og adjektivet *likvid*.

I den nye utgaven skal alle redene oppløses, slik at det blir enklere å finne oppslagsordene, og alle sammensatte ord skrives fullt ut. Bokutgaven får altså glattalfabetisk makrostruktur. I bokutgaven 2005 kan valgfriheten f.eks. være oppført slik:

laga|sild el. **lage-** el. **lågå-**, -a el. -en; - el. -er, -ene (eg. sild fra Lågen).

Her har redaktørene skilt de tre tillatte stammeformene fra hver-

andre. Ved å kombinere de valgfrie stamme- og bøyningsformene, får man følgende valgmuligheter:

lagasild, lagasilda/lagasilden; lagasild/lagasilder, lagasildene
 lagesild, lagesilda/lagesilden; lagesild/lagesilder, lagesildene
 lågåsild, lågåsilda/lågåsilden; lågåsild/lågåsilder, lågåsildene

Under **loft** finner man en rekke sammensetninger med *loft(s)*, bl.a. *glugg(e)*. Ordet kan altså hete *loftglugg*, *loftsglugg*, *loftglugge* og *loftsglugge*. Bøyningsformene finner man så under **glugg**, der *glugge* er oppført som variantform.

Dette oppsettet er lite egnet for en digital utgave, der både alternative stammeformer og bøyingsformer bør skrives fullt ut. I den gamle databasen finnes fulle bøyingsformer av også sammensatte ord, men de har ikke vært oppdatert. Vi har derfor måttet generere nye bøyingsparadigmer for usammensatte ord ut fra oppdaterte endelser, og koble ordleddene i sammensatte ord til riktig grunnord. Redene i Tanums bokutgave inneholder bare unntaksvis opplysninger om ordklasse og homograftilknytning for sisteledd av sammensatte ord. Det har vært et stort arbeid å foreta denne koblingen. I de tilfellene hvor det var bare ett oppslagsord sisteleddet kunne kobles til (altså ingen homografer), som ved *telegraf* (*bildetelegraf*, *billedtelegraf*, *branntelegraf* osv.), kunne koblingen gjøres maskinelt, men ved homografer måtte de behandles manuelt, som ved *råd* (*ambassaderåd*, *ansettelsesråd* osv.). Dette har vært tilfellet med mer enn 46 000 ord. Også mange førsteleddskoblinger er blitt feilaktige ved konverteringen. Redene i Tanum innledes ofte med grunnord av flere ordklasser med ensartet eller nesten ensartet skrivemåte (f.eks. substantivet *stemme* og verbet *stemme*), og analyseprogrammet har naturlig nok ikke kunnet koble sisteleddene til riktig førsteledd her. Disse redene har derfor måttet gjennomgås manuelt.

At også sisteleddene i sammensatte ord kobles til riktig

homograf, har ikke vært gjort tidligere i noen ordbok over norsk språk, så langt vi kjenner til. Vi mener derfor at arbeidet som nå har vært gjort med Tanum, representerer en leksikografisk nyvinning i Norge. I noen tilfeller har vi vært sterkt i tvil om hva som er mest plausible tolkning, og det er ikke alltid mulig å gi en entydig løsning. Et *profilbilde* kan være både et sidebilde (til *profil m*) og et tverrsnittbilde (til *profil n*).

Sammenlignet med *Bokmålsordboka*, der man får korrekt bøyning av sammensatte ord, får man i nye Tanum også sikker (så langt det er mulig) opplysning om hvilke ordledd ordet er sammensatt av. I tillegg er som nevnt ordmassen langt større.

Også redaksjonen for *Svenska Akademiens ordlista* har valgt en løsning for neste utgave der koblingen til riktig homograf vises, se Malmgren (2014).

I forbindelse med at systemet med hovedformer og sideformer i bokmål ble avskaffet i 2005, ble enkelte mindre brukte ordformer tatt helt ut av rettskrivningen, mens noen sideformer ble opphøyet til hovedformer. Av bøyingsformer fikk mange tidligere sideformer status som hovedformer.

De offisielle ordbøkene skal i utgangspunktet ikke gi noen føringer for hvilke former som er mest utbredt og dermed mest brukbare i språksamfunnet. Fra 1959 har myndighetene lagt til grunn for redigering av godkjenningspliktige ordlister at valgfrie former skal oppføres i alfabetisk rekkefølge. Tanums spesielle status har ført til at dette prinsippet har vært fulgt også her.

En naturlig konsekvens av dette har da vært at vi i datarydningen har etablert krysshenvisninger mellom alle valgfrie stammeformer, og at de samme utfyllende opplysninger har måttet tas med under alle variantene, for ikke å gi føringer for valg av former. Dette prinsippet utløser svært mange henvisninger i basen, pr. mars 2014 er det i alt ca. 340 000. Av disse gjelder ca. 53 000 kryssreferanser mellom ulike variantformer. Fra *lagasild* vil det være henvist til *lagesild* og *lågåsild*, fra *lagesild* til *lagasild* og *lågåsild*, fra

lågåsild til *lagasild* og *lagesild*, altså ved n variantformer $n \cdot (n-1)$ henvisninger. For sammensatte ord med valgfrihet i flere ledd blir antallet varianter og dermed henvisninger ganske høyt. Ved de aller mest kompliserte tilfellene har redaksjonen tatt sjansen på å foreta visse forenklinger, og gjøre rede for disse i en brukerveiledning.

Norsk ordbank ved Universitetet i Oslo – en database med fullformer og bøyingsinformasjon for leksemer i bokmål og nynorsk – er knyttet opp mot *Bokmålsordboka* og *Nynorskordboka* og er av Språkrådet anbefalt som referanse for korrekt rettskrivning. Basen er imidlertid ikke fullt oppdatert, og er omtalt slik i Språkrådets notat til Kulturdepartementet om Norsk ordbokpolitikk (03.03.2014):

Språkrådet har lenge hatt et ønske om å forenkle kontrollarbeidet med skoleordlister ved å bruke rettskrivningsinformasjonen i Norsk ordbank som hjelpemiddel i godkjenninga. Norsk ordbank, nynorsk og bokmål, er nå på vei til å bli godt oppdatert, og artikkelhodene, dvs. oppslagsform, ordklasse og bøyning, kan lastes ned gratis fra Språkbanken. Vi tenker oss at ordlisteprodusenter laster ned Ordbanken og bruker materialet (utvalgte deler av det) som basis for egne skoleordlister. Dermed vil Språkrådet bare trenge å vurdere opplegget for ordlista, og ellers ta stikkprøver. (Språkrådet 2014:23)

Det kunne dermed tenkes at Tanum ikke trengte sitt eget bøyingsverk, men kunne hente dette fra Norsk ordbank. Men ettersom Tanum-basen er mer omfattende enn Norsk ordbank (ca. 209 000 oppslagsord mot ca. 146 000 pr. april 2014, og inneholder mye av den samme informasjonen om oppslagsordene, er denne løsningen ikke aktuell:

	Norsk ordbank	Tanum-basen
Antall oppslagsord	146 342	209 125
Antall oppslagsord med liten forbokstav	139 650	202 950
Antall unike oppslagsord med liten forbokstav	49 707	111 258
Bøyningsformer av usammensatte ord	Ja	Ja
Bøyningsformer av sammensatte ord	Ja	Ja, via lenking til riktig homograf
Ordklasseopplysninger	Ja	Ja
Opplysning om normering	Ja	Ja
Krysshenvisninger til alternativformer	Nei	Ja
Uttale av irregulære ord	Nei	Ja
Referanse til bøyningskategori	Ja	Nei
Opplysning om appellativ/proprium	Ja	Delvis, ved at proprium ikke har ordklasse

Tabell 1: Fakta om Norsk ordbank og Tanum-basen.

5. Publisering av den nye basen

Gjennom mange år har det vært spådd at de trykte ordbøkernes tid er forbi, og at digitale hjelpemidler overtar. Nå er det fortsatt slik at det selges godt over 100 000 trykte ordbøker i Norge hvert år, men markedsandelen går stadig ned. Arbeidet med Tanum-prosjektet har vært basert på ønsket om å gjøre basen enklest mulig å redigere i fremtiden, og om å gjøre en ny digital utgave mest mulig brukervennlig. I en digital utgave på Kunnskapsforlagets digitale språktjeneste Ordnett trenger man ikke å lage et plassøkonomisk brukergrensesnitt slik som i en trykt bok. Forkortelser skal unngås eller gjøres klikkbare, bøyningsformer skal skrives fullt ut, med nødvendige ledetekster, og bøyningsopplysningene kan være dy-

namiske, dvs. at bøyningene av sammensatte ord kan hentes fra det usammensatte sisteleddet. For oppslagsordet *lågåsild* henter programmet bøyningen fra *sild*. Dermed blir det langt mindre sjanse for feil etter hvert som rettskrivningen blir endret.

I den planlagte bokutgaven vil det bli en stor utfordring å presentere alt innholdet mellom to permer. Basens nye innretning, der alle sammensatte ord skrives fullt ut, er langt mer plasskrevende enn rede-strukturen. Samtidig innebærer oppløsningen av redene at leserne mister opplysningen om hvilken av flere homografer førsteleddet i et sammensatt ord er dannet til. Vi må derfor sette inn et tall som viser dette, f.eks. **original²språk**, som viser at førsteleddet er adjektivet *original*, ikke substantivet. Så vil det i bokutgaven heller ikke være plass til å vise bøyningsformene av sammensatte ord fullt ut. Vi må nøye oss med å vise bøyningen av sisteleddet. Det vil altså stå **arboret** [arbor'et] S m/f (*arboretet; arboret/arboreter, arboreta/arboretene*), men **original²språk** S n [-språket; -språk, -språka/-språkene]. En ny utgave av Tanum i bokform vil bli langt mer omfangsrik enn de tidligere på grunn av de nye løsningene.

6. Databasen som ressurs for Kunnskapsforlaget

Den ferdige Tanum-basen gir et inventar av norske ordformer med grammatiske og andre opplysninger som blir stort og pålitelig. Den vil dermed være en viktig ressurs for Kunnskapsforlaget, både som fullformsbase som utnyttes i søk på Ordnett og som ressurs for andre ordbøker. En del av våre kunder etterspør grammatiske opplysninger til norske ekvivalenter i de tospråklige ordbøkene; disse vil nå kunne hentes dynamisk fra Tanum-basen. Ettersom ordmassen i Tanum er så stor, vil den utgjøre et arsenal for supplering av ordmassen i andre ordbøker. Samtidig finnes det ord i andre av våre norske ordbaser som ikke er med i Tanum, så

berikelsen kan også gå den andre veien. Tanum-basen kan også være interessant for utlisensiering til språkteknologiske prosjekter og andre formål. Den har allerede vært stilt til rådighet som ressurs for Iness, et stort språkteknologisk prosjekt ved Universitetet i Bergen.

7. Tanum-basen og bokmålsnormen

Bokmålets doble opprinnelse i det dansk-norske riksmål og i folkemålet og samnorskprosjektet fra 1930-årene ligger til grunn for den brede valgfriheten og det ofte dobbelte eller tredobbelte tilfanget av semantisk overlappende ordformer i bokmål. Riksmålets *lagesild* eksisterer side om side med dialektformene *lagasild* og *lågåsild*. I andre tilfeller er det vanskelig å se språkpolitiske motivasjoner for valgfriheten. Dette gjelder bl.a. bruken av valgfri fuge-s, som vi ser i *loftglugg/loftsglugg/loftglugge/loftsglugge*. Det er lett å se at språkbrukerne over det ganske land, som vil skrive et bokmål de kan personlig identifisere seg med, kan ha glede av valgfriheten i det første tilfellet, men mindre innlysende at dette er relevant i det andre eksempelet.

I Språkmeldingen fra 2008 gir Kulturdepartementet følgende føringer for fremtidig normering:

For bokmålet sitt vedkomande reknar departementet med – trass i at det framleis er etter måten stor valfridom på mange punkt i rettskrivinga – at den oppryddinga og forenklinga som vart godkjend i 2005 har vore tilstrekkeleg til at det så langt fram vi kan sjå, ikkje bør vera trong for større endringar. Dermed skulle det no vera lagt grunnlag for ein relativt stabil rettskrivingssituasjon i bokmål.

Truleg vil det også vera med på å sikra slik stabilitet at variasjonen i faktisk språkbruk i bokmål synest å vera klart

mindre enn den valfridommen som rettskrivinga opnar for.

Utover desse overordna prinsippa ønskjer ikkje departementet å leggja meir detaljerte og spesifikke føringar på ein framtidig normeringspolitikk. Det er viktig at ein slik politikk framleis har ei breiast mogleg språkdemokratisk forankring, og mange av dei normeringsprinsippa som vart utvikla av Norsk språkråd, vil truleg vera eit godt grunnlag å byggja vidare på. (Stortingsmelding nr. 35 (2007–2008), kap. 8.4.2.5)

I Språkrådets vedtekter, § 3, står bl.a. følgende: «Dersom faktisk språkbruk eller sterke systemomsyn tilseier det, kan Språkrådet vedta ny stavemåte og bøyning av tidlegare normerte ord og justeringar i skrivereglane.»

Det er altså god hjemmel for å foreta visse forenklinger, så lenge disse ikke er «store endringer», og det er god hjemmel for å undersøke den faktiske språkbruk som begrunnelse for justeringer. Vi er overbevist om at publiseringen av den nye Tanum-basen kan bidra til en fornyet diskusjon om forenkling av bokmålsnormen, basert på empiriske undersøkelser av de valg språksamfunnet gjør. I 9. utgave presenteres sammensetningen *høyt lønnet* slik under oppslagsordet **høyt** el. **høgt** adv.:

høy(t)- el. **høg(t)** lønna el. løn(ne)t;

Under

høg adj., n. -t; komp. **høgere**, sup. **høgest** el.

høy adj., n. -t; komp. **høyere**, sup. **høyest**;

står det slik:

lønna el. løn(ne)t el. høgtlønna el. høgtløn(ne)t el. høyt-
lønna el. høytløn(ne)t

Begge disse oppføringene kommer midt inne i svært omfattende reder.

Når man legger alt dette sammen, blir det klart for bokmålsbrukeren at dette ordet kan skrives på 12 forskjellige måter, noe som gir $12 \cdot (12-1) = 132$ mulige kryssreferanser.

Hvis vi holder formene med *høg/høgt* utenfor, siden de brukes også på nynorsk, gir et enkelt Google-søk 26.03.2014 på disse formene i anførselstegn og på sider på norsk følgende resultat (inkl. tillatte bøyingsformer):

høytlønnen/-ede/-ete	17 029
høytlønt/-e	7 700
høytlønna	2 600
høylønnen/-ede/-ete	640
høylønt/-e	544
høylønna	58

Tabell 2: Google-treff på varianter av *høytlønnen*.

En rimelig forenkling ville da være å beholde de tre øverste, samt den av formene med *høg* som systematisk passer best med disse, nemlig *høgtlønna*. Alle rimelige språkpolitiske hensyn ville da ha vært ivaretatt, de 12 formene redusert til fire, og antallet krysshensvisninger til 12.

I andre språksamfunn vil selv fire valgfrie skrivemåter av samme ord oppleves som komplisert og kanskje også unødvendig, og det er grunn til å spørre seg om vi gjør språket mer eller mindre robust i det multinasjonale Norge i en situasjon der språket er under sterkt press fra engelsk, ved å opprettholde det hevdvunne systemet av valgfrihet.

8. Planer for videre redigering

I forordet til 1. utgave av Tanum (Sverdrup & Sandvei 1940:IV) står det bl.a. at

Ordboka vil gi opplysning om skrivemåte, kjønn og bøyning av flest mulig grunnord. Særlig er da mange særnorske ord, fagord og fremmedord kommet med. [...] så langt det har vært råd å få det til, [er] alle vanlige sammensetninger og avledninger tatt med.

En kan vel legge til at det er en god del *uvanlige* sammensetninger også, dvs. sjeldne. Videre står det i forordet at

Norsk *rettskrivningsordbok* legger fram et rikholdig ordtilfang uten å ville øve press i noen retning. Ordene står der enten de er nye eller gamle, særnorske eller halvdanske, heimlige eller fremmede. Så får den som bruker boka, selv velge.

Derfor finner vi i Tanum fortsatt mange ord fra f.eks. gammelt arbeidsliv (*brøte* 'velte tømmer i elva og fløte', *bakole* 'rem på seletøy'), fra bergverksdrift (*blotning* 'synlig del av formasjon', *jøpel* 'transportmaskin, hestevandring'), arkaiske ord (*kyrr* 'stille, rolig', *liktrå* 'spedalskhet'), en mengde fremmedord og fagord innenfor arkitektur (*peristyl* 'søylegang'), anatomi (*diafyse* 'mellomstykke på lange knokler'), medisin (*effluvium* 'utflod', *flatus* 'avgang av tarmgass') og fra en rekke andre fagområder.

Vi må stille spørsmålet om Tanum fortsatt skal være en ordsamling der en kan finne likt og ulikt (til glede for f.eks. kryssordlødere og scrabblespillere) eller om den skal strammes inn til en rettskrivningsordbok over dagens aktuelle ordforråd. Videre om

vi skal skille innholdsmessig mellom den elektroniske utgaven og bokutgaven.

En mulig vei å gå er å legge inn i ordmassen i Tanum et stort antall aktuelle ord fra avis-korpuset i Bergen, men ikke ta med alle i bokutgaven, og ta ut av bokutgaven et stort antall uaktuelle ord.

Ovenfor er det opplyst litt om hvordan ordforrådet i Tanum ble samlet i slutten av 1930-årene. Dette var lenge før korpusleksikografiens tid. Et korpusbasert Tanum omkring 1940 ville ha gitt et langt magrere resultat. Uansett hvor stort tekstkorpus man hadde hatt tilgang til, ville mye av det som står i boken, ikke ha kommet med. Samlingen er derfor svært verdifull som et ordarsenal. Ordmassen i Tanum overgår langt det man finner i SAOL (125 000 i 13. utgave fra 2006) og *Retskrivningsordboken* (65 000 i 4. utgave fra 2012; dette tilsvarer antallet i *Bokmålsordboka*). Det er neppe noen som vil argumentere for en sanering som vil fjerne alt historisk og kuriøst ordstoff fra basen, men det er neppe heller noen som vil argumentere for å beholde alt i en fremtidig utgivelse. Her finnes for eksempel oppslagsord som vi som arbeider med redigeringen, ikke forstår betydningen av, og som vi heller ikke kan finne i noen tilgjengelige kilder. Det kan godt hende at slike ord skal beholdes i basen, men ikke publiseres. Språkrådet legger ingen føringer for den fremtidige redigeringen av ordforrådet i Tanum, men nøyer seg med å kreve at skoleordlister må dekke det sentrale ordforrådet i språket for å kunne bli godkjent (Stortingsmelding nr. 35 (2007–2008), kap. 8.4.4.5). Ordbokens redaksjon oppfatter at den har stor frihet til å ta ut dødt språkstoff. Redaksjonene for SAOL og *Retskrivningsordboken* har i senere utgaver gjennomført en betydelig sanering, uten protester. Ved å foreta en skjønnsom sanering og samtidig ta inn et større nyordmateriale, slik de andre skandinaviske redaksjonene også har gjort (for eksempel kom 4500 nye oppslagsord inn i 4. utgave av *Retskrivningsordboken*), vil Tanum få et mer moderne preg.

Vi har gjort oss tanker om Språkrådets rådgivende medvirkning i det fortsatte arbeidet med Tanum. Hovedsporet har vært at ordboken fortsatt skal være en forlagsdrevet autoritativ bokmålsordbok med offentlig anbefaling, i pakt med en langvarig tradisjon (jf. Språkmeldingens kap. 8.4.4.7). Vi har invitert Språkrådets fagråd 3, som har språknormering og språkoobservasjon som sitt arbeidsfelt, til å bli med i en prosess der vi prøver å bli enige om føringer for utvikling av denne språkressursen på kort og lengre sikt. Vi vil spille inn forslag om visse forenklinger av bokmålsrettskrivningen, som fagrådet så kan ta opp til behandling i henhold til sine vedtekter. Ut fra Språkrådets notat til Kulturdepartementet mars 2014 (Språkrådet 2014) ser det ut til at man i Rådet for tiden vil prioritere utviklingen av Norsk ordbank. Dette trenger ikke å være til hinder for et fortsatt samarbeid om Tanum, men det gjør det mer usikkert om Rådet ønsker å engasjere seg mer aktivt i et prosjekt som det ikke selv styrer.

Hvis Språkrådet mot formodning verken skulle ønske å engasjere seg aktivt for eller sette sitt stempel på Tanum i fremtiden, vil vi selvsagt stå friere til å redigere en stor rettskrivningsordbok for bokmål ut fra egendefinerte prinsipper. Alt tyder på at denne ressursen i overskuelig fremtid vil være mer omfattende enn databasen Språkrådet rår over, og dermed vil Tanum ha forutsetninger for å bli en nødvendig referanse for språksamfunnet også i fremtiden.

Litteratur

Ordbøker

Alnæs, Ivar (1910): *Norsk uttale-ordbok*. Kristiania: H. Aschehoug & Co. (W. Nygaard).

- Bokmålsordboka* (1986). Marit Ingebjørg Landrø & Boye Wangenstein (red.). Oslo: Universitetsforlaget. 3. utgave (2005) Boye Wangenstein (red.) Oslo: Kunnskapsforlaget. <<http://bit.ly/1muS6dp>> (mai 2014).
- Guttu, Tor & Boye Wangenstein (red.) (2012): *Nyord i norsk*. Oslo: Kunnskapsforlaget.
- Juell Tønnessen, S. (1910): *Retskrivningsordbok over det norske riksmåal*. Kristiania og Kjøbenhavn: Gyldendalske Boghandel Nordisk Forlag.
- Juell Tønnessen, S. (1926): *Norsk og fremmed-ordbok*. Oslo: Gyldendal Norsk Forlag.
- Krogsrud, Torgeir & Didrik Arup Seip (1924): *Norsk riksmålsordbok*. Kristiania: Steenske Forlag.
- Nynorskordboka* (2006). Marit Hovdenak et al. (red.). 4. utgåva. Oslo: Det Norske Samlaget. <<http://bit.ly/1nihodk>> (mai 2014).
- Ordnnett <<http://www.ordnett.no>> (april 2014).
- Retskrivningsordbogen* (2012). Anita Ågerup Jervelund, Jørgen Schack, Jørgen Nørby Jensen & Margrethe Heidemann Andersen (red.). København: Alinea A/S. <<http://www.dsn.dk/ro/ro.htm>> (mai 2014).
- Svenska Akademiens ordlista* (2006). Trettonde upplagan. Stockholm: Norstedts Akademiska Förlag. <<http://bit.ly/SPoES7>> (mai 2014).
- Sverdrup, Jakob & Marius Sandvei (1940): *Norsk rettskrivningsordbok*. Oslo: Johan Grundt Tanum.
- Tanums store rettskrivningsordbok*. 9. utgave. Boye Wangenstein (red.) (2005): Oslo: Kunnskapsforlaget.

Annen litteratur

- avis.uib.no = Norsk aviskorpus <<http://avis.uib.no>> (april 2014).
 Malmgren, Sven-Göran (2014): *Svenska Akademiens ordlista* gjennom 140 år: mot fjortonde opplagan. I: *LexicoNordica* 21 (dette bind).
- Språkrådet 2014 = *Norsk ordbokpolitikk*. (Notat fra Språkrådet til Kulturdepartementet 3. mars 2014) <<http://bit.ly/1nNiK3B>> (mai 2014).
- Språkrådets vedtekter <<http://www.sprakradet.no/Toppmeny/Om-oss/Vedtekter-for-Sprakradet/>> (mars 2014).
- Stortingsmelding nr. 35 (2007–2008): *Mål og mening. Ein heilskapleg norsk språkpolitikk*. Oslo: Det kongelige kultur- og kyrkjedepartement.
- Aars, J. & S.W. Hofgaard (1907): *Norske retskrivnings-regler med alfabetiske ordlister*. Kristiania: W. C. Fabritius & Sønner A/S.

Øystein Eek
 seniorredaktør, cand.philol.
 Kunnskapsforlaget
 Postboks 4432 Nydalen
 NO-0403 Oslo
 oystein.eek@kunnskapsforlaget.no

Boye Wangenstein
 ordbokredaktør, cand.philol.
 Kunnskapsforlaget
 Postboks 4432 Nydalen
 NO-0403 Oslo
 boye.wangenstein@kunnskapsforlaget.no

Oppbyggingen av en relasjonsbasert islandsk ordboksbase

Jón Hilmar Jónsson

The article describes the structuring of a relational dictionary database of Icelandic which partly derives from the lexicographic analysis of material in the Orðabók Háskólans archives and additionally builds upon the material of three related phraseological dictionaries of Icelandic from the years around 2000. The syntactic relations of words and phrases in this material reveal semantic relations building up a big collection of synonyms and other kinds of semantically related lexical items. An open and extendable lemma list (including multi-word lemmas) and a systematic presentation of the lemmas together with a comprehensive analysis of textual occurrences bridges the gap between a corpus and a ready-made dictionary.

1. Bakgrunn og utgangspunkt

Islandsk ordbokstradisjon skiller seg fra tradisjonen i de skandinaviske land der oppmerksomheten tidlig var konsentrert om utgivelsen av store nasjonale ordboksverk. På Island arbeidet man riktignok lenge mot det målet å utgi en historisk ordbok over islandsk fra reformasjonen og boktrykkingens inntog til det samtidige språket, alt fra opprettelsen av Orðabók Háskólans i året 1944. Da hadde Den Arnamagnæanske Kommission i København allerede tatt initiativ til å utarbeide *Ordbog over det norrøne prosasprog* (ONP) som ville dekke det eldre språket (jf. Battista & Johannsson 2014). Denne arbeidsfordelingen innebar at språkhistorien i sin helhet ikke ville få en sammenhengende dekning, og beskrivelsen av det yngre språket ville bli foretatt uten større hensyn til det rike kildematerialet til det eldre språket.

Innsamlingen av materiale til den ordboken *Orðabók Háskólans* skulle utarbeide, ble utført på tradisjonell vis ved en omfattende ekserpering av tekstbelegg, samtidig som det ble lagt vekt på innhenting av opplysninger om ord og ordbruk rundt omkring i landet, især gjennom et regelmessig radioprogram i mange år (se brosjyren *Orðabók Háskólans. Leksikografisk institutt ved Islands universitet* fra 1995).

Da dataalderen meldte seg tidlig i 1980-årene, var materialeinnsamlingen fortsatt i gang og det forelå ingen konkrete planer om utarbeidelse og utgivelse av et større nasjonalt ordboksverk. I stedet for å fastlegge en plan i den retning var det nå blitt mer aktuelt å utnytte den nye teknologien til å få en helhetlig oversikt over det innsamlede materialet og skaffe seg innsikt i materialets sammensetning og ut fra dette ta stilling til en videre planlegging. I samsvar med det ble det laget et oversiktsregister over oppslagsordene i instituttets hovedarkiv med noen grunnleggende markeringer som senere ble etterfulgt med inntasting av størsteparten av beleggmaterialiet.

Det ble etter hvert klart at islandsk ikke ville få en leksikografisk beskrivelse i form av en stor ordbok som kunne måle seg med de store skandinaviske ordboksverk. Det islandske ordforrådet hadde allerede fått en utførlig beskrivelse i *Íslensk-dönsk orðabók* tidlig i 1920-årene, og den lå for en stor del til grunn for den første allmenne enspråklige islandske ordboken, *Íslensk orðabók*, som først utkom i 1963. Denne ordboken er senere blitt utvidet og nylig også utgitt i digital form.

På bakgrunn av dette har man i det akademiske ordboksarbeidet kunnet utvikle prioriteringer som tar mer hensyn til de faktiske forholdene og blir sterkere motivert av utviklingen innenfor leksikografisk databehandling og de muligheter den digitale ordboken byr på. Dette har bl.a. resultert i sammenhengende ordboksprosjekter med en sterk fokusering på leksikografiske relasjoner som foruten å forenes i et konkret ordboksverk gradvis har bygget opp en flerfunksjonell islandsk ordboksbase.

Denne artikkelen skal gi en bred oversikt over utviklingen av dette arbeidet og sammensetningen av ordboksbasen, hvordan de enkelte komponenter er relatert til hverandre og hvordan de spiller sammen i utformingen og fremstillingen av en webbasert digital ordbok.

Innledningsvis er det verdt å fremheve to viktige forhold som har vært toneangivende i hele dette arbeidet. Kartleggingen av Orðabók Háskólans beleggmateriale ble i sin tid foretatt uten begrensninger med hensyn til lemmaseleksjon, og den dekket dermed en kolossal mengde „likestilte“ oppslagsord. Disse oppslagsordene og de belegg de bygget på, hadde ennå ikke gjennomgått betydningsdifferensiering og betydningsbeskrivelse. Dermed var det semantiske aspektet ikke umiddelbart i fokus og det ble aktuelt og hensiktsmessig å anlegge andre synsvinkler.

2. Syntagmatiske relasjoner i forgrunnen

Analysen av beleggmateriale i Orðabók Háskólans hovedarkiv (med belegg fra trykte språkkilder) ble i den første fasen for en stor grad rettet mot verbene, med særlig vekt på deres syntagmatiske egenskaper og mangfold i konstruksjoner (jf. *Sýnihefti sagn-orðabókar* 1993). En systematisk kartlegging av verbene og deres omfattende beleggsamling synliggjorde hvor viktig det er å belyse deres opptreden i syntagmer, ordforbindelser og setninger, også med hensyn til deres semantiske egenskaper og betydningsbeskrivelse. Denne erfaringen bidro til konstrueringen av en særskilt konstruksjons- og kombinasjonsordbok under tittelen *Orðastaður*, som først utkom i 1994 og i 2. utgave i 2001, der verb og verbkonstruksjoner inntar en forholdsvis stor plass i lemmalisten.

Arbeidet med ordforbindelser, kollokasjoner og konstruksjoner som ble innledet med *Orðastaður* ble utvidet og komplettert i 1990-årene. I disse årene ble det laget en omfattende database med ordforbindelser fra Orðabók Háskólans beleggsamling (her-

etter *Ordfb*), i alt omkring 135.000 ordforbindelser knyttet til fast 60.000 lemmaer. I 2002 ble *Orðastaður* komplettert og utvidet i form av en ny ordbok, *Orðaheimur*, der oppmerksomheten for en stor del gjaldt idiomene (se Jónsson 2005). Her var lemmalisten todelt, en hovedlemmaliste med alfabetisk ordnete begrepsbetegnelser og en tilleggsliste (register) med samtlige innholdsord i ordforbindelsene der disse var innordnet etter en fast sorteringsregel. Det siste skrittet i dette arbeidet besto i en sammenslåing av *Orðastaður* og *Orðaheimur* med en kombinert lemmaliste og et felles ordforbindelsesregister under overskriften *Stóra orðabókin um íslenska málnotkun* (heretter *Stob*) i 2005 (se Jónsson 2008).

Selv om det syntagmatiske aspektet står sentralt i alle disse prosjektene der hovedvekten legges på kollokasjoner, idiomer og konstruksjoner, er også andre leksikografiske relasjoner inne i bildet.

Orðastaður inneholder en morfologisk komponent som består i en semantisk grovklassifisert liste med sammensetninger til de enkelte lemmaer og som i alt dekker omtrent 100.000 ord. Foruten å belyse og utdype lemmaenes karakter gjenspeiler disse sammensetningene gjerne semantisk nærhet og likheter gjennom sine paradigmatisk relasjoner.

I *Orðastaður* og *Orðaheimur* (og dermed også i *Stob*) blir ledd i ordforbindelser med varierende leksikalsk innhold markert og avgrenset (med hakeparentes):

- (1) *ávarpa* <*samkomuna*> ‘adressere forsamlingen’
- (2) <*hnífurinn*> *bítur* <*vel*> ‘kniven er skarp’

Her blir det gjerne angitt to eller flere typiske representanter for det ledd det dreier seg om, som dermed står i en innbyrdes paradigmatisk relasjon og som regel vil være nært semantisk beslektet:

- (3) *setja* <*fundinn, ráðstefnuna*> ‘åpne møtet/konferansen’
- (4) <*reiðin, heiftin*> *ólgar í* <*honum, henni*> ‘sinnet koker i ham/henne’

En standardisert fremstilling av ordforbindelsene (med hensyn til faktorer som argumentstruktur, valens, kasusstyring og fakultative ledd) åpner veien for et samlet alfabetisert register over samtlige ordforbindelser der innholdsordene fungerer som lemmaer og i mange tilfeller vil dekke en stor mengde ordforbindelser av ulike slag. Dette kommer klarest til syne i *Stob* der materialet i *Orðastaður* og *Orðaheimur* er blitt slått sammen til et felles register. Her kommer de paradigmatiske relasjoner ganske sterkt til uttrykk innenfor de ulike ordklasser, blant adjektiver som kollokatorer til enkelte substantiver (*þrotlaust/hvíldarlaust/sleitulaust erfiði* ‘uavbrudt/... slit’), blant de substantiver adjektivene står sammen med (*þrotlaust erfiði/strit/þrældómur* ‘uavbrudt slit/strev/slaveri’), blant objekter til enkelte verb (*heyra hávaða/skothvell/fótatak* ‘høre bråk/smell/fottrinn’) osv.

I *Orðaheimur* tar beskrivelsen utgangspunkt i ordforbindelsenes semantiske egenskaper i og med at den i hovedsak gjelder ordforbindelser av metaforisk karakter der de enkelte ordene ikke utgjør en logisk tilgang til vedkommende forbindelse. Isteden blir ordforbindelsene dekket av en alfabetisk lemmaliste med drøyt 800 begrepsbetegnelser hvor det innenfor de enkelte artikler blir foretatt en videre gruppering med hensyn til semantiske og kommunikative faktorer. Naturlig nok dreier det seg her i hovedsak om abstrakte begreper, som f.eks. egenskaper, følelser og menneskelig atferd (med overskrifter som *Ákveðni* ‘Besluttsomhet’, *Heiðarleiki/ráðvendni* ‘Hederlighet’, *Ótti* ‘Frykt’, *Reiði* ‘Sinne’, *Ofbeldi* ‘Vold’, *Mistök* ‘Feil’), mens de konkrete begreper, som i høyere grad er knyttet til enkeltord, får mindre oppmerksomhet.

I denne fasen av ordboksarbeidet, frem til utarbeidelsen og utgivelsen av *Stob* i 2005, kommer dermed samspillet mellom de leksikografiske relasjoner til syne i ulike sammenhenger uten at den kan sies å få en klar og selvstendig fokusering. Behovet for å få til en hensiktsmessig tilgangsstruktur ledet frem til etableringen av et altomfattende alfabetisk register. Men overfor fremveksten av den digitale ordboken med en umiddelbar brukertilgang og sine fleksi-

ble søkemuligheter er en slik tilgangsstruktur ikke lenger tilfredsstillende fra et allment brukerperspektiv. Det krever tilpasning til nye forutsetninger som samtidig innebærer et klart behov for en videre utvikling av en sammensatt ordboksmodell som forener de ulike leksikografiske relasjoner.

3. Omstrukturering, utvidet lemmabegrep og entydiggjøring

Den store utfordringen for den ordboksmodellen som her er under utvikling består i å gi de semantiske relasjonene en mer fremtredende plass og å utforme en semantisk komponent i et aktivt samspill med den syntagmatiske komponenten. En slik målsetning er ikke realiserbar innenfor en tradisjonell semasiologisk ordbok med flertydige ettordslemmaer. Hvis også ordforbindelsene skal kunne spille en aktiv rolle som betydningsbærende enheter må de få direkte tilgang til selve lemmalisten som likeverdige enheter sammen med ettordslemmaene. Innenfor en digital ordbok er dette ikke bare mulig, det er faktisk helt naturlig med hensyn til en fleksibel datastrukturering og brukertilgang.

3.1. Utvidet database, gjenspeiling av semantiske relasjoner og semantisk klassifisering

De leksikografiske data *Stob* og *Ordfb* inneholder kan sies å utgjøre et optimalt grunnlag for en semantisk orientert beskrivelse, ut fra den forutsetning at de syntagmatiske og paradigmatiske relasjoner gjenspeiler relasjoner av semantisk karakter. En fylldig samling med kollokasjoner er f.eks. av stor verdi i denne sammenheng. Med tanke på det og med hensyn til klare strukturelle likheter og en parallell fremstilling ble samtlige data fra *Stob* og *Ordfb* slått sammen i en ny og utvidet database. Det innledet en ny fase i opp-

byggingen av den tilsiktede ordboksbasen som ble markert ved å gi det nye prosjektet overskriften *Íslenskt orðanet* (heretter forkortet til *Íon*, se Jónsson & Úlfarsdóttir 2011).

De semantiske relasjonene gjelder dels det gjensidige forholdet mellom enkelte ord, i form av synonymer og antonymer. Slike relasjoner er sjelden av absolutt karakter og markeringen av synonymer og antonymer vil i svært mange tilfeller være en vurderingssak. Ved en datastøttet analyse kan de semantiske relasjonene derimot vurderes på en bredere skala som antyder ulik grad av semantisk nærhet. Når det gjelder flerordsenheter må de strukturelle egenskapene også trekkes inn i bildet slik at relasjonene gjelder strukturelt likeverdige enheter (som da vil være gjensidig utskiftbare i tekstsammenheng).

Ved siden av denne markeringen og analysen består kartleggingen av de semantiske relasjonene i en overordnet begrepsklassifisering. En slik klassifisering kan bli foretatt på ulike måter og være mer eller mindre overgripende. I *Íon* blir den utført på samme måte som i *Orðaheimur*, slik at lemmaer av samme klasse får en felles overskrift i form av et belysende substantiv som tilhører vedkommende klasse. Disse klassene (begrepsfeltene) er av ulik omfang og overlapper i betydelig grad. I stedet for å foreta en streng hierarkisk systematisering blir begrepsklassifiseringen fremstilt i form av en fri og bevegelig konstellasjon av semantiske grupperinger, noe som i mange tilfeller gjenspeiler faktiske forhold i språkbruken, ikke minst når det gjelder flerordsenheter og ordtilfang som er knyttet til abstrakte begreper.

Det er ikke realistisk å tenke seg at begrepsklassifiseringen vil gi en helhetlig dekning av lemmabestanden til *Íon*. Markeringen av gjensidige relasjoner som synonymi og antynymi vil også bare dekke en del av lemmaene. Men et aktivt samspill mellom disse to operasjoner skal kunne gi en verdifull innsikt i de relasjoner som på ulike plan knytter sammen semantisk beslektede ord og andre leksikalske enheter.

3.2. Lemmatisering av betydningsbærende flerordsenheter, entydige lemmaer

Det mest overgripende og tidkrevende arbeidet ved utformingen av *ÍOn* består i å innføre flerordslemmaer og dermed å trekke betydningsbærende flerordsenheter inn i lemmalisten på samme nivå som ettordslemmaene. Dette henger sammen med kravet om semantisk entydige lemmaer, som utgjør en forutsetning for en problemfri analyse og markering av innbyrdes semantiske relasjoner. Til sammen betyr dette en vesentlig endring av ordklassenes karakter og deres innbyrdes størrelsesforhold, i første omgang særlig med hensyn til verbene som øker drastisk i antall.

Fremstillingen av ordforbindelsene i *Stob* gjelder i hovedsak som mønster for fremstillingen av flerordsverb i *ÍOn*, ved markering av valens og fakultative ledd. Lemmatiseringen betyr bl.a. at en stor mengde idiommer fremstår som selvstendige verblemmaer og kan dermed på naturlig måte inngå i innbyrdes relasjoner. Det samme gjelder for verbkonstruksjoner med en mer varierende leksikalsk representasjon. Her kommer forskjellen gjerne direkte til uttrykk i markeringen av fakultative ledd:

- (5) a *brjóta ísinn* 'bryte isen'
 b *brjóta* <*glerið, bollann, spýtuna*>
 'knuse glasset/koppen/planken'

Innførelsen av flerordslemmaer er også av stor betydning for adverbialer som i stor utstrekning får en klar lemmastatus på linje med ettordsadverb:

- (6) a *líklega, að öllum líkindum* ('sannsynligvis')
 b *undanfarið, upp á síðkastið* ('i det siste')

Blant flerordslemmaene foregår entydiggjøringen stort sett på en

enkel måte og kommer til uttrykk gjennom selve formen, der f.eks. ulik semantisk valens gjenspeiler betydningsmessige skilnader:

- (7) a *rjúfa* <*vegginn, þakið*> ‘slå et hull i veggen/taket’
 b *rjúfa* <*sættina; heitið*> ‘bryte fredsavtalen/løftet’

Ellers er flertydighet med homonyme lemmaformer forholdsvis sjelden. I de tilfeller skilles det mellom lemmaene gjennom en homonymindeks sammen med en identifiserende kommentar til betydningen:

- (8) 1 *bera höfuðið hátt reigja höfuðið* (‘bøye hodet bakover’)
 2 *bera höfuðið hátt vera stoltur* (‘være stolt’)

Blant ettordslemmaene er flertydighet og dermed homonymi mye mer utbredt. Især blant substantivene blir lemmaene adskilt gjennom homonymindekser:

- (9) 1 *hóf hófsemi* (‘måtehold’)
 2 *hóf samkvæmi, veisla* (‘selskap, fest’)

I den faktiske språkbruken er adjektivenes betydning ganske flytende og tar i stor utstrekning farge av de substantiver de står sammen med og karakteriserer. For å gjenspeile dette og dermed få til et fungerende samspill med semantisk beslektede lemmaer får adjektivlemmaene gjerne et tilleggselement i form av et differensierende substantiv:

- (10) a *mildur [dómur]* ‘mild dom’
 b *mildur [hljómur, tónn]* ‘mild klang/tone’
 c *mildur [regn]* ‘mild regn’
 d *mildur [veður, tíðarfar, loftslag]*
 ‘mild vær/værforhold/klima’
 e *mildur [vetur]* ‘mild vinter’

Disse lemmaformene er tydeligvis relatert til adjektivene i deres attributive funksjon. Henvisning til personer blir her ikke markert, og predikativer blir fremstilt som flerordslemmaer av typen *vera* + *Adjektiv*:

- (11) *vera mildur* ‘være mild’

Tilsvarende lemmaformer gjelder også for predikative substantiver:

- (12) *vera ljúfmenni* ‘være et elskverdig menneske’

- (13) *vera eyðslukló* ‘være en sløset person’

Dermed får predikativene en selvstendig lemmastatus som trekker frem deres innbyrdes relasjoner og fremhever deres verbale karakter, især når de inngår i synonymrelasjoner med andre typer verblemmaer:

- (14) *vera vellauðugur* ‘være stenrik’

- (15) *vera margfaldur milljónamæringur* ‘være *multimiljonær*’

- (16) *vita ekki aura sinna tal* ‘ha flust med penger’

Denne omfattende omstruktureringen av den opprinnelige lemmalisten i *ÍOn* åpner ikke bare veien for en bred kartlegging av lemmaenes semantiske relasjoner, den endrer faktisk hele ordboksbasens karakter. Hele den leksikografiske analyse og klassifisering foregår nå på lemmaplanet og dekker en betydelig større mengde lemmaer enn tidligere.

Tilgangen til store islandske tekstsamlinger og korpusmateriale er blitt vesentlig forbedret de siste årene. For oppbyggingen av ordboksbasen har dette vært av stor betydning. Det digitale biblioteket *Tímarit.is* dekker mesteparten av islandske aviser og tidsskrifter helt fra slutten av 1800-tallet til dags dato. Denne store tekstsam-

lingen har vært til stor nytte i ordboksarbeidet. Det gjelder også for den første grammatisk taggede korpusen over islandsk nåtids-språk, *Mörkuð íslensk málheild*, som ble ferdiggjort i 2011. Den leksikografiske analysen har i økende grad hentet tilleggsmateriale fra disse kildene for dermed å få til en bredere dekning av det aktuelle analyseobjektet.

3.3. Grammatisk tagging av flerordslemmaer

Det tilhører den leksikografiske tradisjon å angi ordklassebetegnelse ved de enkelte lemnaer. I en statistisk omgivelse som i en trykt ordbok spiller denne angivelsen ikke noen aktiv rolle, men den utgjør et klassifiseringselement som kan ha en praktisk betydning i andre sammenhenger, f.eks. som søkefaktor i en digital ordbok. Man kan godt tenke seg at ordklasseskilnaden får en klarere funksjon i leksikografisk sammenheng, men uansett dreier det seg om en forholdsvis grov grammatisk klassifisering.

Lematisering av flerordsenheter byr her på en mye mer detaljert grammatisk markering og dermed en mer finkornet klassifisering av lemnaene. En slik markering vil kunne styrke flerordslemnaenes egenverdi og selvstendighet i betydelig grad. Med hensyn til dette er det blitt foretatt en grammatisk tagging av samtlige flerordslemnaer i ordboksbasen, der faktorer som ordklasse, kasus og bestemthet står i forgrunnen. Det betyr f.eks. at verblemnaene kan klassifiseres ut fra kasusstyring og at verb med objekt i akkusativ kan bli avgrenset som en klasse for seg. Det viser at taggen “so <no-ag>” (verb med substantivisk objekt i akkusativ i bestemt form, f.eks. *lesa <bókina; söguna; textann>* ‘lese boken/historien/teksten’) omfatter omtrent 2500 lemnaer, mens søkestrengen “so <no-ag>*” (der objektet kan etterfølges av et komplement i en eller annen form) matcher drøyt 4000 lemnaer. Derimot dekker tilsvarende tagger med dativobjekt, “so <no-dg>” og “so <no-dg>*”, henholdsvis ca. 400 og 1200 lemnaer. Til

sammenlikning inneholder ordboksbasen omtrent 1500 lemmaer der verbet forutsetter et dativsubjekt (lemmaer som <mér> *leiðist* ‘jeg kjeder meg’, <mér> *líður* <vel> ‘jeg har det bra’).

Et slikt leksikalsk uavhengig klassifiseringsgrunnlag er av stor verdi, både med hensyn til oppbyggingen av databasen og med tanke på praktiske brukerbehov overfor en webbasert digital ord-bok. Når det gjelder ordboksbasen, er det bl.a. til direkte nytte ved den semantiske klassifiseringen av lemmaene, for de grammatiske likhetene viser seg i svært mange tilfeller å gjenspeiles i semantiske likheter.

3.4. Semantisk analyse og markering av semantisk nærhet

Markeringen av semantiske relasjoner som synonymi og antonymi vil i mange tilfeller kunne bygge på leksikografens intuisjon og innsikt, ikke minst i de tilfeller der de leksikografiske data ikke bidrar med noen klar vitnesbyrd. Ord og ordforbindelser kan til og med være så sjeldne at de ikke eller kun sporadisk forekommer i store tekstsamlinger. Men ved å vurdere ordenes semantiske egenskaper og relasjoner på bakgrunn av deres forekomst og tilknytninger i tekstsammenheng gis det et bedre grunnlag for den semantiske analyse og beskrivelsen av det innbyrdes forholdet mellom nærstående leksikalske enheter.

Når leksikografens oppmerksomhet gjelder enkeltordenes forekomst i kollokasjoner og konstruksjoner, blir deres opptreden som sideordnete enheter mest ute av fokus. Når fokuseringen overføres til likestilte leksikalske enheter og deres innbyrdes relasjoner slik det skjer ved omstruktureringen av den islandske ordboksbasen, blir forbindelser av denne typen derimot et høyaktuelt objekt i ordboksarbeidet.

I arbeidet med *ÍOn* har parataktiske forbindelser og de paradigmatiske og dermed semantiske relasjoner de avdekker, fått sta-

dig økt betydning, både for den leksikografiske analyse og som et direkte middel til å fortette dekningen av enkelte begrepsfelter og til å trekke inn nye lemmaer. Verdien ligger ikke minst i det forholdet at parataktiske forbindelser karakteriserer jevnt bruken av enkeltord som betydningsbærende flerordsenheter (*frost og kuldi* 'frost og kulde', *styrjaldir og vopnuð átök* 'krig og væpnede konflikter'; *einbeittur og viljasterkur* 'besluttsom og viljesterk', *einbeittur og fastur fyrir* 'besluttsom og bestemt'; *biðjast afsökunar og skammast sín* 'be om unnskyldning og skamme seg', <*aðgerðin*> *heppnast og ber árangur* 'operasjonen lykkes og bærer frukt', *það kólnar og hvessir* 'det blir kaldere og det blåser opp') og de er karakteristiske for alle åpne ordklasser og deres flerordsforbindelser. Ut fra det synspunktet er de av spesiell betydning for analysen av flerordslemmaene, for her er hypotaktiske forbindelser ikke inne i bildet som et alternativt analysegrunnlag på samme måte som i ettordslemmaene.

I mange begrepsfelter i *ÍÖn* er det bygget opp et tettmasket nett av lemmaer som er relatert gjennom parataktiske forbindelser i tekstsammenheng og er dessuten på forskjellig måte knyttet til lemmaer innenfor semantisk nærstående felt. Det gjenspeiler mangfoldigheten og dynamikken i ordboksbasen og knytter sammen beslektet ordforråd over begrepsfelter. Men disse forbindelsene bidrar også med data som inngår i den semantiske analysen og gir markeringen av synonymi og semantisk nærhet en ny dimensjon.

En større samling parataktiske forbindelser kan gi et interessant grunnlag for sammenlikning av lemmaer som gjennom felles partnere er knyttet til hverandre. Her vil antallet felles partnere gi en sterk antydning om semantisk nærhet. Det betyr at sett fra et bestemt lemma kan denne nærheten vises frem i form av en gradert liste over nærstående lemmaer med antatt forskjellig nærhetsgrad. Her må det naturligvis tas hensyn til hvorvidt de aktuelle lemmaer har fått likeverdig dekning slik at deres innbyrdes fre-

kvens i parataktiske forbindelser også blir trukket inn i bildet. Men uansett hvordan denne nærheten blir tolket og fremstilt, dreier det seg om en viktig informasjon som belyser hvor nært enkelte lemmaer med lik opptreden er relatert til hverandre.

Som et eksempel kan vi se på to svært nærstående substantiver, *andvaka* og *svefnleysi* 'søvnløshet'. I databasen forekommer *svefnleysi* i 167 ordpar, *andvaka* i 67. Blant disse dreier det seg om 27 felles partnere: *áhyggjur* ('bekymring') og *svefnleysi*, *áhyggjur* og *andvökur*; *kviði* ('angst') og *svefnleysi*, *kviði* og *andvökur*; *svefnleysi* og *ofþreyta* ('utmattelse'), *ofþreyta* og *andvökur*. Mange av disse felles ordene viser seg også å stå høyt oppe på lista over de ord som har de fleste felles partnere med *andvaka* og *svefnleysi*: *vökur* 'våking' (24 med *andvaka*, 27 med *svefnleysi*), *sálarstríð* 'indre strid' (17 med *andvaka*, 16 med *svefnleysi*), *kviði* 'angst' (11 med *andvaka*, 20 med *svefnleysi*), *þunglyndi* 'depresjon' (9 med *andvaka*, 16 med *svefnleysi*). Et slikt vitnesbyrd understreker en sterk semantisk nærhet mellom vedkommende ord. En analyse av ordenes ulike tilknytninger i forbindelser av denne typen vil dessuten i mange tilfeller kunne trekke fram differensierende faktorer og dermed være til nytte ved presisering av ordenes semantiske egenskaper og karakterer.

Foruten å trekke frem semantisk nærhet på denne måten gir direkte søking på parataktiske forbindelser til de enkelte lemmaer et viktig bidrag til oppbyggingen av ordboksbasen i to henseender. For det første vil det oppdages nye aktuelle lemmakandidater som på en enkel måte kan inkluderes i basen og vil berike vedkommende begrepsfelt eller begrepsfelter. For det andre vil utbyttet i svært mange tilfeller bestå i en lang og innholdsrik liste forbindelser med utstrakt semantisk likhet blant partnerne som i sin tur byr på klassifisering og vil være et viktig input i en nærmere semantisk analyse.

4. Webbasert ordbok, semantiske relasjoner i forgrunnen

Hensikten med etableringen av *ÍOn* var dels å utvikle og levere et konkret leksikografisk produkt i form av en webbasert digital ordbok. Denne ordboken er tilgjengelig på websiden <www.ordanet.is> og er fortsatt under utvikling. Med hensyn til det er det ikke grunn til å gå nøye inn på avgrensingen og utformingen av denne ordboken her. Hittil har den vært konsentrert om bestemte datatyper der de semantiske relasjonene står i forgrunnen. Tilgangen til de enkelte begrepsfelter går gjennom søk på hvilket som helst lemma som tilhører feltet, og det samme gjelder for synonymer og antonymer. Ordklassebetegnelse og den grammatiske markeringen av flerordslemmaene spiller også en aktiv rolle ved at brukeren kan velge hvordan lemmaene innenfor enkelte begrepsfelter skal sorteres, enten alfabetisk eller etter den grammatiske taggen. Her ville det være interessant å by på en mer direkte tilgang til markeringen slik at brukeren kan søke på bestemte tagger eller kombinasjoner av disse (se Jónsson & Úlfarsdóttir 2011:97–99).

De parataktiske forbindelsene setter et sterkt preg på mange lemmaer i ordboken slik den ser ut nå der de enkelte partnerne utgjør interaktive lenker. Men det gjenstår å bygge inn informasjon om hva denne datatypen antyder om semantisk (mest) nærstående lemmaer.

Uansett hvordan den webbaserte ordboken kommer til å utvikles vil den fortsatt bære preg av å bygge på en større ordboksbase som kan kompletteres og utvides i ulike retninger og ikke har noen skarpe grenser. Det vil bl.a. gjenspeiles i betydelig ubalanse mellom de enkelte lemmaer der noen viser mange og mange slags tilknytninger til andre lemmaer, mens andre kan bygge sin tilhørighet på en enkel og ensidig relasjon.

Ordboken inneholder nå drøyt 150.000 lemmaer, mens selve ordboksbasen omfatter ca. 260.000 lemmaer. Denne forskjellen

kan sies å gjenspeile ulik lemmaverdi. De lemmaer som hverken inngår i en semantisk relasjon (synonymi, antonymi, begrepsfelt) eller opptrer som ledd i en parataktisk forbindelse, vil til så lenge stå utenfor ordboken slik den blir definert nå. En del av disse lemmaene kan allikevel forekomme i syntagmatiske relasjoner (kollokasjoner og idiomer) i ordboksbasen, mens andre (og da først og fremst sammensetninger) utelukkende forekommer i morfologiske relasjoner. Ordboksbasen inneholder dermed en hel del latente lemmaer som i første omgang har en liten aktiv funksjon, men som vil aktiveres når de inngår i relevante relasjonstyper.

5. Flerfunksjonell ordboksbase

I det forangående har beskrivelsen av den islandske ordboksbasen vært konsentrert om hvordan den er bygget opp omkring leksikografiske relasjoner med de syntagmatiske relasjonene som grunnlag for en omfattende leksikografisk analyse der de semantiske relasjonene står i sentrum. Termen *ordboksbase* i denne sammenheng har en flertydig funksjon. Den kan simpelthen henvise til det samlede datamateriale som ligger til grunn for utarbeidelsen av en konkret ordbok, altså *ÍO*ns webordbok i dette tilfellet. Men termen har også en bredere henvisning til en mer allmenn funksjon som utgangsmateriale og grunndatabase for leksikografisk arbeid, både med hensyn til forskning og praktiske ordboksprosjekter (se også Jónsson 2013). Denne allmenne funksjonen har etter hvert fått større vekt i utformingen av arbeidet og fortjener derfor en nærmere presentasjon.

5.1. Lemmaseleksjon og lemmaverdi

Behovet for tilgang til en større leksikografisk database kan aktualiseres i ulike sammenhenger. Med hensyn til islandsk vil det i

praksis primært oppstå ved utarbeidelse av konkrete ordbøker, såvel enspråklige som to- eller flerspråklige. Når det gjelder lemaseleksjon og vurdering av lemmaverdi, kan ordboksbasen bidra med veiledende opplysninger. Her kommer bl.a. antall og variasjoner i kollokasjoner, parataktiske forbindelser og konstruksjoner inn i bildet. Ordboksbasen vil også kunne underlette en begrepsmessig oversikt over lemmabestanden og en spesiell vurdering av lemaseleksjonen i forhold til enkelte begrepsfelter. Den semantiske klassifiseringen av lemmaene i basen kan dessuten ha en kontrollerende funksjon ved betydningsbeskrivelsen, og valg og fremstilling av ekvivalenter for behandlingen av enkelte lemmaer kan kontrolleres mot relaterte lemmaer innenfor samme felt. Det betyr til og med at fremdriften av ordboksbeskrivelsen kan led-sages av begrepsklassifiseringen. Det har allerede vist seg å være av praktisk betydning i arbeidet med den islandsk-skandinaviske ordboken *Islex*, der en grov semantisk markering av de islandske lemmaene, som for en stor del er hentet fra ordboksbasen, gjorde det mulig å arbeide med semantisk sammenhengende ordtilfang i hver enkel redigeringsfase (se Úlfarsdóttir 2013).

5.2. Semantisk nærhet

Markering av semantisk nærhet og det datamaterialet den bygger på kan ha stor verdi for ordboksarbeid i ulike sammenhenger. Foruten å utdype en tradisjonell synonymbeskrivelse styrker den grunnlaget for lemaseleksjon og gir viktig veiledning for valg av ekvivalenter bygget på konkrete tall. Her er det verdt å understreke at opplysninger om leksikalsk frekvens er av spesiell betydning når de gjelder innbyrdes sammenhengende ord eller andre leksikalske enheter, men slike opplysninger har lenge vært sparsomme i leksikografisk arbeid.

En ordboksbase som prioriterer en dyptgående og omfattende semantisk klassifisering av ordforrådet, er spesielt verdifull når det

gjelder å finne motsvarigheter på målspråkssiden til kildepråklige elementer i en to- eller flerspråklig ordbok. Her må valget utelukkende styres av innholdet til de aktuelle ekvivalenter, og for å kunne treffe et veloverveid valg må leksikografen ha oversikt over et bredt leksikalsk spektrum. Ut fra en slik oversikt er det også mulig å foreta en mer generell kontrastiv analyse av forholdet mellom to språk der motsvarigheter og likheter inngår i en større kontrastiv sammenheng (jf. Hannesdóttir & Jónsson 2001, Hannesdóttir & Tingsell 2011).

5.3. Grammatisk tagging

Grammatisk tagging av betydningsbærende flerordsenheter gir et sorterings- og klassifiseringsgrunnlag som i flere henseender vil være til nytte i praktisk ordboksarbeid, og kan på ulike måter kobles til leksikalske og semantiske elementer og deres opptreden. Den kan også gi det grammatiske aspektet sterkere fokusering i utformingen av konkrete ordbøker, både i enspråklig og flerspråklig sammenheng.

5.4. Grunnlag for leksikografisk arbeid

Alle disse elementene har en viktig funksjon i den islandske ordboksbasen som her har vært skildret og diskutert. En ordboksbase av denne type skal dermed kunne utgjøre et viktig grunnlag for leksikografisk arbeid og konstrueringen av konkrete ordbøker med ulike genuine formål. Det beror bl.a. på at den gjelder et stadium i ordboksarbeidet som kan sies å stå midt i mellom et grovanalysert datamateriale, f.eks. i form av tagget korpus, og en konkret ordbok med et klart genuint formål. Utarbeidelsen av en relasjonsbasert ordbok som en integrert del av selve ordboksbasen understreker og konkretiserer den rolle ordboksbasen tillegges i denne sammenheng.

6. Sammenfatning

Flere omstendigheter bidro til at islandsk ikke på samme måte som de skandinaviske språk fikk et stort nasjonalt ordboksverk i trykt form. Allikevel ga islandsk ordboksarbeid omkring og etter midten av 1900-tallet, der hensikten var å utarbeide et stort ordboksverk over islandsk helt fra boktrykkingens inntog midt på 1500-tallet til samtiden, et rikelig utbytte, bl.a. i form av en verdifull beleggssamling. Dette materialet ble tilgjengelig på en ny måte da datamaskinell registrering og analyse av ordboksmateriale ble aktuell sent på 1900-tallet.

Denne beleggssamlingen utgjør en viktig kilde i arbeidet med noen islandske ordbøker og leksikografiske databaser som ble ferdiggjort omkring århundreskiftet 2000. Her gjelder oppmerksomheten især ordforbindelser og dermed analyse av syntagmatiske relasjoner. Under det arbeidet kom forholdet til de semantiske relasjonene stadig mer inn i bildet. I ordboksprosjektet *Íslenskt orðanet*, som for en stor del bygger på disse ordbøkene og deres materiale, får de semantiske relasjonene en sterk fokusering, med analyse av synonymi og semantisk nærhet og en omfattende og detaljert begrepsklassifisering. Her har innførelsen av flerordslemmer og deres grammatiske tagging en avgjørende betydning for en effektiv og hensiktsmessig analyse.

Gjennom dette arbeidet er det etter hvert bygget opp en flerfunksjonell ordboksbase som både skal kunne støtte utarbeidelsen av nye islandske ordbøker og ha en generell betydning for forskning omkring islandsk ordforråd og islandsk språkbruk. En webbasert digital ordbok utgjør en integrert komponent i denne ordboksbasen og er med på å fremheve den betydning som de leksikografiske relasjoner og deres innbyrdes samspill tildeles.

Litteratur

Ordbøger og databaser

ÍOn = *Íslenskt orðanet*.

Íslensk-dönsk orðabók (1920–1924): Sigfús Blöndal (red.). Reykjavík.

Íslenskt orðanet. <www.ordanet.is> (april 2014).

Íslenzk orðabók handa skólum og almenningi (1963): Árni Böðvarsson (red.). Reykjavík: Bókaútgáfa Menningarsjóðs. [2. utgave 1983, 3. utgave 2007; nettutgave på <<http://snara.is>>].

Mörkuð íslensk málheild. <mim.arnastofnun.is> [Tagget korpus med islandske tekster fra perioden 2000–2010.] (april 2014).

Ordfb = *Skrá um orðasambönd*. <arnastofnun.is/page/skra_um_ordasambond> [Fraseologisk register til Orðabók Háskólans.] (april 2014).

Orðaheimur (2002) = Jón Hilmar Jónsson: *Orðaheimur. Íslensk hugtakaorðabók með orða- og orðasambandaskrá*. Reykjavík: JPV-útgáfa.

Orðastaður (1994) = Jón Hilmar Jónsson: *Orðastaður. Orðabók um íslenska málnotkun*. Reykjavík: Mál og menning. [2. utgave JPV-útgáfa 2001.]

Stob = *Stóra orðabókin um íslenska málnotkun*.

Stóra orðabókin um íslenska málnotkun (2005) = Jón Hilmar Jónsson: *Stóra orðabókin um íslenska málnotkun*. Reykjavík: JPV-útgáfa. <<http://snara.is>> (april 2014).

Sýnihefti sagnorðabókar (1993). Ásta Svavarsdóttir, Guðrún Kvaran, Jón Hilmar Jónsson & Kristín Bjarnadóttir (red.). Reykjavík: Orðabók Háskólans.

Tímarit.is. <<http://timarit.is/>> [Digitalt bibliotek med islandske avistekster.] (april 2014).

Annen litteratur

- Battista, Simonetta & Ellert Thor Johannsson (2014): Ordbog over det norrøne prosasprog 2004–2014 – fra trykt udgave til net-version. I: *Leda-Nyt* 57, 6–23.
- Hannesdóttir, Anna Helga & Jón Hilmar Jónsson (2001): Að hafa í sig og á. Ísländsk fraseologi i ett isländskt-svenskt perspektiv. I: *LexicoNordica* 8, 67–91.
- Hannesdóttir, Anna Helga & Sofia Tingsell (2011): Svensk fraseologi för islänningar – om den isländsk-svenska ordboken i ISLEX och en onomasiologisk svensk resurs: Frاسبanken. I: *LexicoNordica* 18, 61–86.
- Jónsson, Jón Hilmar (2005): Orðaheimur – en fraseologisk begrepsordbok. I: Ruth Vatvedt Fjeld & Dagfinn Worren (red.): *Nordiska studiar i leksikografi 7. Rapport frá Konferanse om leksikografi i Norden Volda 20.–24. mai 2003*. Oslo, 228–236.
- Jónsson, Jón Hilmar (2008): Lexicographic description. An onomasiological approach on the basis of phraseology. I: Sandro Nielsen & Sven Tarp (red.): *Lexicography in the 21st Century. In honour of Henning Bergenholtz*. Amsterdam/Philadelphia: Benjamins, 257–280.
- Jónsson, Jón Hilmar (2013): Að byggja undir íslensk-erlenda orðabók. Forgreinng og orðabókarefni. I: *Orð og tunga* 15, 1–22.
- Jónsson, Jón Hilmar & Þórdís Úlfarsdóttir (2011): Íslenskt orðanet: Ett skritt mot en allmennspråklig onomasiologisk ordbok. I: *LexicoNordica* 18, 87–109.
- Orðabók Háskólans. Leksikografisk institutt ved Islands universitet* (1995): Ásta Svavarsdóttir (red.). Reykjavík: Orðabók Háskólans.
- Úlfarsdóttir, Þórdís (2013): ISLEX – norræn margmála orðabók. I: *Orð og tunga* 15, 41–71.

Jón Hilmar Jónsson
 forskningsprofessor
 Árni Magnússon-instituttet for islandske studier
 Islands universitet
 Neshagi 16
 IS-107 Reykjavík
 jhj@hi.is

En ”mer l. mindre stor” stor ordbok – om variationerna i SAOB:s omfång och ambitionsnivå

Lennart Larsson

The first section of The Swedish Academy Dictionary (SAOB), describing the Swedish language from 1521 to the present, was published already in 1893. When the dictionary is finished in a few years, it will comprise thirty-eight volumes. During this span of over 120 years, its level of ambition (and accordingly, its scope) has varied considerably. As shown in the paper, four periods can be distinguished: two periods (until c. 1920 and c. 1960–2002) when the ambition was to account for virtually all words and uses of words that were documented, and two periods (c. 1920–1960 and c. 2002 onwards) when quite strict rules of restriction steered the work. After summing up these rules, the paper gives examples of how the varying levels of ambition are reflected both in the selection of lemmas and in the (more or less exhaustive) way they are addressed.

1. Inledning

Svenska Akademiens ordbok (SAOB) utkom med sitt första häfte 1893 och befinner sig nu, 120 år och 36 tunga volymer senare, en bit in på bokstaven V – den kvalar således utan tvekan in i kategorin stora ordböcker. Och det gäller inte bara omfånget utan också ambitionerna: att beskriva i princip hela det (i skrift belagda) nysvenska ordförrådet från 1521 (det år då Gustav Vasas registratur tar sin början) och framåt.

SAOB är alltså vad man brukar kalla en diakronisk samtidsordbok, dvs. den behandlar ordförrådet från en viss tidpunkt i det förflutna och fram till den egna samtiden. Men när – som fallet är

här – utgivningen av en dylik ordbok sträcker sig över en längre tidsrymd så innebär det att den språkliga materia som ska beskrivas efterhand blir allt större. När SAOB började utges 1893 var det 372 år av språkanvändning som skulle beskrivas, medan det 2017, när ordboken enligt planerna ska avslutas, kommer att handla om 496 år; tidsspännet har alltså ökat med 124 år eller 33,3 %.

Den här utsträckningen i tiden innebär i sin tur att om samma principer för beskrivningen, samma begränsningsregler i förhållande till materialet skulle råda verket igenom, då skulle följd- en bli en successiv utvidgning av ordboken, helt enkelt därför att språkets förändringar ger ett allt större material att beskriva. I praktiken tillkommer dessutom ytterligare en faktor som verkar i samma riktning: i ett projekt av det här slaget kan man knappast räkna med att excerperingen av texter från äldre tider är fullständig redan från början, utan också mängden äldre språkprov kommer att öka.

Något mönster av det här slaget – alltså en sakta men säker utvidgning – finns emellertid inte alls i SAOB. I stället varierar dess omfång ganska kraftigt under olika perioder, och det är det jag kommer att syna närmare här: dels visa på hur variationerna ser ut i kvantitativa mått, dels ge vissa glimtar av vad som ligger bakom dessa. Men det rör sig bara om glimtar – ämnets storlek gör att en mera heltäckande och djupgående framställning skulle kräva ett helt annat utrymme.

2. SAOB i jämförelse med SAOL

Ett väl beprövat sätt att få grepp om hur språkbeskrivningen i SAOB omfångsmässigt förhåller sig till det språk som ska beskrivas är att jämföra med Svenska Akademiens ordlista (SAOL), en ordförteckning över samtidspråket som kan antas ge en god bild av hur svenskans ordförråd kvantitativt fördelar sig mellan de oli-

ka initialbokstäverna.¹ Som måttstock här har jag valt SAOL:s nionde upplaga från 1950, dels därför att den innehåller ett ovanligt stort antal uppslagsord, dels därför att den i tiden ligger ungefär mitt i SAOB:s utgivningsperiod. Resultaten av jämförelsen framgår av tabell 1.

	SAOL 1950 A–U		SAOB A–U				jämförelse SAOL/SAOB i procentenheter	
	antal sidor	i %	band	antal spalter	i %	år	SAOL större	SAOB större
A	29	5,00	1–2	2 780	4,50	1893–1903	0,50	
B	33	5,68	2–5	4 928	7,98	1898–1925		2,30
C	6	1,03	5	304	0,49	1903–1905	0,54	
D	20	3,44	6–7	2 624	4,25	1905–1923		0,81
E	14	2,41	7	880	1,42	1917–1922	0,99	
F	42	7,23	8–9	3 616	5,85	1919–1928	1,38	
G	21	3,61	10	1 764	2,86	1928–1929	0,75	
H	24	4,13	11–12	2 432	3,94	1930–1932	0,19	
I	18	3,10	12–13	1 256	2,03	1933	1,07	
J	6	1,03	13	504	0,82	1934	0,21	
K	49	8,43	14–15	3 880	6,28	1935–1939	2,15	
L	24	4,13	15–16	1 972	3,19	1939–1942	0,94	
M	30	5,16	16–17	2 168	3,51	1942–1945	1,65	
N	18	3,10	18	1 308	2,12	1946–1948	0,98	
O	24	4,13	18–19	1 776	2,87	1949–1951	1,26	
P	40	6,88	19–21	3 134	5,07	1952–1955	1,81	
Q	1	0,17	21	4	0,01	1955	0,16	
R	29	5,00	21–23	4 632	7,50	1955–1962		2,50
S	93	16,01	24–33	16 668	26,98	1963–2002		10,97
T	40	6,88	33–36	3 746	6,06	2002–2010	0,82	
U	20	3,44	36–37	1 398	2,26	2010–2013	1,18	

Tabell 1: Jämförelse mellan SAOL och SAOB.

1 En första jämförelse mellan SAOB och SAOL gjordes redan 1908 (se Sigurd 1986:149); för senare exempel se Sigurd (1986:161), Allén (1986:242) och Hast (1993:60f).

Som framgår av tabell 1 finns det två (här skuggade) perioder, dels fram till ca 1920, dels ca 1960–2002, då SAOB mer eller mindre markant tenderar att överskrida SAOL i omfång, och två perioder, dels ca 1920–1960, dels ca 2002 och framåt, då SAOB tvärtom är relativt mindre omfångsrik än SAOL.

Inte oväntat visar sig liknande mönster om man, som i tabell 2, sammanställer enskilda SAOB-artiklar från perioderna 1 och 3 (dvs. de mer omfångsrika) med synonymer eller antonymer eller på annat sätt jämförbara ord från period 2 (dvs. den första av de mindre omfångsrika; i tabellen anges också antalet sammansättningar som orden uppvisar i SAOL 1950).²

	period	år	antal spalter	antal ssgr (och deras spaltantal)	antal ssgr i SAOL 1950
AFTON	1	1896	13	135 (6,5 sp.)	37
KVÄLL	2	1938	7,5	80 (4 sp.)	39
DAG	1	1905	85	335 (33 sp.)	114
NATT	2	1946	46	593 (38 sp.)	245
LÖPA	2	1942	31	43 (3 sp.)	24
SPRINGA	3	1985	64,5	107 (18 sp.)	28
NORD	2	1947	12	160 (10 sp.)	75
SYD	3	2000	16	232 (14 sp.)	62
LIGGA	2	1940	27,5	44 (3 sp.)	32
SITTA	3	1968	62	96 (5,5 sp.)	38
LÄGGA	2	1941	48	13 (1 sp.)	2
SÄTTA	3	2002	122	65 (4 sp.)	14

Tabell 2: Jämförelse mellan SAOB-artiklar från olika perioder.

Ser man till spaltantalet i de i tabell 2 redovisade artiklarna kan

2 Det bör framhållas att tabell 2 inte gör anspråk på någon fullständig exakthet eller konsekvens. Ett av problemen är att man under period 1 var mycket generösare med att låta sammansättningarna bli egna artiklar; exempelvis har DAG drygt 30 sådana utbrutna sammansättningar, och med dem inräknade hade siffrorna skjutit i höjden ännu mycket mer.

man konstatera att utrymmet för de ord som hör hemma i perioderna 1 och 3 tenderar att vara omkring dubbelt så stort som för dem från period 2. En mera specifik slutsats man kan dra är att SAOB tar upp betydligt fler sammansättningar under perioderna 1 och 3: mest slående är kanske NORD och SYD där SAOL redovisar en större mängd sammansättningar under NORD – helt naturligt med tanke på vårt geografiska läge – medan SAOB däremot har 45 % fler sammansättningar under SYD än under NORD.

Men det handlar inte bara om sammansättningarnas antal. Om man jämför DAG med NATT och SPRINGA med LÖPA ser man att de medtagna sammansättningarna också bereds betydligt större utrymme: t.ex. har de 43 sammansättningarna under LÖPA pressats in på 3 spalter, medan de 107 på SPRINGA getts hela 18 spalter. Likaså kan man konstatera att huvudorden fått breda ut sig mer under perioderna 1 och 3: även med sammansättningarna borträknade är t.ex. DAG jämfört med NATT och SITTA och SÄTTA jämfört med LIGGA respektive LÄGGA påfallande mycket större – skillnader som inte kan förklaras på annat sätt än att man hade ganska vitt skilda redovisningsprinciper och ambitionsnivåer under de olika perioderna.³

3. Period 2 och 4: reglementerade begränsningar

En skillnad mellan å ena sidan perioderna 1 och 3 (då SAOB alltså är ”mer stor”) och å andra sidan perioderna 2 och 4 (då den är ”mindre stor”) är att bara de senare har explicit formulerade principer för ordboksarbetets omfattning. Bakgrunden till den första reglementeringen var att SAOB-projektet under 1910-talet befann sig i kris, ordboken stod i det närmaste stilla, det fanns ett högst

3 Att det finns ett dylikt mönster har förstås uppmärksamats förr; se t.ex. Sigurd (1986:161ff.), Allén (1986:241ff.) och Hast (1993:60f.) som behandlar de tre första perioderna.

reellt nedläggningshot, och för att avvärja detta lade en kommitté ledd av den nytillträdde ordbokschefen Ebbe Tuneld år 1920 fram ett betänkande (återgivet i Böök 1960:197–207) som blev grundvalen för den nystart som då tog sin början. Kring millennieskiftet 2000 blev det sedan åter dags att begränsa, och i ett ”PM angående SAOB:s uppgift och omfång” daterat 2003 är de principer som gäller för den fjärde perioden satta på pränt.⁴ För perioderna 1 och 3 saknas däremot något motsvarande: SAOB:s första 30 år ter sig när det gäller artiklarnas innehåll och omfång som ett ganska trevande och styrsellöst projekt, medan den radikala utvidgning av ordboken som tar sin början kring 1960 aldrig blir kodifierad och stadfäst. (I teorin – om än inte i praktiken – gällde alltså Tunelds begränsningsregler också under period 3. Lite ironiskt är att Böök just 1960 i sin minnesskrift över Tuneld nöjt konstaterar att ordboksredaktionen under ledning av dennes efterträdare Pelle Holm och sedan Sven Ekbo ”framgångsrikt kunnat genomföra programmet efter de uppställda riktlinjerna” (s. 180), liksom att han förutspår att Ekbo blir ”den förste i raden [av ordbokschefer], som efter all sannolikhet kan räkna med att få vara med om ordbokens avslutning” (s. 181). Som framgår av Mattisson (2002:121ff.) är det vid denna tid, under bokstaven R, som artiklarna börjar svälla ut på ett sätt som helt ställer Bööks förutsägelse på huvudet.)

I PM:et från 2003 formuleras ordbokens grundläggande uppgift på följande sätt:

Ordboken ska beskriva det under nysvensk tid allmänna och etablerade i ordförrådet, de i språkgemenskapen konventionaliserade formerna, betydelseerna och användningarna. Uppgiften får inte förskjutas till en noggrann redogörelse

4 Ett PM med samma rubrik och till stora delar samma innehåll föreligger redan från 1997. En redogörelse för och diskussion kring vissa av begränsningarna finns också i Lundbladh (2000).

av all iakttagen ordanvändning, till en beskrivning av enskilda kontexter eller tillfälligheter utan slutsatser om generella förhållanden.

Därefter ges ett antal mera handfasta och detaljerade förhållningsregler i syfte att begränsa omfånget. Till de viktigaste hör: det genomsnittliga relationstalet ska vara 8,5 (dvs. 1 000 lappar med språkprov ska motsvara 8,5 tryckta spalter); ett huvudmoment får i normalfallet innehålla högst 7 språkprov och ett undermoment 5; för att ett ord ska komma med krävs att det finns ”två av varandra oberoende språkprov i den egna materialsamlingen”; ord som bara finns i lexikon eller i encyklopedier ska normalt inte tas med, och detsamma gäller för tillfälliga och sekundära sammansättningar liksom för exklusivt fackspråk; införande av nya källor för enstaka belägg” ska bara ske om det är absolut nödvändigt”. Därtill framhålls följande:

Det är ytterst viktigt att kärnbetydelserna framhävs och inte överskuggas av mer eller mindre tillfälliga nyanser. [...] Större restriktivitet än på bokstaven S ska iakttagas i fråga om utbrytning av undermoment. [...] Ordboksartiklarna blir då också överskådligare och lättare att förstå för användarna.

Samma anda som i PM:et 2003 präglar utredningsbetänkandet från 1920 (Böök 1960:197):

Det ordförråd som skall medtagas måste inskränkas. [...] Den behandling, som kommer de upptagna orden till del, bör bli mera översiktlig. En mindre stark uppdelning av betydelserna bör genomföras; mindre viktiga nyanser i betydelse och bruklighet få uteslutas. [...] Slutligen måste i allmänhet knappas av på fordringarna på artiklarnas full-

ständighet. Det finns praktiskt taget ingen gräns för det arbete, som kan nedläggas på en artikels utarbetande.

Även om själva betänkandet inte innehåller några mer detaljerade anvisningar i syfte att begränsa omfånget så fanns det sådana formulerade: t.ex. gällde ett relationstal på 8 (jfr ovan), och reglerna för hur många språkprov per betydelse som fick citeras var betydligt striktare än tidigare (se Mattisson 2002:119).

Det finns alltså mycket som förenar perioderna 2 och 4: antalet redovisade språkprov är starkt begränsat och det är tydligt att många såväl ”mindre viktiga nyanser i betydelse och bruklighet” som tillfälliga sammansättningar har lämnats därhän under båda perioderna (jfr tabell 2). Men det finns också skillnader som faller i ögonen. En är att det uppenbarligen inte fanns några krav på två av varandra oberoende belägg under period 2: att artiklar byggde på ett enda språkprov var inte helt ovanligt. Till de mera frapperande exemplen hör LANNA (L 268) och sammansättningen LÅNG-VARDE (L 1445) där såväl etymologierna som betydelseorna (och i det senare fallet även själva lemmat) mest bara är frågetecken:

LANNA, f. [av okänt urspr.] (†) ofruktbar jord? En gårdh, som haffver sitt namn (*”Klippan”*) ther aff, at på samma Skogh äro många brante och stygge klippor och Bergh .., är mäst en onyttig marck och Lanna. GYLLENIUS *Diar.* 335 (c. 1670).

-VARDE? [möjl. felskrivning] (†) om person som väntat länge? *CivInstr.* 135 (1684).

I fall som dessa handlar det nog om att man i ordbokens tidigare historia hade en betydligt större reverens för enstaka lexikala egendomligheter från äldre tid. Men ibland måste bevekelsegrunderna sökas på annat håll. Så exempelvis i följande sex sammansätt-

ningar från artikeln NORD (N 657ff.), vilka samtliga är försedda med dödtecken. (Ord som anges som obrukliga kräver att både första- och sistabeläggen redovisas, vilket innebär att det bara fanns ett språkprov vardera i materialet.):

-FLICKA. (†) nordisk flicka. Hvad liknar i Södern en Nordflickas kyss?
PoetK 1816, 1: 172.

-GOSSE. (†) nordisk gosse. *PoetK* 1818, 1: 29.

-KVINNA. (†) nordisk kvinna. Så mild och stilla, / .. den redliga
Nordqvinnan är. *LING Riksd.* 34 (1817).

-KÄMPE. (†) nordisk kämpe. *LING Gylfe* 34 (1814).

-TÄRNA, f. (†) jfr -flicka, -mö. *LING Riksd.* 33 (1817).

-YNGLING. (†) nordisk yngling. *Euterpe* 1: 4 (1823).

Gemensamt för dessa sammansättningar är att de tillhör en och samma kulturella sfär, den romantiska göticism som var i svang under åren kring 1820 med Pehr Henrik Ling som portalfigur. Och det är svårt att inte sätta deras förekomst här i samband med de patriotiska ambitioner som följt ordboken alltsedan Akademiens grundande och som redaktören och excerperingschefen Theodor Hjelmqvist ånyo formulerade 1893: SAOB skulle vara ”en källa för vår fosterländska odling” genom att belysa hur ”det är nordman-nakraft och nordmannastål, som ännu i dag förnimmas i vårt modersmål” (citerat efter Loman 1986:115). Att dessa ambitioner ännu var vid liv på 1940-talet indikeras också av att den sammansättning på NORD som uppvisar flest sekundära sammansättningar inte är någon av de förväntade NORD-OST eller NORD-VÄST med 10 respektive 9 sådana, utan i stället NORD-MAN med inte mindre

än 11; även om Hjelmqvists ”nordmannastål” saknas så hittar man vid sidan av NORDMANNA-KRAFT också t.ex. NORDMANNA-HARPA, NORDMANNA-LYNNE och NORDMANNA-SÅTT (samtliga med bruklighetsangivelsen ”i sht i poesi o. vitter stil”).⁵

Ett annat för period 4 helt främmande inslag är det betydande antal ord, i första hand sammansättningar, som överhuvudtaget inte har belagts med språkprov. Ett exempel är den av ordbokschefen Tuneld skrivna artikeln FÅGEL (F 1968ff.) där 18 av de ca 190 sammansättningarna (alltså 10 %) saknar beläggsangivelser; allt ordboken har att förmedla när det gäller t.ex. FÅGEL-HONA, FÅGEL-PIP och FÅGEL-STJÄRT är själva lemmat. Huruvida ett dylikt tillvägagångssätt bidrar till att minska eller öka ordbokens omfång beror förstås på om orden är belagda i materialet eller inte. Några stickprovskontroller tyder på att det kunde variera. I artikeln FÅGEL handlade det tveklöst om att spara utrymme: av de 18 sammansättningarna var det bara två som saknade belägg i samlingarna. Mot slutet av period 2 tycks de obelagda sammansättningarna däremot – liksom under period 1 (jfr nedan) – i betydligt större utsträckning ha saknats i materialsamlingen. Det gäller t.ex. samtliga obelagda sammansättningar på PLAST (P 1072f.) och RAS (R 325ff.) (jfr Mattisson 2002:125) – och i sådana fall leder redovisningen av orden givetvis till en utvidgning av ordboken.

Samtidigt som SAOB under period 2 alltså har vissa tendenser att ta med sådant som ur ett kritiskt nutidsperspektiv förefaller något perifert eller överflödigt kan man – ur samma perspektiv – ibland önska att den varit betydligt mer utförlig än vad som är fallet. Ett exempel är deponens verbet *finnas* med dess nog så spretiga och mångfasetterade betydelser, där redaktören har valt att sammanföra allt under ett av momenten i artikeln FINNA (F 597). Definitionen lyder på följande sätt:

5 Ett annat intressant exempel på hur redaktörernas världsbild tycks ha styrt lemmaurvalet ger sammansättningarna i artikeln NEGER (se Larsson 2009:45ff.).

9) i pass. i deponentiell anv. (förr äv. refl.): vara att se l. påträffa; stå att få l. erhålla l. bekomma; vara, vara för handen; förefinnas; förekomma; föreligga; vara till, finnas till, existera; oftaopers. i förb. *det finnes* med följ. logiskt subj. (förr äv. *det finnas*, då det logiska subjektet står i pl.).

Den här lösningen sparar förstås både arbete och utrymme – där emot är det knappast något särskilt lyckat exempel på semantisk analys. Att upplägget sviktar framgår för övrigt av att redaktören, för att täcka in de många disparata betydelser och användningar som synonymdefinitionerna anger, har känt sig tvungen att anföra så många som 16 belägg i ett och samma moment, något som annars knappast torde förekomma under period 2. Hade artikeln FINNA skrivits under senare tid skulle den deponentiella användningen med all sannolikhet ha behandlats utförligare, antagligen som ett överordnat moment inlett av romersk siffra och uppdelat i ett antal huvud- och undermoment.

4. Period 1 och 3: Allt ska med!

Även om perioderna 1 och 3 uppvisar mycket stora skillnader sinsemellan så har de ändå ett grundläggande drag gemensamt: en stark strävan efter fullständighet, att allt ska med, också det mera perifera och tillfälliga – att formuleringarna i 1920 och 2003 års reglementeringar i mångt och mycket liknar varandra beror ju på att de var reaktioner på likartade fenomen.⁶

Samtidigt finns det förstås mycket som skiljer dessa två perioder åt. Något som i hög grad bidrar till det stora omfånget under

6 Denna strävan är förstås en viktig förklaring till varför period 3 sticker ut så kraftigt när det gäller omfånget, också i förhållande till period 1: en dylik fullständighet kräver ju större och större utrymme ju längre fram i SAOB:s historia man kommer.

period 1 är de många språkprov som redovisas. Den begränsning till sju språkprov för ett huvudmoment som är reglerat i PM:et 2003 och som i stor utsträckning också tycks ha gällt under perioderna 2 och 3 finns inte alls i SAOB:s begynnelse: t.ex. illustreras det första huvudmomentet på DANS (D 317f.) av hela 27 språkprov (vartill kommer ett antal redaktionsexempel, alltså språkprov som redaktören själv har konstruerat).

Ett annat särdrag för perioden är vad man skulle kunna kalla för bristande ekonomisering. Ett av uttrycken för detta är att samtliga sammansättningar utom de helt obrukliga är försedda med uttalsangivelser; fr.o.m. period 2 ges uttalet bara i sådana fall där efterledens betoningsmönster avviker från motsvarande simplex. Ett annat är att avledningar behandlas som egna huvudord. Exempelvis står inte bara AFSKILJARE, AFSKILJBAR och AFSKILJELIG utan också AFSKILJBARHET (för övrigt utan några belägg) och AFSKILJELIGHET som egna ord (A 500); från period 2 och framåt hade de här orden stått som avledningar inom artikeln AFSKILJA (om de överhuvudtaget hade tagits upp).

Ett tredje kännetecken för period 1 är de många och långa utvikingarna i form av anmärkningar som innehåller dels olika slags språkvetenskapliga specialutredningar, dels encyklopediska upplysningar (jfr Sigurd 1986:186ff.). Ett exempel på det förra är artikeln DIKTARE (D 1354f.) med sina två långa anmärkningar på sammanlagt nära två spalter som behandlar ordets bruklighet och förhållande till synonyma uttryckssätt – i och för sig intressanta utredningar men knappast något man numera träffar på (eller förväntar sig att träffa på) i en ordbok. Ett exempel på det mer encyklopediska kan man finna i artikeln DALER (D 199f.) där det första huvudmomentet avslutas med en exposé på en dryg spalt över myntets historik och värde under olika tider, en sakprosatext av ett slag som snarast ser ut att höra hemma i ett uppslagsverk som Nordisk familjebok.

Vid sidan av den ymnighet som exemplifierats ovan kan SAOB

i sin inledningsfas också uppvisa en påfallande knapphet. Man slås ofta inte bara av de kortfattade definitionerna – t.ex. avklaras betydelsebeskrivningen av det ovannämnda DALER på bara två rader – utan också av frånvaron av sådant som kollokationsangivelser och redovisning av stavningsvarianter; ett exempel på det senare är ADRESS (A 65ff.) där förekomsten av en stavning som ”adresse” endast framgår av de citerade språkproven.

Medan SAOB under period 1 i mångt och mycket ger intryck av en viss löslighet i konturerna, av att ännu inte riktigt ha hittat sin form, har den under period 3 däremot en synnerligen strikt och koncentrerad och formaliserad karaktär. Men samtidigt förenas alltså de två perioderna i sin strävan efter fullständighet; och därmed går det också – alla skillnader till trots – att urskilja vissa likheter dem emellan.

En likhet är att man, i sin önskan att vara uttömmande, tenderar att ge utrymme inte bara åt det faktiska utan också åt det blott möjliga eller hypotetiska språkbruket: man kompletterar det som är belagt med det som *skulle* kunna vara det. Hur det kunde te sig under period 1 får illustreras av de två sammansättningarna AFTON-SÖL och AFTON-TODDY (A 673):

-SÖL⁻². 1) söl l. senfärdighet om aftonen. *Hata aftonsöl o. morgonfjäsk*. jfr NATT-SÖL. 2) (†) dryckeslag, som hålles på aftonen. Abend-Zeche .. afton-söl, dryckesmål. LIND (1749).

-TODDY⁻²⁰.

När det gäller AFTON-SÖL bygger artikeln på ett enda belägg, anträffat i Linds tysk-svenska ordbok från 1749 och med en helt obsolet betydelse. (Till saken hör att man i SAOB:s tidigare historia lade en helt annan vikt vid ordboksbelägg än som numera är fallet: det är inte utan orsak som PM:et 2003 innehåller en särskild punkt om att ”ord som bara finns som lexikonbelägg” nor-

malt inte tas med (jfr ovan.) Det moment som bygger på detta enda språkprov har man sedan kompletterat med ytterligare ett, grundat på en i samtidsspråket möjlig (om än nog så tillfällig) användning av ordet; men eftersom några reella belägg inte fanns att tillgå fick det illustreras med ett redaktionsexempel. Inte heller för AFTON-TODDY (en kulturföreteelse som man anar låg redaktörerna varmt om hjärtat) fanns några belägg i arkivet, men här valde man – som ofta annars under period 1 – att nöja sig med att bara ge själva lemmat (kompletterat med den sedermera bortrationaliserade betoningsangivelsen, jfr ovan).

Att man i ordbokens inledningsskede tog till dylika lösningar skulle kunna förklaras av den bristfälliga excerperingen, att man ännu inte hunnit samla in några större mängder språkprov. Men också under period 3 träffar man på ord och betydelser som saknar reella belägg. Att man på det här sättet bara ger redaktionsexempel eller inte ens det är visserligen ganska sällsynt även om det förekommer (så t.ex. SPIREA-VÄXT (S 9789) som endast åtföljs av redaktionsexemplet ”Det finns många vackra spireaväxter”, samt de (utpräglat fackspråkliga) sekundära sammansättningarna SNÖRSTÄMPEL-ORNAMENT, SNÖRSTÄMPEL-ORNERING och SNÖRSTÄMPEL-TEKNIK (S 8447) där språkprov helt saknas). I stället utvecklade man under bokstaven S ett nytt grepp för att få fram belägg för sådant som man tyckte borde komma med i ordboken, ett grepp som följande citat får exemplifiera:

-BLUS, i sht *mil.* i snödräkt (se d. o. 1) ingående blus (se blus, sbst.¹ 1); jfr -byxor. SAOBArkSakkSvar (1977).

-BYXOR, pl. i sht *mil.* i snödräkt (se d. o. 1) ingående byxor; jfr -blus. SAOBArkSakkSvar (1977).

Särsk. förb. (till 1; fotogr.): SOLARISERA BORT¹⁰⁰¹⁰ 4. gm solarisering borttaga. SAOBArkSakkSvar (1979).

Ssg: sovjetforsknings-institut. institut för bedrivande av sovjetforskning, sovjetologinstitut. *SAOBArkSakkSvar* (1978).

2) [eg. specialfall av 1] om ansträngande marsch(prov) som soldater i förband förr hade att utföra i senare skedet av sin utbildning. Soldatmarsch .. var inget reglementerat men användes som ett (med dagens ögon sett) ”jippobetonat” kraftprov. Ganska disciplinerande eftersom alla måste göra provet. *SAOBArkSakkSvar* (1977).

4) *mus.* motsv. STILISERA 3 *b*, om jazzmusiker med god l. uppdriven stilkänsla inom den jazzstil som han l. hon representerar. *SAOBArkSakkSvar* (1986).

Gemensamt för de här exemplen är att de har en källa benämnd ”SAOBArkSakkSvar” (dvs. ”Sakkunnigsvar i SAOB:s arkiv”) som första och enda belägg. Vad det handlade om var att man skickade en fråga till en sakkunnig på området, och dennas svar fick bli det belägg som saknades. På så sätt kunde man komplettera ordboken med t.ex. sammansättningarna SNÖ-BLUS och SNÖ-BYXOR (S 8364ff.), den sekundära sammansättningen SOVJET-FORSKNINGS-INSTITUT (S 9055), den s.k. särskilda förbindelsen SOLARISERA BORT (S 8667), liksom de som egna moment redovisade specialbetydelserna av SOLDAT-MARSCH (S 8687) och STILIST (S 11773) – och allt detta utan att behöva ge avkall på kravet att de ord och betydelser som tas upp ska kunna beläggas.

Det här förfarings sättet, att man anför sakkunnigsvar som belägg i ordboken, förekommer i ca 280 fall mellan 1974 och 1999, och i nära 70 av dessa är sakkunnigsvaret – som i exemplen ovan – det enda belägget för ordet eller momentet i fråga.

En annan likhet mellan perioderna 1 och 3 handlar om det encyklopediska inslaget. Hur detta gör sig gällande också under den senare perioden kan illustreras av följande två definitioner, dels av SOUL-MUSIK (S 9016), dels av det andra betydelsemomentet i

den sekundära sammansättningen STAT(S)REGLERINGS-FOND (S 11146f.):

-MUSIK. [efter amerik.-eng. *soul music*] i sht *mus.* musik som har l. kännetecknas av soul; i sht inskränkta o. speciellare, om ett slags (av svarta musiker i USA omkr. 1955 introducerad o. 1969 ss. genre etablerad) av stark känsla o. intensivt samspel mellan artist(er) o. publik präglad (form av) populärmusik, som har sina rötter i gospelmusik (dvs. en i slutet av 1800-talet i USA framvuxen kristet religiös svart musikform som förmedlar evangeliets budskap gm hymnliknande melodier l. gm sånger i snabbt tempo med växelsång mellan predikant o. församling) o. i rytm-and-blues (utgörande en i amerikansk stadsmiljö uppkommen sammansmältning mellan blues (dvs. en i USA i början av 1900-talet ur starkt känslfull (ofta om (tragiska) personliga erfarenheter handlande) svart sång i kombination med europeisk harmonik utvecklad musikform) o. jazzmusik).

-fond. (om ä. förh. i Sv.) [...] 2) (om förh. 1875—1980, då ifrågavarande fond avskaffades) i sg. best., om den (gm sammanslagning år 1875 av ordinarie o. extra stat(s)regleringsfonderna (se 1) bildade) fond varpå utfallet av statens skattebudget l. (efter 1937) driftbudget (se under skatte-budget) samt medel av olika slag (inklusive tillgångar o. skulder) som icke kunde hänföras till någon kapitalfond i rikshuvudboken samlades; äv. i utvidgad anv., om konto(rubrik) l. fondtitel i rikshuvudbok varunder ovan angivna fond redovisades; jfr 3.

Liksom fallet var i t.ex. det ovannämnda DALER får vi här en nog så detaljerad skildring av dessa företeasers utvecklingshistoria. Men det finns en grundläggande skillnad: medan den encyklopediska

informationen under period 1 anträffas i separata anmärkningar av sakprosakaraktär utgör den under period 3 en integrerad del av definitionstexten. Särskilt lyckad kan denna återkomst för det encyklopediska knappast sägas ha varit. För det första kan man förstås ifrågasätta det rimliga i att så pass perifera ord får ett så stort spaltutrymme (sålunda upptar de tre momenten i STAT(S)-REGLERINGS-FOND – med den övergripande bruklighetsangivelsen ”numera bl. i vissa kretsar, särsk. med ålderdomlig prägel, o. i skildring av ä. förh.” – sammantaget nästan en hel spalt). För det andra lär det hur som helst inte vara i en ordbok som den intresserade söker upplysningar av det här slaget. För det tredje handlar det om SAOB-definitionernas – ur ett allmänspråkligt perspektiv – mycket särpräglade stil, där allt ska fogas in i *en* syntaktisk enhet, fylld av parentetiska inskott och underordningar. En sådan text är till sin natur inte särskilt lättläst; och används den för att på det här sättet presentera encyklopedisk information riskerar den att bli mer eller mindre oläslig (jfr Sigurd 1986:182).

5. Avslutning

Det är lätt att konstatera att om man faktiskt har ambitionen att SAOB inom överskådlig tid ska ha tagit sig igenom bokstaven Ö, då råder det ingen tvekan om att de begränsningsregler som finns formulerade i PM:et 2003 är nödvändiga (även om de förstås kan diskuteras i sina detaljer).

Dels blir ju det tidsspann som ordboken omfattar större för varje år – så för varje år finns det allt mer att beskriva.

Dels har vi sedan ett antal år tillbaka en närmast explosionsartad tillväxt av antalet presumtiva källor och språkprov. I och med internet har mängden tillgängligt språkmaterial i praktiken blivit oändligt, och skulle man ha fortsatt som under S, att man tog upp nära nog samtliga ord och användningar som man stötte på (och

mer än så ibland), liksom att man fick lägga till och utnyttja nya källor utan några egentliga inskränkningar, då skulle den ordbok vi arbetar på nu ha ett omfång som vida överstiger det som gällde för S – och den skulle bara ha hunnit en bit in på bokstaven T.

Källor och litteratur

- Allén, Sture (1986): Inför tredje seklet. I: Sture Allén, Bengt Loman & Bengt Sigurd: *Svenska Akademien och svenska språket. Tre studier*. Stockholm: Norstedts, 233–275.
- Böök, Fredrik (1960): Minne av Ebbe Tuneld. I: *Svenska Akademiens handlingar ifrån år 1886*, del 68, 1959. Stockholm.
- Hast, Sture (1993): SAOB och dess likar. De fyra stora historiska nyspråkliga germanska ordböckerna. En kort redogörelse och jämförelse. I: *Ord och lexikon. Festskrift till Hans Jonsson 10 juni 1993*. Red.: Lars Svensson, Ingrid Hansson & Bo Ruthström. Stockholm: Norstedts, 55–84.
- Larsson, Lennart (2009): Hur märker användaren av SAOB de 110 åren mellan A och TYNA? I: *Språk och stil* NF 19, 35–52.
- Loman, Bengt (1986): ”En Inrättning, ägnad endast til Språkets förbättring”. I: Sture Allén, Bengt Loman & Bengt Sigurd: *Svenska Akademien och svenska språket. Tre studier*. Stockholm: Norstedts, 1–141.
- Lundbladh, Carl-Erik (2000): När och hur kan SAOB bli klar? I: *Studier i svensk språkhistoria* 5. Förhandlingar vid Femte sammankomsten för svenska språkets historia Umeå 20–22 november 1997. Utg. av Lars-Erik Edlund (Nordsvenska 11), 287–293.
- Mattisson, Anki (2002): P och R i SAOB. I: *Alla ord är lika roliga. Festskrift till Lars Svensson 28 februari 2002*. Red.: Anki Mattisson, Per Stille, Gunilla Swietlicki & Bo-A. Wendt. Stockholm: Norstedts, 118–126.

PM angående SAOB:s uppgift och omfång. Otryckt källa, daterad
Lund 15 december 2003. I: SAOBArkiv.

SAOB = *Ordbok över svenska språket utgiven av Svenska Akademien*
(Svenska Akademiens ordbok) 1–, 1893–. Lund.

SAOBArkiv = Svenska Akademiens ordboks arkiv, Lund.

SAOBs = Svenska Akademiens ordboks samlingar, Lund.

SAOL = *Svenska Akademiens ordlista över svenska språket*, 1950.
9 uppl. Stockholm.

Sigurd, Bengt (1986): *Ordboken, ordlistan och några andra av Svenska Akademiens språkliga insatser under 1900-talet. I: Sture Allén, Bengt Loman & Bengt Sigurd: Svenska Akademien och svenska språket. Tre studier.* Stockholm: Norstedts, 143–231.

Lennart Larsson
universitetslektor, docent, redaktör på SAOB
Institutionen för nordiska språk
Uppsala universitet
Box 527
SE-751 20 Uppsala
lennart.larsson@nordiska.uu.se
lennart.larsson@svenskaakademien.se

Svenska Akademiens ordlista genom 140 år: mot fjortonde upplagan

Sven-Göran Malmgren

Physically speaking, the *Glossary of the Swedish Academy* (SAOL, *Svenska Akademiens ordlista*) is not an imposing dictionary. It is, however, a minor classic within Swedish lexicography. Beginning in 1874 with a small word list of about 35,000 lemmas, it is now a more dictionary-like product with about 125,000 entries (13th edition 2006). This paper gives a historical survey of the SAOL focused on the upcoming 14th edition (2015). Apart from the increase in the number of lemmas, the improvement of the semantic descriptions is worth noting. The main task of the SAOL is still to function as an orthographic and inflectional norm, but since the 10th edition, the need for more – and somewhat more elaborate – definitions has gradually grown stronger. In the 14th edition, this trend will continue. Thus, apart from a set of very basic (and monosemous) words, every entry word should have some kind of explanation, explicitly or implicitly (by means of references etc.). To achieve this within the limited space available, certain default conventions are necessary.

1. Inledning

Svenska Akademiens ordlista, eller SAOL, är inte en stor ordbok i fysisk mening. Men den är en svensk klassiker med många upplagor på nacken. Den trettonde publicerades 2006, och den fjortonde är på gång. SAOL tillhandahåller den allmänt accepterade normen för svensk stavning och böjning, och de numera ca 125 000 ordartiklarna ger ofta även annan nyttig information, t.ex. av semantiskt slag. En ny upplaga av SAOL väcker alltid ett betydande intresse, större än t.o.m. förstaupplagor av alla andra svenska ordböcker kan glädja sig åt.

I det följande ges först en kort historik över SAOL från upplaga 1 (1874) till och med upplaga 11 (1986). Sedan diskuteras, något mer ingående, de viktiga strukturella förändringar som genomfördes i SAOL 12 (1998) och konsoliderades i SAOL 13 (2006). Slutligen redogörs, i artikelns huvuddel, för de principiella nyheter som kommer att införas i SAOL 14 (2015).¹

2. Från begynnelsen till 1986

Den som är van vid nutidens bastanta SAOL-volymer kan finna de första sex upplagorna av SAOL (1874–89) förvånansvärt tunna. De innehöll mindre än eller omkring 40 000 lemman, att jämföra med över 50 000 i Dalins stora svenska definitionsordbok från 1850-talet. Delvis berodde det begränsade lemmaantalet på en negativ inställning till lånord. Vad SAOL kunde erbjuda – som inte fanns hos Dalin – var historiskt välgrundade böjningsmönster och stavningar; ordlistans initiativtagare Johan Erik Rydqvist var den historiska språkvetenskapliga skolans främste företrädare i Sverige.

Men detta historiska synsätt började på språkvårdens område komma i konflikt med ett mer pragmatiskt, ett synsätt som förespråkade en förenklad och mer ljudenlig ortografi. Inte minst aktiva var företrädare för skolvärlden, som framhöll elevernas problem med ljud som kunde motsvaras av tre, fem eller ännu fler olika bokstäver eller bokstavskombinationer (sj-ljudet var och är notoriskt i detta avseende).

SAOL:s sjunde upplaga (1900) innebar på flera sätt ett lyft, med bl.a. en mer liberal syn på lånord och därmed en kraftigt utökad lemmauppsättning. På böjningsmorfologins område innebar den en viss modernisering. Men den tillmötesgick inte kravet på en stavningsreform.

En sådan kom i stället sex år senare, men då genom ett rege-

¹ I en ganska färsk samlingsvolym, *SAOL och tidens flykt* (Gellerstam (red.) 2009), tas många olika aspekter på SAOL upp.

ringsdekret. 1906 års stavningsreform innebar en kompromiss mellan de radikala reformivrarnas krav och Svenska Akademiens och SAOL:s mer konservativa ståndpunkt. De viktigaste förändringarna gällde v- och t-ljudet, som nu fick en mer ljudenlig stavning (*af* > *av*, *godt* > *gott*, *hvilken* > *vilken*). Däremot lämnades alla de olika stavningarna av sj-ljudet oförändrade.

Den nya ortografin slog snabbt igenom, och därmed riskerade SAOL att förlora sin tidigare status som normgivare. Nästa SAOL-upplaga lät också vänta på sig, men när den kom 1923, visade det sig att dess rekommendationer i allt väsentligt stämde med 1906 års stavningsreform. Och alltsedan dess har SAOL:s roll som stavnings- och böjningsnormerare varit oomstridd.

Efter åttonde upplagan dröjde det åter ungefär ett kvartssekel till nionde upplagan (1950) och ytterligare 23 år till den tionde (1973). Med dessa upplagor hade SAOL kommit upp till sitt nuvarande omfång och väl det: lemmaantalet i SAOL 9 var över 155 000 och i SAOL 10 över 135 000. Noterbart var en liberalare inställning till slangord och en uppgradering av betydelseangivelserna. Denna tendens fortsatte i SAOL 11 (1986), där också tidigare tabubelagda (sexuella) ord fick plats.

Sammanfattningsvis kan det sägas, att de första upplagorna av SAOL i viss mån utgjorde försök att – med tillbakablickar på en tidigare (förment) språklig guldålder – motverka starka språkliga och sociala tendenser i tiden. Det yttrade sig i hävdandet av en historiskt motiverad ortografi (och i viss mån morfologi), liksom i inställningen till lånorden. Från och med åttonde upplagan har SAOL varit mer i harmoni med sin samtid. Visserligen höjs då och då röster för en ny stavningsreform, men SAOL:s och den svenska språkvårdens auktoritet har varit starka nog att ganska snabbt neutralisera sådana krav. Varje ny SAOL-upplaga innehåller ett betydande antal av de modernaste lånorden, och en av redaktionens viktigaste uppgifter är att i möjligaste mån anpassa dem till svenskan. En viktig princip är t.ex. att engelskt plural-s i möjligaste mån bör undvikas.

3. De lexikografiska förändringarna i SAOL 12

Från lexikografisk synpunkt förändrades inte mycket mellan SAOL 1 och SAOL 11. Alla dessa upplagor är nischalfabetiska, dvs. sammansättningar och avledningar återfinns normalt i samma stycke som sina resp. grundord (förleder). Detta kan givetvis medföra att den alfabetiska ordningen bryts, inte minst när förleden är homograf. Här följer ett antal artiklar i SAOL 11 med utgångspunkt i artiklarna ¹val (da., no. *valg*) och ²val (da., no. *hval*):

¹val -et, pl. = • *vara i valet och kvalet*
 -affisch -agitation -allians -apparat
 -appell -arbetare -beredning /.../

²val -en -ar däggdjur /.../ -bard -djur -fisk
 /.../ -fångare -fångst /.../

Figur 1: Ur SAOL 11 (något förenklat).

Det är uppenbart att den nischalfabetiska uppställningen hade sina fördelar. SAOL-användaren fick t.ex. omedelbart information om att *valaffisch*, *valallians* och *valarbetare* hade att göra med väljande. Ännu bättre hade det visserligen varit om det hade framgått att alla orden under ¹val i figuren refererar till företeelser i samband med *politiska* val, i motsats till idiomet *vara i valet och kvalet*; kanske hade även detta varit möjligt med i stort sett bevarad utrymmesekonomi? Det antyds också av figuren att ¹val genererar långt fler sammansättningar än ²val, något som skulle kunna ha betydelse för statistiskt baserad gissning av okända sammansättnings betydelse.

Men den nischalfabetiska makrostrukturen hade en stor nackdel: många SAOL-användare hade svårt att hitta det ord de sökte. Den som av någon anledning letade efter ordet *valbard* och först fick syn på ordet *valberedning* kanske inte kom på att det gällde att

gå nedåt, inte uppåt, i ordlistan. Efter att ha fått svara på många frågor med anknytning till detta problem beslöt ordlistans dåvarande huvudredaktör Martin Gellerstam att införa en rent alfabetisk makrostruktur i SAOL 12. Så här blev resultatet, vad våra *val*-ord beträffar:

¹val s. -et, pl. = • *vara i valet och kvalet*
²val s. -en -ar däggdjur /.../
 val|af·fisch s.
 val|ag·it·at·ion s.
 /.../
 val|fångst s.
 val|för·rätt·are s.

Figur 2: Ur SAOL 12 (något förenklat).

Det är onekligen lättare att hitta i SAOL 12 än i SAOL 11, men den ökade tillgängligheten har i viss mån uppnåtts på bekostnad av värdefull information. Det finns inte längre något som antyder att *valaffisch* och *valagitation* hör samman med ¹val, eller att *valfångst* och *valförrättare* har olika förleder.² Denna informationsförlust påtalades i två recensioner, Reuter (1999) och framför allt Lundblad (2000). Som framgår av Lundbladhs recension, kunde den förlorade informationen ganska lätt ha återställts, något som dock inte skedde i SAOL 13 (2006).

En viss ersättning för den förlorade informationen får man genom de lodstreck som i SAOL 12 och 13 utmärker den stora gränsen i sammansatta ord (se *valaffisch* och andra sammansatta ord i figur 2). De små punkterna i flera av uppslagsorden utmärker övriga morfologiska gränser.

2 Det kan hävdas att den information som gått förlorad i detta fall är umberlig för de flesta svenskar. Det finns dock många exempel på att också svenska användare kan tveka om förledens betydelse, och dessutom bör SAOL även tillmötesgå behoven hos användare med annat modersmål än svenska (jfr Allén 2009:172).

En annan stor förändring i SAOL 12 gäller principerna för upp- rättande av lemman. I SAOL 1–11 hade etymologin varit avgörande, och den kunde även leda till att ordklassgränser överskreds (ljus, adjektiv och substantiv, var bara ett lemma). Från och med SAOL 12 tillämpas strängt formella principer; ord från olika ordklasser (t.ex. *ljus*) blir i princip alltid två lemman, och ord (”homografkandidater”) som har samma stavning, böjning och uttal förs till samma lemma. Ett exempel är *by* (’samhälle’; ’vindkast’), som blir ett lemma fastän vi har att göra med två etymologiskt skilda ord. (Om denna modell, den s.k. lemma-lexemmodellen, se t.ex. Allén 1970:XVIIff.)

I SAOL 13 genomfördes inga egentliga strukturella förändringar. Viktig var en större sovring bland mer eller mindre obrukliga ord, av vilka många kanske slentrianmässigt hängt med från upplaga till upplaga (se Gellerstam 2004). Denna sovring underlättades av starkt förbättrade datamaskinella hjälpmedel som snabbt kunde tillhandahålla viktiga frekvensuppgifter. Samma hjälpmedel möjliggjorde också rationellare sökning efter nya ord. Förutsättningen för dessa hjälpmedel var i sin tur den s.k. SMDB (Svensk morfologisk databas, med hela ordförrådet indelat i böjningsklasser), som också möjliggjorde större konsekvens i böjningsangivelserna (se t.ex. Berg 2009).

SAOL 13 hade också en något tydligare språkvårdsprofil än tidigare upplagor. När två konkurrerande ordformer förelåg, kunde den ena prioriteras ner genom att föregås av angivelsen ”hellre än” (tidigare ”äv(en)”).

SAOL 13 recenserades utförligt av Nina Martola i *LexicoNordica* (Martola 2007). Bland Martolas synpunkter märktes kanske främst hennes förslag om förbättrad behandling av partikelverben (jfr nedan).

4. Elektroniska biprodukter av SAOL 13

Med utgångspunkt i SAOL 13 har flera elektroniska produkter skapats. Den produkt som har fått störst spridning är den (kostnadsfria) mobilapp som släpptes 2011 och som har funnit ca en halv miljon användare. I appen finns all information från pappersversionen med, och dessutom mycket mer. Exempelvis ges fullständig böjning på alla relevanta ord, inte bara det urval av former som brukar ges i pappersordböcker. Denna komplettering av böjningsformer var långtifrån trivial; exempelvis krävdes många beslut huruvida passiv- eller perfektparticipformer existerade eller ej (finns *regnas*, *regnad*?). SAOL-appen innehåller dessutom bonusmaterial, bl.a. en korsordshjälp byggd på SAOL 13.

5. Mot SAOL 14

I SAOL 14 kommer följande förändringar att genomföras (jämfört med pappersversionen av SAOL 13):

1. Alla substantiv, verb och adjektiv får böjningsuppgifter.
2. Sambandet mellan sammansättningsled och grundord återställs (jfr ovan, Lundbladh 2000).
3. Utökade betydelseangivelser; utökad differentiering (med momentsiffror) av betydelser för att möjliggöra hänvisningar.
4. En ny makrostrukturell kategori, *sublemman*, etableras för s.k. löst sammansatta partikelverb (*tycka om*, *komma ut* etc.); jfr ovan, Martola (2007).
5. Uppgifter om tonaccent (akut och grav accent) ges.
6. Vissa namn tas upp, bl.a. alla namn på länder.

5.1. Böjningsangivelser

I pappersversionen av SAOL 13, och i alla tidigare SAOL-upplagor, gavs i princip böjning bara för enkla substantiv och adjektiv, t.ex. för *medicin* men inte för *hostmedicin* och för *röd* men inte för *blåröd*. I SAOL 14 kommer alla, även sammansatta, substantiv och adjektiv att få böjningsuppgifter. Det är inte bara en service till läsaren som kanske slipper göra en extra uppslagning, utan det innebär också ett avsevärt informationstillskott. Det är nämligen inte alls säkert att alla sammansättningar böjs på samma sätt som de enkla ord som utgör deras efterleder. Det framgår t.ex. knappast i SAOL 13 att *hostmedicin* kan stå i plural, men inte *allmänmedicin*. Också de sammansatta verben, som i en del fall fick explicita böjningsuppgifter i SAOL 13, kommer att få fullständigare angivelser i SAOL 14.

5.2. Hänvisningar till för- och efterleder

Vi återvänder till orden med förleden ¹val och ²val i avsnitt 3. Vi såg där att den nischalfabetiska makrostrukturen i SAOL 11 hade vissa fördelar som gick förlorade vid övergången till en rent alfabetisk struktur i SAOL 12. Frågan är om det går att förena den rikare informationen i den tidigare upplagan med den ökade tillgängligheten i den senare. Svaret framgår av figur 3, där också de fullständigare böjningsuppgifterna visas:

¹val s. ~et, pl. ~ 1 allmänt: det att välja;
vara i valet och kvalet inte kunna bestämma
 sig ² utseende av politiska representanter
 genom allmän röstning: *allmänna v.*;
valets segrare – Nästan alla sms med *val-*
 hör till ¹val 2.

²val s. ~en ~ar 1 ett däggdjur /.../
 val|af·fisch s. ~en ~er
 val|ag·it·at·ion s. ~en ~er
 val|al·li·ans s. ~en ~er
 val|ap·par·at s. ~en ~er
 /.../
 val|fångst s. ~en ~er • till ²val 1
 val|för·lust s. ~en ~er
 /.../
 val|situ·at·ion s. ~en ~er • till ¹val 1

Figur 3: Ur databasen till den kommande SAOL 14.

Den givna lösningen på problemet är alltså att använda sig av hänvisningar till numrerade lemman och betydelsemoment. Dock tillåter utrymmet i pappersordboken kanske inte explicita hänvisningar varje gång de egentligen skulle behövas (vid tiotusentals sammansättningar). Men många av dessa hänvisningar kan sparas in genom en liten upplysning vid grundordet. I det aktuella fallet får läsaren informationen att om inget annat anges, så hör sammansättningen till ¹val 2. Man kan alltså säga att orden *valaffisch*, *valagitation*, *valallians*, *valapparat* och *valförlust* är försedda en osynlig uppgift om att de hänför sig till politiska val. Däremot måste *valfångst* och *valsituation* få explicita angivelser; i det senare fallet handlar det inte längre om ett politiskt val.

Därmed är den förlorade informationen återställd och mer till; i SAOL 11 gjordes ju ingen skillnad mellan de två betydelsemomenten av ¹val. Det ursprungliga problemet är löst, men en berättigad fråga kvarstår: om man ger så pass noggranna hänvisningar till förlederna, varför inte också till efterlederna? Detta har aldrig tidigare gjorts i SAOL (och knappast helt konsekvent i någon annan svensk ordbok), och det är behäftat med betydligt större problem. Eftersom sammansättningarna med en given efterled kan vara spridda

över hela alfabetet, är det ganska poänglöst att – så som gjordes i artikeln 'val – ange att den eller den betydelsen av efterleden är vanligast. Och det finns ungefär 30 000 sammansättningar med flertydiga efterleder i SAOL, varför det skulle krävas alltför mycket utrymme för att ge explicita hänvisningar vid dem alla.

Lösningen på problemet bygger på en enkel iakttagelse: när en efterled är flertydig, och teoretiskt kan höra t.ex. till betydelsemoment 1, 2 och 3 enligt SAOL:s analys, så är det i verkligheten i fyra fall av fem fråga om betydelsemoment 1. Detta är ett intressant faktum i sig, och dessutom lexikografiskt fördelaktigt. Om man godtar konventionen att avsaknad av hänvisning är liktydigt med hänvisning till betydelse 1, så kan man därför spara ca 80 procent av de explicita hänvisningarna från de flertydiga efterlederna. Exempelvis är ordet *föreställning*, som liksom danskans och norskans *forestilling* har två betydelser (se figur 4), efterled i ca 40 sammansättningar i SAOL 14. Här visas ett urval av motsvarande artiklar jämte artikeln om grundordet:

före|ställ·ning s. ~en ~ar 1 uppförande av pjäs e.d.:
*en sceniskt lyckad f. 2 idé, tankebild: det är svårt
 att göra sig en f. om Rysslands storlek*
 fam·ilje|före·ställ·ning s. ~en ~ar
 far·väl|före·ställ·ning s. ~en ~ar
 fest|före·ställ·ning s. ~en ~ar
 film|före·ställ·ning s. ~en ~ar
 gala|före·ställ·ning s. ~en ~ar
 guds|före·ställ·ning s. ~en ~ar • till *föreställning* 2

Figur 4: Ur databasen till den kommande SAOL 14.

Precis som väntat utgör fem av de första sex sammansättningarna exempel på någon form av sceniskt evenemang, dvs. på betydelse 1 av *föreställning*. Hänvisningen blir därför ”osynlig”. Det sjätte

ordet, *guds föreställning*, betecknar inte en teaterpjäs om utan en mental bild av Gud. Det måste därför ha en explicit hänvisning till betydelse 2.

På det här sättet kan i princip alla de mer än 50 000 sammansättningarna i SAOL med flertydig för- och/eller efterled disambigueras.³ Förutsättningen är givetvis att artiklarna om grundorden, alltså de enkla orden, är väl strukturerade, med numrerade betydelsemoment. Disambigueringen av *familjeföreställning*, *guds föreställning* etc. var inte möjlig i SAOL 13, där betydelsemomenten behandlades inom ramen för ett betydelsemoment. Ännu mindre var en liknande disambiguering möjlig när betydelsemomenten av för- eller efterlederna betraktades som allmänt bekanta och därför inte fick någon definition alls. Om detta i nästa avsnitt.

5.3. Mer om utökade betydelseangivelser

SAOL är som framgår av namnet en *ordlista*, inte en fullfjädrad definitionsordbok. Ofta ges enbart synonymer i betydelsebeskrivningarna, och traditionellt har definitioner ofta utelämnats helt vid de vanligaste och allmänt bekanta orden, t.ex. *musik*. Så kommer också att ske i SAOL 14, med ett viktigt förbehåll: flertydiga ord, främst substantiv och adjektiv, får alltid förklaring och normalt minst två numrerade betydelsemoment.⁴ Exempelvis förklaras inte ordet *musik*, som är entydigt, men däremot ordet *affär*, som är flertydigt ('butik'; 'ekonomisk transaktion' m.fl. betydelse) och ingår i ett stort antal sammansättningar som därför behöver disambigueras.

3 Detta är egentligen en förenkling. En sammansättning kan vara flertydig även om både för- och efterleden är entydiga. Det kan finnas flera möjliga semantiska relationer mellan förleden och efterleden. Detta problem kan inte alltid lösas i SAOL 14. Jfr även nästa avsnitt.

4 Intressant nog tillämpades samma princip i tidigare upplagor av *Nudansk ordbog* (se Becker-Christensen 2013:84). Det bör tilläggas att de allra centralaste verben i allmänhet får en mer rudimentär betydelseanalys i SAOL 14, liksom i tidigare upplagor. De ingår ju också mera sällan i sammansättningar.

Som nämndes tidigare, kan en sammansättning behöva förklaras, även om både för- och efterleden är entydiga eller disambiguerade. Omvänt kan en förklaring av en sammansättning göra en hänvisning till rätt betydelse av för- eller efterleden onödig. En viss betydelse av för- eller efterleden kanske dessutom bara förekommer i en enda sammansättning, och då är det knappast motiverat att upprätta den i artikeln om det enkla ordet. Ordet *dörrspegel* betecknar inte en vanlig (reflekterande) spegel utan ett avgränsat fält (oftast av trä liksom den övriga dörren). Den betydelsen återkommer inte i någon annan *spegel*-sammansättning. Då är det förmodligen bättre att ge en förklaring i artikeln *dörrspegel* än att ansätta en ny betydelse (2) i artikeln *spegel* och hänvisa till *spegel* 2 från *dörrspegel*. På samma sätt betecknar älgtest ett avancerat test som inte direkt har med (djuren) älgar att göra. Det är då knappast motiverat att ansätta en ny betydelse av det enkla ordet *älg*.

Något bör också sägas om relationerna mellan grundord och avledning. I de nischalfabetiskt ordnade SAOL-upplagorna var detta kanske det snårigaste kapitlet. Likafullt fick användaren värdefulla upplysningar om samband som var allt annat än glasklara (jfr Lundbladh 2000). Att *diskussion* hör ihop med *diskutera* är säkert känt av nästan alla, men hur många är säkra på sambandet mellan *digerera* och *digestion*? Här bör man också skilja mellan passiv och aktiv kunskap: av dem som förstår *digestion* (givet att de förstår *digerera*) kan nog inte alla bilda substantivet utifrån verbet. De äldre SAOL-upplagorna tillhandahöll båda dessa kunskaper, men inte SAOL 12 och 13. I SAOL 14 kommer i varje fall det passiva kunskapsbehovet att tillgodoses (med hänvisning t.ex. från *digestion* till *digerera*), möjligen också det aktiva (t.ex. med en angivelse ”jfr *digestion*” i artikeln *digerera*).

Till sist ska ytterligare en liten komplettering av betydelseangivelserna i SAOL 13 nämnas. Det rör sig om vissa semantiska analogier som tidigare varit underförstådda. Om man förstår det viktiga lagspelsordet *hemmaplan*, så bör man också, utan någon längre

förklaring, förstå ord som *hemmamatch*, *hemmaspelare* och *hem-maseger*. Men det är fördelaktigt om den gemensamma idén hos dessa ord framgår av ordlistan. I SAOL 14 kommer *hemmaplan* att förklaras, och i artiklarna om *hemmamatch* etc. kommer hänvisningen ”jfr *hemmaplan*” att stå. Artikeln *hemmaplan* blir en s.k. synopsisartikel. Samma teknik används också i mer komplicerade fall. Om ordet *matchboll* är förklarat, räcker det med hänvisningen ”jfr *matchboll*” i artiklarna *setboll* och *gameboll*.

5.4. Behandlingen av partikelverb

I sin recension av SAOL 13 (Martola 2007:252) skriver Nina Martola: ”Om jag av någon anledning skulle slå upp *torka av*, så skulle jag knappast kontrollera under *avtorka* /.../”. Så är det onekligen, men i SAOL 13 liksom i alla tidigare SAOL-upplagor är artikeln *avtorka* den enda ingången till *torka av*, via en svårtydd upplysning om att *avtorka* även förekommer i s.k. lös förbindelse (eller lös sammansättning). I eller nära artikeln *torka* står ingenting om *torka av*, och man får heller inte veta att *avtorka* i motsats till *torka av* är praktiskt taget obrukligt.

När ett partikelverb enbart förekommer i lös förbindelse, som t.ex. *hålla med* (’samtycka’) och *ligga i* (’arbeta hårt’), saknades det ofta i tidigare SAOL-upplagor. I bästa fall kunde det komma med som ett syntaktiskt exempel i kursiv stil, t.ex. *tycka om* under *tycka*.

Allt detta är i enlighet med en gedigen SAOL-tradition, som säkert bottenar i en föreställning om att ett uppslagsord (lemma) helst inte bör bestå av flera ord. Som Martola riktigt påpekar, fungerade SAOL:s modell för partikelverb hyggligt långt fram på 1900-talet, när de fast sammansatta varianterna av partikelverben ännu inte var så stilistiskt markerade som nu. Men när många av dem t.o.m. börjar bli obrukliga (som *avtorka*) är det kanske dags att hitta på något nytt.

För att i viss mån anknyta till SAOL-traditionen, har det be-

funnits lämpligt att ta upp de löst sammansatta partikelverben inte som fullvärdiga lemman utan som *sublemman*, något som också föreslås av Martola. Ungefär så här kommer artiklarna om *torka* och tillhörande löst sammansatta partikelverb att se ut i SAOL 14 (**torka* är ett substantiv):

**torka* v. *~de ~t* • få att bli torrare
torka av • torka ordentligt: *t. el. (sällan)*
avtorka bordet
torka in • 1 fastna efter att ha torkat
 2 bli inställd: *festen torkade in*
torka inne • bli inställd
torka upp • torka ordentligt

Figur 5: Ur databasen till SAOL 14.

Partikelverbens sublemmastatus markeras med ett litet indrag och med att de förutsätts ha samma böjning som det enkla verbet. Så snart det finns en fast sammansatt motsvarighet, markeras det som i artikeln *torka av*. Frekvensförhållandena mellan den lösa och den fasta varianten anges också. I artikeln *avtorka* ges samma upplysning, fast spegelvänt: *avtorka* el. vanl. *torka av bordet*.

Drygt 2000 löst sammansatta partikelverb kommer att tas upp som sublemman i SAOL 14. Självfallet är det ändå bara fråga om ett urval.

5.5. Uppgifter om tonaccent

I likhet med de flesta andra svenska ordböcker har SAOL hittills inte tillhandahållit uppgifter om tonaccent. Artiklarna *Uppland* (ett svenskt landskap) och *uppland* (område som är knutet till en tätort) innehåller samma uttalsangivelse i SAOL 13, [up´-]. Men de båda orden uttalas inte på samma sätt; det första har akut accent

och det andra grav. Även om den musikaliska accenten inte spelar någon stor roll för att skilja svenska ord åt (det finns några hundra minimala par), så är den en del av det svenska uttalssystemet och det är svårt att se något skäl till att utelämna uppgifter om den. Oftast kräver det ingen extra plats, t.ex. i fallet *Uppland/uppland* ovan (det är bara att vända på det ena accenttecknet). Däremot skulle det vara alltför redundant att ge accentuppgift för varje uppslagsord; av de över 80 000 sammansättningarna i SAOL har kanske 99,9 procent grav accent, och det stora flertalet av dem har betoning på första stavelsen.

Detta kan utnyttjas i följande konvention: med ett fåtal angivna undantag, som berör några tusen ord,⁵ ses flerstaviga ord med akut accent, och/eller med accent på senare stavelse än den första, som undantag från huvudregeln. Enbart dessa ord behöver då accentuppgift. En viktig grupp är ord på *-el*, *-er* och *-en*, av vilka en stor majoritet (mot huvudregeln) har akut accent. Exempelvis har *nyckel* grav accent (och behöver ingen accentuppgift), medan *cykel* har akut accent (och därför får accentangivelse).

På det här sättet får samtliga uppslagsord i SAOL 14 en tidigare inte meddelad upplysning om musikalisk accent, till priset av en mycket liten ökning av omfånget.

5.6. Namn

Som framgår av föregående avsnitt togs ett fåtal namn upp i SAOL 13, nämligen svenska landskapsnamn. I SAOL 14 kommer ytterligare några hundra namn att tas med, främst landsnamn men även vissa andra namn. Den bakomliggande tanken är att ord som används i definitionerna, med undantag för helt transparenta sammansättningar, själva ska förklaras i ordlistan. Om man läser

5 Framför allt handlar det om verb med de obetonade prefixen *be-* och *för-*, som nästan alltid har akut accent och betoning på andra stavelsen. Detta kommer inte att anges vid varje ord.

definitionen ”person från Rwanda” i artikeln *rwandier*, så ska man inte behöva slå i ett annat uppslagsverk för att få veta vad *Rwanda* är för något. I artikeln *Rwanda* kommer det helt enkelt att stå ”ett land i Afrika”.

Att ta med landsnamn har också en annan fördel. Antag att man vill veta vad invånarna i Elfenbenskusten kallas. Det är då värdefullt att få hänvisningen ”jfr *ivorian*” i artikeln *Elfenbenskusten*. Det är ett extremt exempel, men liknande hänvisningar är också värdefulla från *Polen* till *polack* och *polska*, från *Frankrike* till *fransman* och *fransyska* och t.o.m. från *Åland* till *ålänning*. I samtliga dessa fall står artiklarna om invånarbeteckningarna långt från artikeln om landsbeteckningen; användaren är betjänt av en extra ingång.⁶

6. Slutord

Utrymmet har bara tillåtit en kortfattad redovisning av nyheterna i SAOL 14. Det bör ändå tilläggas, att en större inventering av sammansättningar i de svenska textmaterialen (SAOL-kandidater) genomförs. Det bör också sägas, att den pappersversion av SAOL 14 som stått i fokus i den här artikeln med all säkerhet kommer att följas av elektroniska utgåvor. I sådana versioner av SAOL 14 blir de utrymmesproblem som berörts i artikeln givetvis mindre kännbara.

6 En variant av denna idé tillämpas i den svenska Skolordlistan (se Malmgren 2006).

Litteratur

Ordböcker

- Dalin, A.F. (1850–53): *Ordbok öfver Svenska språket*. Stockholm.
Nudansk ordbog (1967). 5. uppl. København.
 SAOL = *Svenska Akademiens ordlista* (1874–2006). Uppl. 1–13.
 Stockholm.
Svensk skolordlista (2004). Utarbetad av Svenska Akademien och
 Svenska språknämnden. 4. uppl. Stockholm.

Övrig litteratur

- Allén, Sture (1970): *Nusvensk frekvensordbok*. 1. Graford. Stockholm.
 Allén, Sture (2009): Framsyn. I: Gellerstam (2009), 167–173.
 Becker-Christensen, Christian (2013): *Nudansk ordbog 1953–2010*.
 I: *Danske studier*, 66–96.
 Berg, Sture (2009): Om ordböjning och SAOL Plus. I: Gellerstam
 (2009), 139–165.
 Gellerstam, Martin (2004): Ord som går i pension. Om arbetet
 med att rensa bort gamla ord i SAOL. I: Ruth Vatvedt Fjeld &
 Dagfinn Worren (red.), *Rapport fra Konferanse om leksikografi
 i Norden, Volda 20.–24. mai 2003*. Oslo, 133–139. (Nordiske stu-
 dier i leksikografi 7.)
 Gellerstam, Martin (red., 2009): *SAOL och tidens flykt*. Stockholm.
 Lundbladh, Carl-Erik (2000): Ord- och ordledsidentifiering i
 SAOL. I: *Språkvård* 1, 6–14.
 Malmgren, Sven-Göran (2006): Skolordlistan i ny tappning. I:
LexicoNordica 13, 261–271.

Martola, Nina (2007): Den trettonde upplagan av SAOL. I: *LexicoNordica* 14, 245–266.

Reuter, Mikael (1999): SAOL 12. I: *LexicoNordica* 6, 211–226.

Sven-Göran Malmgren
professor
Lexikaliska institutet
Institutionen för svenska språket
Göteborgs universitet
Box 200
SE-405 30 Göteborg
malmgren@svenska.gu.se

Moths ordbog og dens virkningshistorie

Marita Akhøj Nielsen

Matthias Moth (1649-1719) worked out the earliest comprehensive dictionary of Danish, based on collections of words and sayings, predominantly of ordinary people's daily language. The final edition contains 109,096 headwords. Danish is the core of the dictionary, but is accompanied by Latin equivalents. The dictionary was not published until 2013 and is now online. During 300 years the manuscript has served as an important source for later Danish dictionaries, and Moth's impact is examined in the paper, especially with regard to *Videnskabernes Selskabs Ordbog* (1793-1905), Chr. Molbech's *Dansk Ordbog* (1833), Otto Kalkar's *Ordbog til det ældre danske sprog* (1881-1918) and *Ordbog over det danske Sprog* (1918-56).

1. Præsentation

Moths ordbog er et begreb i dansk sprogvidenskab. Værkets forfatter er udråbt til "den danske Lexikographies Fader" (MO 1833:VII), og "alle senere større danske ordbøger står i dyb gæld til det" (Lindegård Hjorth 1983:26). Det er hensigten med denne artikel at belyse Moths indflydelse på fire af disse storværker: *Videnskabernes Selskabs Ordbog* (VSO 1793-1905), Christian Molbechs *Dansk Ordbog* (MO 1833), Otto Kalkars *Ordbog til det ældre danske sprog* (KalkO 1881-1918) og *Ordbog over det danske Sprog* (ODS 1918-56) med supplement (ODS-S 1992-2005). Anledningen er den glædelige, at det omsider efter flere forgæves forsøg er lykkedes at publicere Moths ordbog. I 2011 bevilgede Carlsbergfondet penge til en digital udgave, og i oktober 2013 offentliggjorde Det Danske Sprog- og Litteraturselskab (DSL) en betaversion af en onlineudgave (Moths Ordbog 2013), siden fulgt op af apps. Den har vakt

betydelig opmærksomhed i de danske medier, og ved lanceringen var der if. Google Analytics mere end 40.000 opslag på et døgn; for tiden er der ca. 400.

Ordbogen har navn efter sin forfatter, Mathias Moth (1649-1719), som fra 1670'erne og indtil Christian V's død i 1699 beklædte en række høje embeder. På toppen af sin karriere, i 1697, iværksatte han en storstilet indsamling af danske ord, hvis mål var en stor national ordbog med en fuldstændig beskrivelse af samtidens ordforråd i alle sociale lag og i hele det dansk-norske rige. Instruktionen til optegnerne fremhævede specielt bønders og håndværkeres fagsprog, og det indskærpes, at man skulle optegne både ordenes betydning, vendinger og faste udtryk.

Der indkom et omfattende materiale, som Moth selv omhyggeligt gennemgik. Ved Christian V's død i 1699 mistede han alle sine poster, og i de næste mange år var han beskæftiget med ordbogsarbejdet (se Eegholm-Pedersen 1997 og 2013). Foruden hans endelige redaktion, som er emnet for det følgende, er der bevaret en række prøveredaktioner og andre forarbejder, det hele nu indbundet i 62 foliobind. De findes alle på Det Kongelige Bibliotek i København under signaturerne Gammel Kongelig Samling (GkS) 769-777.

Den endelige redaktion er næsten trykklar, men værket blev aldrig publiceret, rimeligvis pga. omfanget: ca. 9.000 foliosider. Efter Moths død gik manuskriptet i arv til hans efterkommere, som i 1753 solgte det til kongen. Hermed begynder værkets virkningshistorie, som behandles i afsnit 3.

2. Anlæg og struktur

Selv benævnte Moth sin ordbog: "en Dansk Glosebog på Dansk og Latin og Latin og Dansk" (GkS 774, fol. 1 bl. 1r), og betegnelsen skal tages bogstaveligt: Der er tale om en bilingval og biskopal ordbog.

I sidsteredaktionen opdelte Moth værket i to parter, en verbalordbog og en art realleksikon med stednavne, fremmedord, botaniske, zoologiske, medicinske og skibstekniske betegnelser.

Makrostrukturen nødvendiggjorde en lang række henvisningsartikler, hvad der naturligvis forøger lemmatallet betragteligt. Det samme gør de lange lister af ord med samme førsteled, som Moth i sin iver efter at demonstrere det danske sprogs rigdom har opstillet under det, han anså for hovedord. Der er ca. 79.000 fuldar-tikler, defineret som artikler med i det mindste latinsk ækvivalent. Artikelstrukturen er righoldig: Efter lemma, evt. med sideformer, anføres ordklasse og bøjning, dernæst danske betydninger, frekvente ordforbindelser og endelig eksempler, hyppigt meget talesprogsnære fraser. Det er ikke mindst de sidste, der har begejstret senere tider. Konsekvent følges definitioner, ordforbindelser og eksempler af latinske modsvarigheder. Det er dog kun hovedordene, der behandles i så omfattende artikler, mens mange sammensætninger og afledninger klares med ordklasse og latinsk ækvivalent.

3. Virkningshistorien

3.1. Moths ordbog bliver offentlig ejendom

Da Moths arvinger i 1753 tilbød kongen ordbogsmanuskriptet, blev det pålagt gehejmarkivaren, Jacob Langebek (1710-75), at vurdere materialet. Langebek var historiker, men havde i en år-række arbejdet på et større ordbogsprojekt, som var iværksat af den lærde topembedsmand Frederik Rostgaard (1671-1745). Rostgaard og Langebeks plan (som er udgivet i Molbech 1827:271-76 og delvist optrykt i Lomholt 1960:219) afviger markant fra Moths. Ikke desto mindre erkendte Langebek de store kvaliteter i Moths efterladte manuskript, som nødvendigvis skulle udnyttes til at for-

bedre hans og Rostgaards forarbejder, før den planlagte ordbog kunne trykkes (Rørdam 1895:177, 218). Men Langebek var led og ked af ordbogsarbejdet, så Moths materiale henlå ubrugt i arkivet (Lomholt 1960:222-24).

3.2. Videnskabernes Selskabs Ordbog

Efter Langebeks død tog Videnskabernes Selskab ordbogssagen op. I løbet af vinteren 1776-77 udarbejdede selskabet flere planer, men uden at Moths manuskript inddrages i drøftelserne (Lomholt 1960:224-34) – man fokuserede på at få fuldført Rostgaard og Langebeks projekt. I begges arbejder er latinske ækvivalenter gennemført, men da planen blev forelagt kongen, Christian VII, befalede “Allerhøistsamme” (og det vil reelt sige kronprinsen, jf. Molbech 1843:98f.), at “Ordbogen maatte være i Landets eget Sprog, og blive lige kier for lærde og ulærde” (Lomholt 1960:230). Derfor valgte Selskabet i den endelige plan “i sær ... den methode, som af Johnson er brugt i hans Dictionary of the English Language” (fra 1755): Definitionerne skal være på dansk, ordbogen skal dække det almene ordforråd bredt, ordnes alfabetisk, følge Langebeks ortografi, medtage etymologier – og kun give få eksempler (Lomholt 1960:232f.). Der blev udpeget en redaktør, Ole Strøm (1726-82), og først i hans drøftelser med Selskabets nyligt udpegede ordbogs- eller revisionskommission blev det knæsat, at grundlaget for ordbogen skulle være Langebeks, Moths og Rostgaards manuskripter (Lomholt 1960:234f.).

Arbejdsgangen blev kompliceret: Redaktøren udarbejdede en kladde på grundlag af Moths, Rostgaards og Langebeks manuskripter, den blev dernæst suppleret med excerpter fra et antal kilder, hvorefter denne første redaktion blev indsendt til kommissionen, hvis medlemmer alle kommenterede udkastet. Eventuelt indhentes yderligere materiale, inden manuskriptet blev sendt tilbage til redaktøren, som så kunne udarbejde trykmanus (Lom-

holt 1960:244f.). Antallet af supplerende kilder steg jævnt gennem ordbogens levetid, men det blev fastholdt, at basis var de tre gamle manuskripter.

Første hæfte, bogstav A, forelå trykt i 1780 (Molbech 1843:108). Ligesom de følgende hæfter var det uden forord eller dedikation af nogen art. Manglen på oplysninger om arbejdet fremprovokerede en række spørgsmål om ordbogens "Grundregler", tidsplan og medarbejdere, fremsat af en anonym forfatter (F. Thaarup) i tidskriftet *Minerva* 1792 (Thaarup 1792:40, jf. Lomholt 1960:245f.). Til bygen af spørgsmål hører også formodningen "Moths, Rostgaards og Langebeks Arbeider letter vel meget Arbeidet?" (Thaarup 1792:41). Nicolai Elert, der havde afløst Strøm som redaktør, publicerede et svar i den følgende årgang af *Minerva*. Her bekendtgjorde han, at der ikke vil være nogen fortale til det samlede bind I,

thi Revisions-Commissionen i Videnskabernes Selskab har ... med megen Grund besluttet, at en fuldstændig Underretning ikke kunde gives Læserne, førend ved Slutningen af det heele Verk, saa som der stedse under Arbeidets Fortgang vilde forekomme Ting, hvorfor der behøvedes at gjøre Rede (Elert 1793:174).

Han opridser dernæst principperne for ordbogen, som de var fastlagt i Selskabets ovenfor omtalte plan, suppleret med nogle detaljer, bl.a. at der ikke anføres kilde til eksemplerne, undtagen i de tilfælde, hvor et lemma kun er belagt hos Moth, Rostgaard eller Langebek, som da nævnes (176).

I slutningen af 1793 udkom da første bind, A-E, af VSO, som bebudet uden nogen form for indledende oplysninger. Heller ikke bind II, F-H fra 1802, har nogen fortale el.lign. Anmelderne af de to bind efterlyste gang på gang en fortale (Lomholt 1960:245f., 250f.), og et første tilløb til oplysning af den undrende almenhed blev trykt på omslaget af første hæfte til bind III fra 1808 (aftrykt

i Müller 1810:85f.). Oplysningerne er hovedsagelig forklaringer på, at ordbogsarbejdet skrider så langsomt frem.

Fra 1806 fik projektet et væsentligt løft med ansættelsen af Christian Molbech (1783-1857), først som "Extrahent" altså excerpist (Lomholt 1960:264f.), siden som redaktør, medlem af kommissionen og til sidst dens formand til hans død. Hermed var en kyndig, energisk og interesseret filolog for første gang nogensinde blevet knyttet til projektet.

I VSO bind III fra 1820 kom omsider den længe savnede fortale, som gør rede for værkets historie, grundsætninger og hjælpemidler. Her er Moths ordbog anført som første nummer under VSO's "Hielpemidler" (XIII f.), og efter en præcis karakteristik, som dog kun gælder verbalordbogen i sidste redaktion, hedder det:

Værket ... er i denne Ordbog omhyggeligen bleven benyttet, især med Hensyn til de gamle Ord, Talemaader og Ord-sprog, hvorimod **Moths** Beskrivelse oftest maatte forandres. Saa ofte den var rigtig og ei altfor vidtløftig, beholdtes den dog formedelst det Naive og Charakteristiske i den høit fortiente Mands Udtryk (VSO bind III 1820:XIV).

Hermed var Moths betydning for VSO omsider lagt frem for offentligheden. Mere præcist og kritisk fremlægges denne afhængighed i fortalen til Molbechs egen indflydelsesrige ordbog MO (1833): Moths ordbog er "den oprindelige Grundvold" for VSO (MO 1833:IV), og der skelnes ikke mellem Moths og de senere redaktørers arbejde. Et af Moths væsentligste misgreb er, at han sammenblander ægte og uægte sammensatte verber, og den fejl er "paa en næsten ubegribelig Maade" forplantet til VSO. Desuden anholder Molbech, at Moth selv har dannet en del ord, at han ikke angiver kilder til sine eksempler, at hans materiale stammer fra forskellige sprogsfærer og -områder: "Hverdagstalen", norsk, islandsk og danske dialekter. Alt dette er der slet ikke taget højde for

i VSO, som generelt har alt for få brugsoplysninger (MO 1833: Vf.).

Molbechs arbejde på VSO var naturligvis bundet af de principper, der var fastlagt før hans fødsel, og kritikken af VSO's brug af Moth kommer ikke til orde i den fortale til VSO bind VI fra 1848, som han selv forfattede (jf. Lomholt 1960:296-99). Her hedder det tværtimod, at VSO trods sine mange mangler altid vil være brugbar som "historisk Sprog-Magazin", ikke mindst fordi den "i Følge Planen, har optaget den allerstørste Deel af den Sprogskat for Modersmaalets Lexikographie og Sproghistorie, som glemmes i *Moths* danske Ordbogs-Arbeide". Det er i høj grad optagelsen af "en stor Mængde sieldne og forældede danske Ord, Udtryk og Talemaader, som ikkun ere bevarede ved Moth", der gør VSO "umistelig i den danske Sproglitteratur" (VSO bind VI 1848:VI). Med næsten samme ord afrundede ordbogskommissionen ved Vilhelm Thomsen og Ludvig Wimmer det hidtil længst varende ordbogsprojekt i Danmarks historie (fortalen til bind VIII, VSO 1905:Vif.).

Moths status som en af hovedkilderne til VSO er altså indis-kutabel, men det er vanskeligt at danne sig et konkret indtryk af udnyttelsesgraden. Derfor har jeg udvalgt et undersøgelsesområde, bestående af fem afsnit a 100 artikler fra bind I, II, IV, VI og VIII. Da Moth kun i yderst begrænset omfang nævnes som kilde, har jeg sammenlignet lemmer og artikelindhold med sidstereaktionen af Moth's verbalordbog. I optællingerne nedenfor er lemmer i de to ordbøger regnet som identiske, hvis de udelukkende afviger mht. ortografi og for verbernes vedkommende bøjningsform; Moth har infinitiv som indgang, VSO præsens. Til lemmer regnes også sublemmer. VSO behandler forbindelser af verbum + partikel som sublemmer, og den praksis er formentlig inspireret af Moth. I opgørelserne er rene henvisningsartikler ikke medtalt. Først citeres VSO, dernæst i parentes Moth.

I undersøgelsesområdet i bind I, **Ag-Agtersting**, er 62 af de 100 lemmer i VSO identiske med Moth's, bl.a. **Agter efter** (**Agte**

efter), **Agestol** (**Age-stól**) og **Agn** (**Agn**). I 21 af de 62 artikler er der yderligere overensstemmelse. Oftest er det definitionerne, der er helt eller næsten ordret identiske, fx "**Agtendes** ... Som er i **Agt** og **Ære**" (Moth: "**Agtende** ... Er den som er i **agt** og **ære**"). Også en del kollokationer går igen, bl.a. "**Agter** ... Han **agter** hverken Gud eller Mennesker" (Moth: "**Agte** ... Hand **agter** hverken gud eller mennesker"). Det er bemærkelsesværdigt, at kun to artikler eksplicit nævner Moth som kilde, **Agerdøvel** og **Agersiig**. Selv om en del af de sammenfaldende lemmaer er frekvente ord, må overensstemmelserne karakteriseres som meget markante. Det betyder dog ikke, at VSO kritikløst har overtaget alle Moths lemmaer. I det pågældende afsnit af alfabetet har Moth 58 ord, som ikke findes i VSO. Mere end halvdelen er sammensætninger som **Agedyne**, **Ager-skreppe** og **Agter-kastel**, 15 er forbindelser af verbet **Age** + partikel, fx **forbi**, **hen** og **sammen**. Dertil kommer en håndfuld personnavne, **Agnete** m.fl., og afledninger som **Agtelig**. Forholdet demonstrerer tydeligt, at VSO's principper for lemmasektion afviger fra Moths, især mht. sammensætninger, forbindelser af verbum + partikel og proprier.

I bind II udgør undersøgelsesområdets 100 artikler **Hed-Heller**, vb. Der tegner sig ganske samme billede som i bind I, kun skal det nævnes, at kildegrundlaget kommenteres i en enkelt artikel, verbet **Heller**: "formodentligen et Provincial-Ord af uvis Oprindelse, som findes hos Moth Rostgaard og Langebek".

I bind IV, hvor **Møddingbrønd-Møllepæl** er undersøgt, er udnyttelsesgraden fortsat betydelig. Her er 63 lemmaer sammenfaldende med Moth, og 32 af dem har yderligere overensstemmelser. Den mest markante nyhed er, at Moth i dette afsnit er nævnt som kilde i 13 artikler, bl.a. "**Møddingskrab** .. Det sidste, som skræbes sammen i en Møddingpyt, og som er det fedeste af Møget. (Moth.)" (Moth: "**Møg-dynge-skrabe** l. **Mødding-skrabe** ... kaldes det sidste i en mødding put, som er det fedeste af møgget"). Også ved underbetydninger henvises der nu til Moth, bl.a. efter betydn-

ning 4 af **Møder**: “At møde for En, hindre ham i hans Gang eller Fart. (Moth.)” (Moth: “**Møde** ... at møde for en. *Impedire cursum cujus*”).

I bind VI, hvor **Skaal-Skademund** er undersøgt, er antallet af sammenfaldende lemmaer faldet noget, til 49, og kun 15 af disse artikler har yderligere materiale, nemlig definitioner, ordforbindelser eller eksempler, der er identiske med Moths. Der er også i dette afsnit 13 henvisninger til Moth, bl.a. i artiklen **Skaanlig**, hvor begge underbetydninger afsluttes med henvisningen, selv om kun definitionen af den anden er delvist sammenfaldende med Moths **Skånlig**. Fra VSO's sidste bind er afsnittet **Vo-Vognrummel** undersøgt. Heraf svarer 37 lemmaer til Moths, og af dem har 11 yderligere sammenfald. Der er også her 13 henvisninger til Moth.

Selv om det undersøgte materiale er stærkt begrænset, i alt 500 artikler i VSO, kan det konkluderes, at sammenfaldene med Moth hele vejen igennem er betydelige. Når andelen falder i de senere bind, kan det hænge sammen med, at redaktionen efterhånden inddrog flere kilder. Det stigende antal henvisninger er nok udtryk for, at Moths sprog var kommet på lang afstand af redaktørernes og kommissionens samtid. Eksempelvis afspejler en betragtelig del af sammensætningerne med **Vogn** en anden tid end gehejmerådens; som illustration skal kun en enkelt anføres: “**Vogn**klasse ... Ved Iernbanen en af de paa forskiellig Maade udstyrede Arter af Personvogne”.

3.3. Molbechs ordbog

Sideløbende med sit engagement i VSO fik Christian Molbech selv udarbejdet adskillige ordbøger, hvoraf MO (1833) er mest interessant i forbindelse med Moths virkningshistorie. Den kom i en forøget udgave 1854-59, som ikke bliver inddraget her. I fortællingen meddeler Molbech, at han systematisk har gennemgået Moths manuskript (MO 1833:III). Dernæst følger hans kritik af VSO's

brug af Moth (se ovenfor 3.2). Selv har han undladt “at optage noget Ord eller nogen Bemærkelse, der ikke fandt Hiemmel i den nærværende almindelige Sprogbrug, uden at udmærke det som *forældet, sieldent* eller *usædvanligt*”. Moth anføres ikke konsekvent som kilde, kun “der, hvor jeg har fundet Anledning til af samme at optage sieldnere, og mindre bekiendte danske Ord og Udtryk, eller hans egne uforandrede Ordforklaringer” (VI f.).

Molbechs artikler er generelt kortere end VSO's, men det er slående, at brugsoplysningerne er langt hyppigere – hvad man også måtte forvente ud fra fortalens fremlæggelse af ordbogens principper og dens kritik af VSO. Kildeangivelser findes i noget videre omfang end i VSO, men langt fra konsekvent. Som i VSO er Moth en meget central kilde, men dog systematisk suppleret med især yngre kilder i et helt andet omfang end VSO.

For at få et håndterligt mål for Molbechs udnyttelse af Moth har jeg foretaget en undersøgelse efter nogenlunde samme retningslinjer som for VSO; da kilde- og brugsangivelserne er langt hyppigere i MO, har jeg ikke fundet grund til at notere overensstemmelser i artikelindhold, medmindre Moth udtrykkeligt er nævnt som kilde. De fem undersøgelsesområder er udvalgt ved at begynde samme sted i alfabetet som i VSO og tælle 100 artikler frem; konkret drejer det sig om **Agat-Alfarvei**, **Hed-Helleflynder**, **Møddingbrønd-Mølletømmer**, **Skaal-Skalk** og **Vocal-Voldsmand**. Her er der ingen grund til særskilte opgørelser for hvert afsnit, fordi de ikke afviger markant fra hinanden, en følge af, at hele ordbogen (i to dele) udkom samme år og er udarbejdet af én redaktør efter samme principper.

Af de 500 lemmaer falder 311 sammen med lemmaer hos Moth, men kun i 46 artikler anføres Moth som kilde; i to af disse artikler optræder han to gange. Det er slående, at sammenfaldet af lemmaer er meget højt, på niveau med de tre første, mest Moth-prægede undersøgelsesområder i VSO. Forklaringen er, at MO (1833) ligesom VSO bygger på Moth, og at det var Molbechs

ambition at dække ikke bare det samtidige, men også i et vist omfang det ældre ordforråd. Endelig skal det ikke glemmes, at ordforrådet fra 1700 til 1833 nok havde undergået en udvikling, men dog også rummede en stor, stabil kerne, som ville være kommet med i ordbogen, selv om Moth ikke var blevet udnyttet.

Molbech oplyser i sin fortale, at han er gået bag om VSO direkte til Moth, og det viser sig da også, at han henviser til Moth, hvor VSO ikke gør det: i artiklen **Helde**, sb., henviser Molbech to gange til Moth, første gang under betydning 1: “et brat, ell. steilt Sted, hvor Veien gaaer nedad”, jf. Moths definition af **Helde**: “Er et steilt sted nêd ad”. VSO har under **Helde** bl.a. “Et Sted i en Vei, hvor Vognen holder til een af Siderne”, men uden henvisning til Moth. Normalt anfører Molbech ikke, at et ord, en betydning eller en ordforbindelse er forældet, når Moth er kilden, men det sker undertiden, bl.a. ved betydning 2 af **Held**: “Ynde, Behagelighed; Yndest, Gunst”. En enkelt gang tillader han sig en kritisk kommentar: Under **Skaldet** anfører han afledningen **Skaldethed** med kilde og noterer derefter: “Moth har det selvgiorte: **Skaldenhed**”.

Kun for afsnittet i bogstav A har jeg optalt, hvor mange af Moths artikler Molbech ikke har overtaget. Det drejer sig om i alt 123. Hovedparten, nemlig 55, er sammensætninger, 22 er verbum + partikel, 14 er afledninger, og hertil kommer bøjningsformer og proprier foruden en broget restgruppe (bl.a. fisken **Agnete**, træet **Ahorn** og **Ail** ‘ajle’). Det afspejler generelle forskelle i lemmatiseringspraksis (bøjningsformer og verbum + partikel) og i lem-maselektion (sammensætninger, afledninger, proprier og ord, der ikke var almindeligt udbredte omkring 1830).

3.4. Otto Kalkars Ordbog til det ældre danske sprog

Otto Kalkar (1837-1926) arbejdede størstedelen af sit liv på den ordbog til det ældre danske sprog, som i daglig tale blot kaldes Kalkars Ordbog, KalkO (1881-1918). Han fortæller selv i forordet til

første bind fra 1885 (optrykt i KalkO VI 1976), at han påbegyndte indsamlingen af ord i 1868. Fire år senere forelagde han projektet for Videnskabernes Selskab, der nedsatte en bedømmelseskomité. I deres betænkning nævnes Moths ordbog for første gang i forbindelse med Kalkars plan; man anser det “i det mindste som saare ønskeligt, at ogsaa *Mathias Moths* haandskrevne Ordbog ... helt igjennem blev tagen paa Raad med” (KalkO VI 1976:29). Kalkar imødekom dette ønske. I fortalen til bind I nævnes Moth i en lidt uklar note: I de første ark har Kalkar jævnlig anført Moths forklaringer “ordlydende og med uforandret skrivemåde, hvad jeg senere af forskellige grunde har opgivet” (KalkO VI 1976:138 note). Meningen er vistnok, at Moth i de første ark af bind I er citeret uden kildeangivelse; det er i hvert fald et ankepunkt mod Kalkar i en anmeldelse af værkets fire første bind (KalkO VI 1976:92).

Moth blev en af de vigtigste håndskrevne kilder til KalkO (Bjerrum 1976:95). Undervejs i arbejdet blev han flere gange bebrejdet, at han stolede for meget på Moth. Kritikken gjaldt oftest konkrete enkeltord og blev fremsat både i anmeldelser og i private breve (fx Bjerrum 1976:48, 60, 69f.). Kalkar tog øjensynlig ikke hensyn til kritikken, men i ansøgningen om støtte til supplementsbindet fremhæver han dog, at han “nu jævnlig har prøvesteder fra den egenlige litteratur for, hvad der hidtil er anført alene efter Moth” (Bjerrum 1967:90).

I kildelisten (KalkO VI 1976) oplyses det, at “Moth” – eller “M”, som forkortelsen blev, da økonomien krævede pladsbesparelser – er en “Håndskreven ordbog på Det kgl. Bibliotek i København (Gl. kgl. Saml. nr. 769-77)” (211 og 214). Det er altså hele komplekset af Moths ordbogsmanuskripter, der henvises til; det skal næppe tages bogstaveligt, da der indgår et meget broget materiale i flere af katalognumrene, men det må tages som udtryk for, at begge parter af slutredaktionen og i hvert fald dele af de tidligere redaktioner er blevet excerperet.

For at få et mere præcist begreb om Kalkars udnyttelse af Moth

har jeg foretaget en undersøgelse efter delvist samme retningslinjer som ved VSO og MO. KalkO har principielt altid kildeangivelse (med forordets lille forbehold for netop Moth i de første ark), og derfor har jeg kun noteret artikler, hvor Moth er angivet som kilde. Kritikernes fremhævelse af Kalkars overdrevne tillid til Moth har motiveret en opgørelse af, hvor mange artikler der udelukkende bygger på Moth, jf. også hans egen bemærkning i ansøgningen om supplementet. Desuden har jeg noteret, om Moth anføres som kilde til betydninger, ordforbindelser eller eksempler. Pga. værkets langstrakte og komplicerede historie anføres de enkelte undersøgelsesområder separat.

I bind I er afsnittet **Age-Alderd** undersøgt. Det indeholder 91 artikler i det oprindelige bind og yderligere 9 i tillægget. Af dem angives Moth som kilde i 44 artikler, heraf to gange i én artikel (**Agerfred**). Af de 44 er Moth eneste kilde til 29 artikler, bl.a. **Agegab**, **Aglevorn** og **Agte**. Det er stort set kun Moths definitioner, ikke hans fraser, der er blevet udnyttet (i 42 artikler), bl.a. "**Agle**, go. *at rokke*" (Moth: "**Agle** ... Er at rokke"). De latinske ækvivalenter er bevaret i en del artikler, fx "**Agerhumle**, no. *centauria*" (Moth: "**Ager-Humle** /en/ Er en urt. *Centaurium. Plin.*"). I andre oversættes de: "**Agerskræppe**, no. *det samme som agerburre*" (Moth: "**Ager-skreppe** /en/ Er en urt. *Lappa major*"). Kun undtagelsesvist kommenterer Kalkar Moth, fx i "**Ankertrold**, no. *en stor fisk, som bag til skal ligne et anker. Moth; vistnok = Aasen: anker-troll, d.v.s. blæksprutte*" (Moth: "**Anker-trold** /en/ skal være en stôr fisk, som båg til skal ligne et anker").

Undersøgelsesområdet i bind II, **Hed-He(de)n**lægge har noget færre henvisninger til Moth, nemlig i 18 artikler, hvoraf de 12 udelukkende har denne ene kilde, bl.a. **Hedblæsende** og **Hed(e)knok**. Udnyttelsen af Moth er i øvrigt som i første bind.

Fra bind III er **Møddingslag-Møsse** blevet undersøgt (da **Mødding(-)** er behandlet under **Mug** i KalkO, har jeg inddraget disse ord og er fortsat med **Mød-Møsse**). Her er Moth kilde til 41 ar-

tikler, heraf eneste belæg i 22. Hvor Moth nævnes jævnsides med andre kilder, sker det gerne for at dokumentere den kronologiske bredde i ordets forekomst. Som noget nyt i de undersøgte afsnit citeres Moth her 6 gange som belæg på en bestemt form, fx anføres under **Mødelig** “Hos *Moth*: *møjelig*”, og i ét tilfælde, **Mølledam**, dokumenteres genus (fælleskøn) bl.a. med Moth. Samme mønster genfindes i alt væsentligt i bindets andet undersøgelsesområde, **Skål(e)-Skåte** og **Skade-Skaldenhed** (Moth optræder i 37 artikler, heraf i 19 som eneste kilde, og udnyttes til samme oplysningstyper som i M).

I det sidste af de oprindelige bind, bind IV, er **Vo-Voldsdele** udtaget, og her optræder kilden “M.” i 50 artikler, heraf i 24 som eneste belæg. I betragtning af den store tiltro til Moth, som disse tal er udtryk for, er det bemærkelsesværdigt, at Kalkar to steder lægger luft til sin kilde: “**Voldkrone**, no. *krone, der gaves til den, der først kom op på en fjendes vold. M. Vel kun overs. af l corona vallaris*”. Moth antages altså at have opfundet et dansk lemma for at få en velkendt latinsk betegnelse med i sin ordbog. En lidt anden drejning har “**Voldrov**, no. *plyndring af døde på åben mark. M. Vel kun en forklaring af valrov*”. Her formodes Moth altså at have givet ordet en etymologisk bestemt form. En sådan kritik har jeg ikke noteret i de andre bind, men undersøgelsesområdet stærkt begrænsede omfang tillader ikke vidtgående konklusioner.

For fuldstændighedens skyld har jeg gennemlæst 100 artikler i supplementsbindet, **A(g)borre-Algode**, men uden at finde Moth nævnt.

Bortset fra supplementsbindet er udnyttelsen af Moths ordbog ganske massiv. Især er det påfaldende, at Kalkar øjensynlig har nærret så stor tiltro til Moth, at han ikke har taget i betænkning at oprette mange artikler alene på Moths autoritet. Til gengæld kan man ærgre sig over, at ingen af Moths kraftfulde eksempler er sluppet ind i KalkO.

Som i VSO og MO har jeg kun i bogstav A noteret de artikler

hos Moth, som Kalkar ikke medtager. Sammentællingen er ikke triviel, da både bind I, tillægget til bind I og supplementet, bind V, skal gennemgås under de lemmaer og stavemåder, hvor Kalkar kan tænkes at have placeret ordene. Derfor må tallene tages med forbehold. Jeg har noteret i alt 88 artikler hos Moth, som ikke er med i KalkO. Den største gruppe er som sædvanlig sammensætningerne, hvoraf 36 er negligeret, dernæst forbindelserne af verbum + partikel, 25. Resten fordeler sig jævnt på proprier, afledninger og fremmedord, foruden et par besynderlige forbigåelser, som kan skyldes, at jeg ikke har fundet ordene, nemlig interjektionen *Aja* og substantivet *Al*, if. Moth "en sort og hård jørd". Det skal bemærkes, at Kalkar tydeligvis i første omgang har forbigået umiddelbart begribelige sammensætninger som *Ager-jørd*, der siden er blevet optaget i supplementet. Det afspejler den ændrede plan for hans ordbog: Oprindeligt ville han kun medtage forældede ord og betydninger, men siden ønskede han at dække hele det hjemlige ordforråd i ordbogens periode. Tillægget til bind I retter i nogen grad op på den oprindelige skævhed, men først med supplementsbindet er hans mål nogenlunde nået (jf. Bjerrum 1976:134f. og 139).

3.5. Ordbog over det danske Sprog

Allerede inden afslutningen af VSO var de første forarbejder gået i gang til det, der skulle blive den største danske ordbog, ODS. Første bind kom i 1918, og i den stofrige indledning gives en oversigt over ODS's forgængere, hvoriblandt Moth naturligvis omtales med den største veneration (ODS I 1918:XV). Forbeholdene over for værket kommer siden i en fodnote som en begrundelse for ODS's ret restriktive retningslinjer for udnyttelsen af Moth:

Da Moth har medtaget mange dialektord, norske ord og adskillige af ham selv dannede sammensætninger, og da vi

ikke har noget sikkert middel til at udskille disse fra de ord, der tilhørte hin tids danske rigssprog, idet brugsbetegnelser er yderst sjældne i Moths ordbog, medtages fra denne kun de ord, der i *forvejen* findes i vort materiale, og som vi saaledes andetstedsfra har vished for tilhører (ell. har tilhørt) gængs dansk sprogbrug. Værdien af Moths værk for vor ordbog ligger altsaa ikke i, at det tilfører os nyt ordstof, men deri at vi ved hjælp af Moth meget ofte kan datere et ord tidligere, end det var muligt paa grundlag af vort øvrige materiale (ODS I 1918:XLV note 1).

Med denne principerklæring ville man forvente en ret beskeden indflydelse fra Moth, og det gælder da ogsaa på strukturelt niveau. Men henvisningerne til Moth er uhyre frekvente. En optælling i den elektroniske version viser, at der er ca. 38.000 referencer til Moth – i gennemsnit én pr. spalte. Både sidstereaktionens to parter (hhv. "*Moth.*" og "*Moth.Conv.*") og den næstsidste redaktion ("*Moth.1.*") er benyttet, i kildelisten specificeret på håndskriftnumre (ODS XXVIII 1956:59).

For at karakterisere udnyttelsen nærmere har jeg foretaget samme type undersøgelse af ODS som af de ældre her omtalte ordbøger. **Agat-Agraman** er undersøgelsesområdet i bind I. Her er Moth nævnt i 25 artikler, i alt 26 gange. Oftest tjener Moth som ældste belæg på et ord, ganske som man måtte forvente det efter indledningen. I de helt korte artikler uden definition bringes henvisningen umiddelbart efter ordklassen: "**Ager-jord**, *en. Moth. A67. MøllH.I.34.*" I de længere er henvisningen placeret mellem betydning og citater: "**Ager-ryg**, *en ... agerens ved pløjning højnedes midtlinie. Moth. A68. Sneen var forsvunden . . . fra Bakker, Digekamme og Agerrygge. Blich.II.492. LandbO.I.54.*" Når Moth udnyttes på denne måde, bringes kun tomme henvisninger. I en håndfuld artikler er det en underbetydning, der dokumenteres med en ligeledes tom henvisning til Moth, fx "**Agerhøne** ...

2) (*dial.*) *mariehøne*. *Moth.A66*”. Kun én gang i undersøgelsesområdet bringes et citat fra *Moth*, i artiklen *age* (betydning 1): “den mand som en gød kone fåer, got leß hand ager i sin gård. *Moth.A62*.”

Forholdene er noget anderledes i de senere bind, hvor redaktionen i højere grad havde fundet sin form. Her er afsnittene **III. hed-Hedningeland** (bind VII 1925), **Møddingbrønd-Møgso** (bind XIV 1933), **Skaal-Skabelsesberetning** (bind XIX 1940) blevet undersøgt. Kun sidste bind skiller sig så meget ud, at det bedst omtales for sig. I de 300 artikler har jeg noteret 123 henvisninger til *Moth*, fordelt på 64 artikler. Som i bind I anføres værket hyppigt som ældste belæg til et ord (fx **I. Hedme**) eller til en underbetydning (bl.a. **møget** bet. 2: “*snavset; skiden. Moth.¹M199*”). Nyt er det, at *Moth* er udnyttet som kilde til ordformer (9 gange, bl.a. “† **Hedenblomster**, en (*Moth.Conv.H48*)” under **Hedblomst**). Også bøjningsformer citeres han for (6 steder, fx “*flt. (i rigsspr. nu ikke i bet. (2.4 og) 4) -e (tidligere ogs. i bet. 4: Moth.S230 ...)*” under **Skaal**). Lidt sjældnere bruges han som belæg på genus (5 gange, fx “*et ell. (i bet. 1; nu l. br.) en (3Mos.13.7(Chr.VI). Moth.S212.*” under **I. Skab**). Og hans ordforbindelser er brugt markant hyppigere end i bind I, nemlig 42 gange, gerne med citat og undertiden kommenteret, fx “det mødte (*nu: gik*) ham ligesom hand tænkte. *Moth.M209.*”, **V. møde** bet. 2.4.

Det er velkendt, at redaktionen af ODS var underlagt skrappe krav til omfang, da værket nærmede sig sin slutning (se fx ODS XXVIII 1956:107). Det sidste undersøgelsesområde, **Vod-Vognlygte**, afslører tydeligt, at restriktionerne også ramte udnyttelsen af *Moth*. Der henvises til værket 26 gange i 14 artikler ved samme oplysningstyper som tidligere. Kun to steder bringes citater, således **Vogn** bet. 2.5 “Jeg er ikke mere tabt båg af en vogn end hand (*dvs.: jeg kan hamle op med ham*). *Moth.T3*”.

En optælling af de *Moth*-artikler, ODS har fravalgt i det første undersøgelsesområde, **Agat-Agraman**, giver 65, heraf 33 sammen-

sætninger, 22 forbindelser af verbet **Age** + partikel, 5 personnavne og et par afledninger og bøjningsformer. Langt de fleste udeladte sammensætninger er umiddelbart gennemskuelige som **Ager-skel** og **Age-vei**, og det er i overensstemmelse med ordbogens principper, at sådanne ord ikke nødvendigvis medtages. Kun en enkelt forbigåelse forekommer diskutabel: **Ager-fred** kendes fra ældre lovsprog (KalkO citerer således Osterssøn Weyles glossar), men er muligvis udeladt, fordi det er skønnet forældet allerede omkring 1700.

Hele vejen igennem ODS's mange bind er Moth en prominent kilde, og han udnyttes i videre omfang, end indledningen til bind I lægger op til. Undersøgelsen af de udeladte artikler tyder på, at man har medtaget det, der ikke decideret stred imod ordbogens principper. På den anden side kan man ikke beskyldte ODS for at nære overdreven tillid til Moth, for materialet er omhyggeligt vurderet og i det fornødne omfang kommenteret.

Trods den ret intense udnyttelse af Moth fandt der en yderligere inddragelse sted i supplementet til ordbogen, ODS-S: "Matthias Moths ordbogsarbejder ... er der ofte grund til at citere" (ODS-S 1 1992:XXX). I praksis gælder det vistnok hovedsagelig de første bogstaver, som generelt trængte mest til en udvidelse af kildegrundlaget. I hvert fald viser en gennemgang af de tidligere nævnte undersøgelsesområder, at kun A-afsnittet præges af Moth-referencer, nemlig 11 i lige så mange artikler. De er alle nyoptagne, flest sammensætninger, bl.a. de netop omtalte **Agerfred**, **Agerskel** og **Agevej**. De resterende nyheder er to fremmedord, **aggravere** og **Aggressor**, som begge henviser til *Moth.Conv.*, foruden **agering**. Der bringes i alt fem Moth-citater. De øvrige fire undersøgte afsnit har tilsammen kun én reference, i artiklen **II.Hede**, hvor der bringes et herligt citat (fra Moths artikel **Søm**): "slåe to søm af ên hede .. føde to børn på ên gang".

4. Moths ordbog i dansk leksikografi

Den langtrækkende virkning af Matthias Moths ordbogsarbejde er for længst erkendt (se fx Lindegård Hjorth 1983:24-26). De undersøgelser, der er fremlagt i denne artikel, kan tjene til en præcisering af Moths impact, som det ville hedde i dagens jargon. Og den er helt overvældende, ikke kun i VSO og MO, der proklamerer, at Moth indgår i selve deres grundlag, men også i KalkO og ODS. I yderligere nogle danske ordbogsværker er Moth blevet udnyttet, men slet ikke i en grad, der tåler sammenligning med de fire her undersøgte ordbøger.

Det er håbet, at onlineudgaven af Moths Ordbog vil betyde, at værkets virkningshistorie fortsætter, at det med de kongelige ord må "blive lige kier for lærde og ulærde".

Litteratur

Ordbøger

KalkO (1881-1918) = Otto Kalkar: *Ordbog til det ældre danske sprog (1300-1700)*. København: Universitets-Jubilæets danske Samfund I-V; fotografisk optryk København: Universitets-Jubilæets danske Samfund/Akademisk Forlag 1976.

KalkO VI (1976) = Otto Kalkar: *Ordbog til det ældre danske sprog (1300-1700)* VI. Marie Bjerrum: "Otto Kalkar og hans ordbog." Kilder og hjælpemidler. Efterskrift. København: Universitets-Jubilæets danske Samfund · Akademisk Forlag.

MO (1833) = Christian Molbech: *Dansk Ordbog indeholdende det danske Sprogs Stammeord, tilligemed afledede og sammensatte Ord, efter den nuværende Sprogbrug forklarede i deres forskellige Betydninger, og ved Talemaader og Exempler oplyste*. København: Gyldendal.

- Moths Ordbog 2013 = <mothsordbog.dk>, tilgængelig også som app.
- ODS (1918-56) = *Ordbog over det danske Sprog* I-XXVIII. København: Det danske Sprog- og Litteraturselskab · Gyldendal.
- ODS-S (1992-2005) = *Ordbog over det danske Sprog. Supplement* I-V. København: Det danske Sprog- og Litteraturselskab · Gyldendal.
- VSO (1793-1905) = *Dansk Ordbog udgiven under Videnskabernes Selskabs Bestyrelse* I-VIII. København: Videnskabernes Selskab.

Anden litteratur

- Bjerrum, Marie (1976): *Otto Kalkar og hans ordbog. I: KalkO VI*, 5-167.
- Eegholm-Pedersen, Svend (1997): *Fra Moths værksted. Lidt om den første store danske ordbog. I: Flemming Lundgreen-Nielsen, John Kousgård Sørensen & Marita Akhøj Nielsen (red.): Ord, Sprog oc artige Dict. Et overblik og 28 indblik 1500-1700. Festskrift til Poul Lindegård Hjorth*. København: UJDS · C.A. Reitzel, 484-505.
- Eegholm-Pedersen, Svend (2013): *Ministerens ordbog. Matthias Moths store danske ordbog 1700-1717. I: Magasin fra Det Kongelige Bibliotek* 4:23-37.
- Elert, Nicolai (1793): *Brev til Forfatteren af Minerva, angaaende den Danske Ordbog. I: Minerva et Maanedsskrift*. Januar, Februar og Marts, 174-189.
- Lindegård Hjorth, Poul (1983): *Danske ordbøgers historie. DSL's præsentationshæfte nr. 12*. København: DSL · C.A. Reitzel.
- Lomholt, Asger (1960): *Det Kongelige Danske Videnskabernes Selskab 1742-1942. Samlinger til Selskabets Historie* III. København: Det Kongelige Danske Videnskabernes Selskab.

- Molbech, Christian (1827): Historisk Udsigt over de danske Ordbogsarbejder i det 17. og 18. Aarhundrede. I: *Nye Danske Magazin* 2. Rk., 245-67.
- Molbech, Christian (1843): *Det Kongelige Danske Videnskabernes Selskabs Historie i dets første Aarhundrede 1742-1842*. København: Det Kongelige Danske Videnskabernes Selskab.
- Müller, P.E. (1810): Dansk Ordbog, udgiven under Videnskabernes Selskabs Bestyrelse. Anden Tome (Andet Bind), F-H Kiøbh. 1802 [anmeldelse]. I: *Kjøbenhavnske lærde Efterretninger for Aar 1810*, 81-94, 97-107, 107-11, 117-25 og 149-57.
- Rørdam, Holger Fr. (1895), udg.: *Breve fra Jacob Langebek*. København: Det Kongelige Danske Selskab for Fædrelandets Historie og Sprog · Gyldendal.
- Thaarup, F. (1792): Noget om det Kiøbenhavnske Videnskabers Selskab. I: *Minerva et Maanedsskrift*. October, November og December, 37-46.

Marita Akhøj Nielsen
ledende redaktør, dr.phil.
Det Danske Sprog- og Litteraturselskab
Christians Brygge 1
DK-1219 København K
man@dsl.dk

Norsk Ordbok i den digitale tidsalderen

Christian-Emil S. Ore & Åse Wetås

A common challenge for the editing of historical and dialect dictionaries is the heterogeneity of the source material. *Norsk Ordbok* (NO) aims at providing a scholarly and exhaustive account of the vocabulary of Norwegian dialects from 1600 to the present and of the modern written standard *Nynorsk*, thus facing the heterogeneity in both dimensions. In 2002 a new and completely digital editorial system with a meta-index as a pivot part was implemented for NO. This index enables linking of each dictionary entry to its relevant available digital source material. The online dictionary has become an academic 'enhanced publication' with reproducible results. This well-defined interlinking of the digital material opens up for user-friendly interfaces and advanced search functions, illustrations, maps and publishing of the dictionary as Linked Open Data.

1. Innleiing

Denne artikkelen vil ta for seg dokumentasjonsverket *Norsk Ordbok* (NO) og ut frå både innhald og databasearkitektur drøfta korleis verket som digital ressurs kan formidlast på nye måtar og på nye publiseringsplattformer. Me vil dessutan sjå nærare på korleis ny funksjonalitet kan leggja til rette for at verket får auka bruksverdi også for nye brukargrupper.

Etter ei kortare innleiing om NO vil me sjå nærare på dokumentasjonsordboka som database. Struktureringa av informasjonen som ligg i det digitale kjeldematerialet og i dei ferdige digitale ordartiklane har gjort det enkelt å setja opp ein funksjonell nettversjon av verket. Artikkelen vil drøfta vidare utviklingsmoglegheit og utnyttingspotensiale for desse ressursane. Siste avsnitt inneheld oppsummering og konkluderande merknader.

2. *Norsk Ordbok* – bakgrunn og arbeidsform

2.1. Verket *Norsk Ordbok*

NO er eit dokumenterende ordboksverk som gjev ei vitenskapleg framstilling av det norske folkemålet – alle dei norske dialektane i by og bygd over heile landet – og det nynorske skriftmålet. Verket representerer ei vidareføring av Ivar Aasens leksikografiske verk (Aasen 1850, 1873), slik det m.a. er framstilt i Skard (1932:17ff.). NO byggjer på eit stort empirisk grunnlagsmateriale, og dette materialet er samla inn og systematisert gjennom heile utgjevingsperioden for verket.

Frå starten av var tanken å gje ut eit verk på 2–3 ordboksband, men redaksjonen såg snart at materialet og artikkellengda gjorde eit slikt omfang heilt urealistisk (jf. Hellevik 1966). Då første band kom ut, dekte dette alfabetbolken *a–doktrinær*, og grunnlagsmaterialet var i stadig auke. På opplauspsida av arbeidet med den trykte utgåva av verket nå i 2014, er den endelege fasiten at NO blir eit dokumentasjonsverk i tolv band. Kvart av banda inneheld om lag 25 000 ordartiklar, fordelte på 800 sider. I den moderne prosjektorganisasjonen NO 2014 har kvart band ei produksjonstid på 15 månader, og det omfattar om lag 35 årsverk fagleksikografisk arbeid.

I perioden 1966–2002 kom dei fire første banda av NO ut. Denne eldste delen av verket er ikkje digitalt tilgjengeleg, og det inneber at NO som digital ressurs førebels ikkje omfattar alfabetstrekket *a–h*.

På 1990-talet blei vesentlege delar av grunnlagsmaterialet til NO digitalisert gjennom det dåverande Dokumentasjonsprosjektet (sjå Ore & Kristiansen 1998). Då det norske Kulturdepartementet og Universitetet i Oslo ved tusenårsskiftet blei samde om at NO i 12 band skulle gjerast ferdig til feiringa av 200-årsjubileet for den norske Grunnlova i 2014, gjorde digitaliseringa av kjeldene

bak verket det mogleg å gå over til fulldigital ordbokproduksjon, der mange forskarar kunne arbeida inn i eitt felles elektronisk produksjonssystem.

Det tidsavgrensa prosjektet NO 2014 starta 1. juni 2002. I startfasen blei det lagt ned mykje arbeid i å utvikla ei elektronisk redigeringsplattform for ordboka. Arbeidet skjedde i nært samarbeid mellom Eining for Digital Dokumentasjon (EDD) og fagleksikografane i NO-redaksjonen. Resultatet er blitt ein virtuell fabrikk til bruk i systematisering, prosessering, tolking og sluttlagring av store mengder språkdata. I løpet av prosjektperioden er dette systemet blitt bygd ut med nye komponentar, m.a. med ein eigen modul for semantisk sortering (jf. Bakken & Grønvik 2008:31ff.). Verktya er òg blitt lagde til rette for lagring av data i nye format. Den nyaste versjonen av NO 2014s kjeldeapplikasjon tillèt til dømes lagring av bilete og websider. Dette mogleggjer innhausting av veldokumenterte grunnlagsdata frå kjelder som ikkje er knytte til ordningar for stabil langtidslagring (jf. drøftinga av problemstillingar som gjeld tekstinnahausting frå nettkjelder i Karlsen & Wetås 2010:523f.).

Parallelt med det digitale utviklingsarbeidet har NO 2014-organisasjonen utvikla redaksjonelle retningslinjer tilpassa den digitale produksjonen. Prosjektet har òg frå 2003 av bygd opp eit elektronisk tekstkorpus over nynorsk tekst frå 1860 til i dag. Dette korpuset er fullintegrrert i redigeringsverktya og passerte tidleg i 2014 100 millionar ord.

Oppdraget frå Kulturdepartementet var opphavleg eit trykt ordboksverk i tolv band, medan ei nettutgåve av verket ikkje var med i dei originale planane. Med tida blei det meir og meir klårt at nettet i den digitale tidsalderen er ein viktigare formidlingskanal enn dei trykte bøkene. Kulturdepartementet har difor utvida «bestillinga» til å omfatta eit «gratis digitalt oppslagsverk på Internett». Kva dette inneber, er omtalt seinare i denne artikkelen.

Produksjonen av trykksatsen i NO 2014s redigeringsssystem er

basert på XML-teknologi, og det var difor rett fram å publisera artiklane på nettet. I mars 2012 lanserte NO 2014 ei første nettutgåve av ordboksverket. I 2014 er prosjektorganisasjonen inne i sitt siste ordinære driftsår, og arbeidet med vidare tilrettelegging og raffinering av nettutgåva av ordboka og nettpublisering av det store grunnlagsmaterialet i databasane har høg prioritet. Ordboka blir publisert både som trykt tekst og som enkeltvise ordartiklar i nettutgåva. Band 12, som dekkjer alfabetstrekket *u–å*, skal etter planen ferdigstillast til desember 2014. Nettpubliseringa av artiklane i dette alfabetstrekket skjer samstundes med publiseringa av det trykte bandet.

2.2. Innhaldet i databasen og mellom permanente

I ferdig utgåve vil NO innehalda meir enn 300 000 store og små portrett av ord frå norsk folkemål og nynorsk skriftmål. I den elektroniske ordbasen av ferdige artiklar er alle ordartiklane lenkja opp til relevant grunnlagsmateriale. Dermed skal alle tolkingane redaktørane har gjort i arbeidet med ordartiklane i verket, kunna etterprøvast mot kjeldematerialet som ligg føre. Denne koplinga har vore med sidan starten av NO 2014-prosjektet i 2002. I dei seinare åra har slik utvida publisering («enhanced publication») blitt stadig meir utbreidd og metodar knytte til denne er etablerte som eit eige felt innanfor digital humaniora (sjå til dømes Ver-nooy-Gerritsen 2009).

I den leksikalske databasen er ein stor del av det norske ordtilfanget vitskapleg behandla og standardisert for første gong. Normeringa av ord frå folkemålet skjer ut frå ei grunnleggjande oppfatning om at alle orda i talt norsk kan ha eit skriftleg, nynorsk korrelat. Sjølv om delar av tilfanget er i geografisk avgrensa bruk, gjeld dokumentasjonen i NO i hovudsak det sentrale og mest brukte ordtilfanget i norsk. Intensjonen er at ordboksverket skal vera eit referanseverk for norsk språk i høve til både fonologi,

morfologi, semantikk og syntaks, slik det kjem til uttrykk gjennom både einskildord og (leksikaliserte) frasar.

3. *Norsk Ordbok* som database

3.1. Metaordboka – ein felles indeks og eit ordboks-univers

Som nemnt i punkt 2.1 blei vesentlege delar av grunnlagsmaterialet til NO digitalisert på 1990-talet gjennom det dåverande Dokumentasjonsprosjektet. I hovudsak dreidde dette seg om setelsamlinga med 3,2 millionar ordsetlar, men òg ei lang rekkje eldre og nyare ordsamlingar og ordbøker blei digitaliserte. Materialet var svært heterogent, og ikkje minst framstod det svært ueinsarta når det galdt tilhøvet til gjeldande eller tidlegare rettskrivingar. Oppslagsorda på setlane og i ordsamlingane var ofte ikkje i samsvar med noka gjeldande rettskriving. For å samordna eit så stort og mangfaldig materiale blei det laga ein felles indeks – Metaordboka (MO), der alle førekomstar av eit gjeve ord er samsorterte under grovnormerte oppslagsord (jf. Ore 2001:267ff.).

MO er meir enn ei felles søkjeliste. Artikkane her inneheld både ein oppslagsdel med informasjon om det normerte oppslagsordet (ordklasse, leddeling og rettskrivingsnorm) og lenkjer til ulike skriftrealiseringar og bruksdøme i dei enkelte deldatabasane som setelarkiv, eldre og nyare ordsamlingar og ordbøker. Ein artikkel i MO kan i prinsippet ha fleire oppslagsord som følgjer ulike rettskrivingsnormer. Dermed kan den til dømes utgjera ein felles indeks over bokmål og nynorsk. Den nåverande MO har oppslagsord på nynorsk etter 1938-rettskrivinga, men me planlegg å leggja til oppslagsord i moderne nynorsk rettskriving.¹

1 Det norske Språkrådet har i eit notat til Kulturdepartementet om norsk ordbokpolitikk (Språkrådet 2014) tilrådd at *Norsk Ordboks* nettutgåve

MO omfattar alt frå gamle ordsamlingar frå 1600-talet til moderne dialektsamlingar og ordbøker. I tillegg kan ein også lagra konkordansar frå det nynorske tekstkorpuset direkte i systemet. Metaordboka har vist seg å vera eit sær sars verdfullt verkty i redigeringa av NO, men er òg i seg sjølv ei verdfull samling (jf. her òg konklusjonane om denne ressursen i evalueringa av dei vitenskaplege språksamlingane ved Universitetet i Oslo av 21.3.14 (Kristoffersen, Ebeling, Vangsnæs & Wetås 2014)). Både dei eldre og nyare ordsamlingane er lagra i original form så langt som råd. Dei uttrykkjer språksynet til forfattaren og tida dei blei skrivne i, i tillegg til å gje lingvistisk og leksikografisk informasjon. Dei er òg historiske tidsvitne og utgjer ei viktig kjelde til kulturhistorisk dokumentasjon.

Lenkjene internt i MO og mellom Metaordboka og NO går begge vegar. Eit oppslagsord i ei dialektordbok gjev på denne måten tilgjenge til alle dei andre ressursane i systemet, inklusive den tilhøyrande artikkelen i NO. Dette gjev dialektbrukarane ei eineståande moglegheit til å sjå eigen dialekt i ein større kontekst.

3.2. *Norsk Ordbok* frå tekst til tre – analyse og resultat

I dag blir dei fleste ordbøker redigerte ved hjelp av XML-teknologi. Det vil seia at ein definerer ein temmeleg fast hierarkisk trestruktur som artiklane må følgja. Kor fast denne strukturen skal vera, varierer frå ordbok til ordbok, men XML-teknologien gjer det enkelt å produsera trykksats og publisera ordboksartiklane på nett. For eldre ordboksverk som er retrodigitaliserte, er det ikkje alltid mogleg å finna eit stramt XML-skjema. Det er nok ikkje utan grunn at tilrådingane frå Text Encoding Initiative (TEI P5) har eit eige XML-element «entryFree» for tekst-

bør ligga føre i nynorsk etter den vedtekte 2012-norma. For å leggja til rette for brei bruk av grunnlagsmaterialet, er det viktig at også MO får oppslagsord etter gjeldande norm.

koding av ordbøker, der delelementa kan komma i vilkårleg rekkjefølgje.

Dei to første banda av NO blei redigerte manuelt (skrivemaskin og tradisjonell sats). Manuskriptteksten til band 3, 4 og halve band 5 blei som ein del av redigeringsarbeidet skrivne inn i eit tekstbehandlingsprogram og sett på grunnlag av eit enkelt kodesystem. Tekstflyten i manuskripta følgjer den tradisjonelle tanken om at ordboka er ein tekst, og ikkje nokon database eller trestruktur. Å tilpassa teksten i desse banda til eit moderne redigeringsssystem krev difor same analysen som for andre eldre ordboksverk.

Då redigeringsystemet for NO 2014 skulle lagast i 2002, blei teksten i dei eksisterande, eldre banda djupanalyser. Tanken var å identifisera og vidareføra den beste praksisen i systematisk form (jf. òg Grønvik 2005:161), og prosjektet var ikkje bunde til å laga eit system som kunne handtera den eldre teksten fullt ut. Mellom anna såg ein at redaktørane mange gonger i dei manuelt utforma artiklane hadde lagt inn tilvisingar til ordsamansetningar som høyrde til lengre ute i alfabetet. I eit komplett verk og i ei elektronisk ordbok må dette gjerast gjennom elektroniske krysstilvisingar.

I analyseprosessen valda òg faste ordsamband og frasar ein god del bry. I dei fleste ordbøker blir slike handsama som bruksdøme med forklåring. Eit døme er frasen *panser og plate*. I NO 2014-løysinga har denne fått ein eigen artikkel i artikkelformatet «underartikkel», og med lenkje til hovudartikkelen som fleirordsfrasen høyrer saman med. I dette tilfellet kunne ein tenkja seg lenkje både til ordartikkelen **panser** og artikkelen **plate** (som begge er med i NO), men ut frå dei vedtekne redaksjonsreglane i NO 2014s redigeringsmanual (RH), skal underartikkelen **panser og plate** lenkjast opp mot artikkelen **panser**.² På trykk er slike underartiklar innbakte i teksten til hovudartikkelen. I den elektroniske ordboka er dei moglege å søkja opp både som ein del av hovudartikkelen

² Nedfelt i RH kap. 4.3.8.4.1–2.

og som eigne, sjølvstendige artiklar med lenkje tilbake til hovudartikkelen, sjå figur 1. Alle andre tilvisingar, inklusive fullsynonym, er òg uttrykte gjennom lenkjer mellom artiklar.

The screenshot shows the interface of the Norwegian Dictionary (Norsk Ordbok 2014). At the top, the title 'Norsk Ordbok 2014' is displayed in a stylized font, with the subtitle 'Ordbok over det norske folkemålet og det nynorske skriftmålet' below it. The search bar contains the text 'panser og plate'. To the right of the search bar is a 'Søk' button and a 'Bibliografi Avansert søk' link. Below the search bar are navigation links: 'FØRRE', 'TREFFLISTE', 'ARTIKKEL', and 'NESTE'. The search results are displayed in a list format. The first result is '1. full rustning (Litterær kjelde: Fedrah 1878)', with a sub-entry 'krigarar til hest vart kledd i panser og plate (Litterær kjelde: Stokk.N1 72)'. The second result is '2. OVERFORD TYDING', with a sub-entry '(dei har) stengt humaniteten sin inne bak panser og plater (Litterær kjelde: Borg.NKB 396)'. There are also icons for audio, image, and other functions next to the search results.

Figur 1: Underartikkelen *panser og plate* i NOs nettutgåve.

I tillegg til tilvisingane mellom artiklane er alle brukseksempel kjeldefeste og kopla til ein bibliografibase med bibliografiske opplysningar uttrykte i formatet til biblioteksstandarden MARC (Machine-Readable Cataloging). Geografisk heimfesting er kopla til ein database over den norske heradsinndelinga frå 1947. Kvar heimfesting viser anten til eit herad eller eit større område som er ei gruppe av slike. Sidan polygonane som definerer grensene til herada er georefererte (koordinatfeste), er det mogleg å visa heimfestingane på kart, som vist i figur 2.

smøye-stol substantiv, hankjønn

1. hempe el lydkje på reim til å smøya (1a), stikka reimenden (på belte) inn i; smygstol (Heimfestingar: heller vanl, Rom R, Fet, Odal, Sol, Stange, Ringsaker, Åmot, Solli, Dovre, Heidal, VGausdal, NLand, ØToten, Ri, Lier, YSandavær, Tel R, Orangedal, Gjerstad, Hjelmeland, Suldal, Kvinnherad, Tysnes, Strandebrann, Vossestrand, Lindås, Eid i N5, Haram, Roms, Ndm sumst, nFjells **Kårt**)
du må trø reima igjennom smøvestolen (Litterær kjelde: SuldalsmS.Sandv.)

2. ring til å smøya (1a), stikka eitkvart gjennom, (serf:) ring, bøyle på hestesele som taumane går gjennom; smettestol (2), taumstol (Heimfestingar: Rom R, Eidsvoll, Odal R, Heidal, Hall A2, Tel R, Kvitesæid, Ryf R4, Sunndal, Trå **Kårt**; Litterær kjelde: NFL11Ød 26); òg nytta som stadnamn (Litterær kjelde: NI II,170)

3. stol av korgfletting (Heimfesting: Lærdal **Kårt**); jfr smøya (1b)



Statens kartverk

Figur 2: Ordartikkelen **smøvestol** i NOs nettutgåve. Kartet viser utbreiing av definisjonen i tydingsnummer 1.

Redigeringsystemet til NO er bygd opp av ulike modular som er knytte saman ved hjelp av referansar. Det har fleire gonger blitt sett spørjeteikn ved kvifor relasjonsdatabasesystem er valt i staden for XML. Svaret på dette er at NO 2014s redigeringsystem nyttar begge typane. Grunnen er at relasjon databasar er best til noko av funksjonaliteten den store vitskaplege ordboka må ha, medan XML-teknologi er betre til andre delar av denne. Kryssreferansar og dataintegritet blir best teke hand om av relasjonsteknologien. Komplekse og raske søk er enklast å få til med ein relasjonsdatabase. XML-teknologien er særst god til å lagra tekstleg informasjon og manipulera han. Det er ikkje uvanleg at informasjonssystem er baserte på begge teknologiane (sjå også drøftinga av dette i Ore & Tvedt 2006:318f.).

I NO 2014-systemet er ordboksteksten under redigering lagra i ein relasjonsdatabase, medan artiklane i ulike stadium i redigeringsprosessen (frå ny til ferdig) er lagra i ein logg som XML-dokument. I prinsippet kunne me utan problem erstatta redigeringsgrensesnittet vårt med ein såkalla XML-editor. Men ein slepp uansett ikkje unna programmodular som held orden på og sjekkar konsistensen på peikarane mellom artiklar, til andre delbasar som inneheld informasjon om stad, bibliografi, personar, eller som omfattar meir administrative metadata. Då trengst noko meir, dvs. tabellar og relasjonar. Redigeringsgrensesnittet for artiklane i NO er gjennom åra finslipt slik at det skal vera lett å læra opp nye redaktørar. I tillegg skal systemet støtta opp under rask produksjon og sørgja for at sluttproduktet er konsistent og i samsvar med grunnlagsempirien.

Det viktigaste skiljet er om ein vel å skriva ordartiklane på tradisjonell måte som fritekst utan formell struktur, eller om ein ser nytte i å organisera innhaldet som ein gjennomtenkt formell struktur. Debatten om XML kontra relasjonsdatabasar er etter vårt syn lite relevant for fagleksikografer. Avgjerande for ei ordbok sin systematikk og akribi er ein godt gjennomtenkt datamodell og struktur, kombinert med solid fagleg arbeid.

4. Nettutgåva av ordboksverket

Nettutgåva av NO er fritt tilgjengeleg og omfattar dei delane av ordboksverket som er produserte på ei digital plattform. Dette gjeld alfabetstrekke frå og med *i*. Det redaksjonelle produksjonssystemet som blei utvikla frå og med 2002, blei laga for plattformuavhengig publisering. Dette medførte at arbeidet med å laga nettpublikasjonen NO var ein relativt liten og enkel jobb (sjå Ore 2012:492f., jf. òg Ore & Tvedt 2006:319f.). Den store faglege vinsten i å gje ut ordboka også som nettpublikasjon, ligg i at dette gjev oss høve til å ekspandera informasjon, slik det i NO på nett er gjort m.a. med kjelder, heimfestingar, etymologiopplysningar og ordklasseinformasjon. Det gjev vidare høve til å publisera visse kategoriar som tilvalsstoff for lesaren, slik at navigeringa i artikkelen blir enklare. I NO på nett er opplysningar om etymologi, uttale, alternative skriftformer og målførevariasjon lagde inn som informasjonskategoriar ordbokbrukaren må navigera seg inn i ved bruk av spesialutvikla ikon.

Nettutgåva har òg gjort det mogleg å leggja inn tilleggsfunksjonalitet som forklårande rettleiingstekst og aktive lenkjer på alle punkt der ordartiklar er kopla elektronisk saman på artikkel- eller tydingsnivå. I tillegg er kartvisningar frå 2014 integrerte for alle artiklar som har opplysningar om målførevariasjon eller geografisk avgrensa utbreiing for ein gjeven definisjon. Søkjefunksjonaliteten gjer det mogleg å søkja både på normalartiklar og på fleirordsfrasar i underartiklar (typen *slå an tonen, stolpar opp og ned, på naud og neppe* osv., jf. punkt 3.2).

Nettutgåva omfattar ikkje berre den teksten ein kan lesa i den trykte utgåva. Den skal vera eit digitalt oppslagsverk som omfattar opplysningane ein kan nå gjennom MO-koplingane, bibliografi- og kartdatabasane og dei store underliggjande samlingane av grunnlagsdokumentasjon. For den framtidige forvaltninga og utviklinga av NO er dette eit svært viktig poeng.

5. Utviklingsmoglegheiter og utnyttingspotensiale

5.1. Ny informasjon og nye plattformer

Den gjennomarbeidde datamodellen som ligg bak redigeringsprogrammet til Norsk Ordbok 2014 gjer det enkelt å kopla ny informasjon og nye funksjonar saman med dei eksisterande informasjonskategoriane. Det er mogleg å supplera ordartiklane med lyd, bilete og video, og det ligg eit stort potensiale både for forskning og formidling i å kopla saman den digitale ordbokressursen med andre system (til dømes Artsdatabanken og ensyklopediske ressursar som *allkunne.no* og *Store norske leksikon*). Ordboka i fullelektronisk format vil det òg vera mogleg å kopla saman med andre ordbøker i Norden. Her kan ein tenkja seg å bruka Linked Open Data-teknologi og såkalla semantiske teknologiar som RDF (Resource Description Framework), til eit slikt felles ordboksnett i Norden. Moglegheitene er legio, så lenge ein har veldefinerte modellar for ordboksdata.

Utviklinga av nye publiseringssplattformer bør også gjera at ordboksressursane blir lagde til rette for slike nye format. Særleg aktuelt vil det vera å laga ein versjon tilpassa smarttelefonar.

5.2. Språkvitskapleg utnyttingspotensiale

NOs digitale ordbase er utforma slik at han gjev eit svært godt utgangspunkt for utarbeiding av ordnett. Den definerte artikkelkategorien «underartikkel», som er til bruk på alle leksikaliserte fleirordsfrasar i NO-basen, gjer det òg mogleg å ekstrahera desse til oppbygging av eit separat konstruktikon (jf. Lyngfelt & Sköldberg 2013). I ei leksikalsk basert framstilling som ei ordbok vil dei minst generelle konstruksjonane vera dei mest tilgjengelege og dei som i størst omfang er skildra i eigne artikkelstrukturar, medan

konstruksjonsmønster med større generalitet i mindre grad er omtalte, og kanskje berre kjem til uttrykk som bruksdøme (jf. til dømes Berg-Olsen 2014).

Organiseringa av informasjonen i ordbasen, med aktive digitale koplingar mellom ordartiklar, gjer det også mogleg å ekstrahe- ra taksonomiske nettverk for ulike fagfelt. Til dømes er folkelege artsnemningar for dyr og plantar alltid kopla saman med vit- skaplege artsnemningar, og dei vitenskaplege nemningane ligg inne i basen med søkbare bruksmerkingar (bot. og zool). Dette gjer basen verdfull for felt som til dømes etnobiologisk gransking (jf. Wetås 2013). Liknande utnyttingspotensiale finst på ei rekkje andre felt, og desse vil vera tilgjengelege innanfor rammene av ordboka som digital ressurs.

6. Oppsummering og konkluderande merknader

6.1. Det beste av gammalt og nytt

Å gje ut verket NO parallelt som trykt publikasjon og som nettpublikasjon kan framstå som unødvendig dobbeltarbeid. Me har i denne artikkelen vist korleis dei redaksjonelle redigeringsverktya er utvikla slik at resultatata kan publiserast på fleire plattformer utan ekstra tilrettelegging. I praksis viser det seg dessutan at den fleirmediale publiseringa hentar ut det beste av den gamle og den nye verda. Fordi NO 2014s oppdrag er å levera eit ordboksverk i 12 band à 800 sider trykt tekst til ein gjeven dato, må organisasjonen heile tida følgja strenge og ufråvikelege krav til omfangsdimensjonering og tidsbruk. Resultatet er informasjonstette og stringente artiklar.

Publiseringa på digitale plattformer gjev høve til å laga utvi- da og meir brukarvennelege artikkelvisningar, til å tilby brukarane

aktive, klikkbare lenkjer mellom artiklar og til å kopla saman den ferdige ordboksressursen med andre digitale ressursar. Med ein ferdig ordbase som dekkjer alfabetstrekket frå *a* til *å*, og der heile tekstmassen er utforma etter dei same redaksjonelle retningslinene, vil det opna seg nye moglegheiter for å ekstrahera delmengder av artiklane til annan bruk. Denne kan vera geografisk eller tematisk avgrensa, eller òg til dømes avgrensa av bruksfrekvens, til bruk i ordboksprodukt som skal dekkja kjerneordforrådet.

Den store ordboka på digital plattform har på denne måten gjeve tre viktige resultat: For det første er store mengder språkdokumentasjon gjort digitalt tilgjengeleg både for forskarsamfunnet i inn- og utland og for ålmenta. Dette gjeld både dei ferdige leksikografiske produkta og dei store underliggjande relasjonsdatabasane med digitalt tilgjengelege grunnlagskjelder.

For det andre har EDD og NO 2014 utvikla leksikografiske forskingsverktøy som møter behovet for lett etterprøvar dokumentasjon av språk. Desse verktøya har eit stort utnyttingspotensiale både for norsk og for andre språk med ulik grad av standardisering, og dei er dessutan sette opp slik at dei kan handtera diakron og synkron heterogenitet (jf. òg Grønvik & Ore 2013:244f.).

For det tredje har overgangen til digitale publiseringplattformer for den store ordboka gjort NO til ein ressurs som stadig kan byggjast ut og som i tillegg kan koplast saman med andre digitale ressursar. Ein kan her nemna ressursane publiserte på ordnet.dk (særleg DDO og ODS) og ei framtidig digital utgåve av SAOB. Desse inneheld relevant informasjon om nærskylde ord i dei skandinaviske språka. Sist men ikkje minst har ein den framtidige vitenskaplege ordboka over norsk bokmål (jf. Vonen 2014). Det er særskild viktig at ein planlegg eit dokumenterande verk over bokmålet slik at det blir lett å kopla denne ressursen opp mot *Norsk Ordbok*. Resultatet vil bli ei best mogleg felles skildring av dei to norske skriftspråka.

6.2. Sluttord om store ordbøker som digital ressurs

Fulldigital ordboksproduksjon bidreg vesentleg til ei systematisk og stringent behandling av ordmaterialet. I tillegg gjev den digitale produksjonen store vinstar i form av høve til å behandla store mengder data. For NOs del har omlegginga til fulldigital produksjon gjort at redaksjonen kan ha eit høgt arbeidstempo. Dette skuldast ikkje minst at mange tidkrevjande rutineoppgåver nå blir løyste av produksjonssystemet. Eit enkelt døme er måten litteraturreferansar, heimfestingar, grammatiske kategoriar, bruksmerkingar osv. er gjorde tilgjengelege som val frå faste menyar. Det sparar arbeid i redigeringsfasen, det sparar arbeid i korrekturfasen og det gjer det mogleg å enkelt ekspandera referansar og kategori-markeringar i digitale visningsformat.

Store og dokumenterande ordboksverk er av natur komplekse og informasjonstette. Dette gjer fleirbrukspotensialet deira som digitale ressursar endå større.

Litteratur

Ordbøker

DDO (2003–05) = *Den Danske Ordbog*. København: Det Danske Sprog- og Litteraturselskab og Gyldendal. <ordnet.dk/ddo> (august 2014).

NO (1966–) = *Norsk Ordbok. Ordbok over det norske folkemålet og det nynorske skriftmålet*. Oslo: Det Norske Samlaget. Nettutgåve: <no2014.uio.no> (juni 2014).

ODS (1918–56) = *Ordbog over det danske Sprog*. København: Det Danske Sprog- og Litteraturselskab og Gyldendal. <ordnet.dk/ods> (august 2014).

- SAOB (1898–) = *Ordbok över svenska språket utgiven av Svenska Akademien (Svenska Akademiens Ordbok)*. Lund: Gleerups förlag.
- Aasen, Ivar (1850): *Ordbog over det norske Folkemaal*. Kristiania: Carl C. Werner & Comp.
- Aasen, Ivar (1873): *Norsk Ordbog*. Christiania: P.T. Mallings Boghandel.

Annan litteratur

- allkunne.no. Nynorsk digitalt oppslagsverk. <allkunne.no> (juni 2014).
- Artsdatabanken. <www.artsdatabanken.no> (august 2014).
- Bakken, Kristin & Oddrun Grønvik (2008): Materialsortering på digital plattform. Eit steg mot høgare dataintegritet i den vitenskaplege leksikografien. I: Asta Svavarsdóttir, Guðrún Kvaran, Gunnlaugur Ingólfsson & Jón Hilmar Jónsson (red.): *Nordiske Studier i Leksikografi 9. Rapport fra konference om leksikografi i Norden. Akureyri 22.–26. maj 2007*. Reykjavík 2008, 31–42.
- Berg-Olsen, Sturla (2014): Skjematiske fleirordskonstruksjonar i *Norsk Ordbok*. I: Ruth Vatvedt Fjeld & Marit Hovdenak (red.): *Nordiske studier i leksikografi 12. Rapport fra Konferanse om leksikografi i Norden. Oslo 13.–16. august 2013*. Oslo 2014, 80–94.
- Grønvik, Oddrun (2005): Redigeringsprogrammet for Norsk Ordbok 2014. I skjæringspunktet mellom menneskeleg skjønn og automatisering. I: Ruth Vatvedt Fjeld & Dagfinn Worren (red.): *Nordiske studiar i leksikografi 7. Rapport frå Konferanse om leksikografi i Norden. Volda 20.–24. mai 2003*. Oslo 2005, 156–165.
- Grønvik, Oddrun & Christian-Emil Smith Ore (2013): What should the electronic dictionary do for you – and how? I: Iztok Kosem et al. (eds.): *Electronic lexicography in the 21st century: thinking outside the paper. Proceedings of the eLex 2013 conference, 17–19 October 2013, Tallinn, Estonia*. Tallinn: Eesti Keele

- Instituut, 243–260. <eki.ee/elex2013/conf-proceedings/> (juni 2014).
- Hellevik, Alf (1966): *Innleiing*. I: Olav T. Beito et al. (red.): *Norsk Ordbok*. Band 1: *a–doktrinær*. Oslo: Det Norske Samlaget, I–XVI.
- Karlsen, Knut E. & Åse Wetås (2010): Internettpublikasjonar som kjelde til dokumentasjonsordbøker. Status og bruk. I: Harry Lönnroth & Kristina Nikula (red.): *Nordiska studier i lexikografi 10. Rapport frå Konferensen om lexikografi i Norden. Tammerfors 3–5 juni 2009*. Tammerfors 2010, 522–529.
- Kristoffersen, Kristian Emil, Signe Oksefjell Ebeling, Øystein Vangsnes & Åse Wetås (2014): Den forskings- og undervisningsmessige verdien av eit utval samlingsressursar ved ILN. Rapport frå ei arbeidsgruppe nedsett av instituttleiar ved ILN hausten 2013. <<http://bit.ly/1vdjAFS>> (juni 2014).
- Lyngfelt, Benjamin & Emma Sköldberg (2013): Lexikon och konstruktikon – ett konstruktionsgrammatiskt perspektiv på lexikografi. I: *LexicoNordica* 20, 75–92.
- MARC Machine-Readable Cataloging. <www.loc.gov/marc/umb/> (juni 2014).
- Ore, Christian-Emil (2001): Metaordboken – et rammeverk for Norsk Ordbok? I: Martin Gellerstam, Kristinn Jóhannesson, Bo Ralph & Lena Rogström (red.): *Nordiska studier i lexikografi 5. Rapport frå Konferens om lexikografi i Norden. Göteborg 27–29 maj 1999*. Göteborg 2001, 250–270.
- Ore, Christian-Emil (2012): Nettordbøker og Norsk Ordbok – hvordan etablere en vitenskapelig nettordbok. I: Birgit Eaker, Lennart Larsson & Anki Mattisson (red.): *Nordiska studier i lexikografi 11. Rapport frå Konferensen om lexikografi i Norden. Lund 24–27 maj 2011*. Lund 2012, 488–499.
- Ore, Christian-Emil & Nina Kristiansen (1998): *Sluttrapport frå Dokumentasjonsprosjektet*. <www.dokpro.uio.no/sluttrapp.pdf> (april 2014).

- Ore, Christian-Emil & Lars Jørgen Tvedt (2006): Digital sats eller digital satsing. I: Henrik Lorentzen & Lars Trap-Jensen (red.): *Nordiske Studier i Leksikografi 8. Rapport fra Konferanse om Leksikografi i Norden. Sønderborg 24.–28. maj 2005*. København 2006, 315–322.
- RDF = Resource Definition Framework. <www.w3.org/RDF/> (april 2014).
- RH = *Redigeringshandbok for Norsk Ordbok 2014*. <<http://bit.ly/109tCa1>> (juni 2014).
- Skard, Sigmund (1932): *Norsk Ordbok. Historie – plan – arbeidskipnad*. Oslo: Det Norske Samlaget.
- Språkrådet (2014): Norsk ordbokpolitikk. Notat fra Språkrådet til Kulturdepartementet 3. mars 2014. <<http://bit.ly/OBoRYU>> (april 2014).
- Store norske leksikon*. <snl.no> (juni 2014).
- TEI P5: *Guidelines for Electronic Text Encoding and Interchange*. TEI Consortium, Charlottesville, Virginia, USA 2014. <www.tei-c.org/release/doc/tei-p5-doc/en/Guidelines.pdf> (juni 2014).
- Vernooy-Gerritsen, Marjan (2009): *Emerging standards for enhanced publications and repository technology: Survey on technology*. Amsterdam: Amsterdam University Press.
- Vonen, Arnfinn Muruvik (2014): Behovet for store ordbøker i en samlet norsk ordbokpolitikk. I: *LexicoNordica 21* (dette nummeret).
- Wetås, Åse (2013): Plantenemningar i Norsk Ordbok. I: Lars-Erik Edlund, Inger Larsson & Staffan Nyström (red.): *Växter och växtnamn – ett möte mellan botanik och språkvetenskap*. Acta Academiae Regiae Gustavi Adolphi CXXIII. Uppsala: Kungl. Gustaf Adolfs Akademien för svensk folkkultur, 165–174.

Christian-Emil S. Ore
førsteamanuensis, leiar for Eining
for Digital Dokumentasjon,
dr.scient.
Institutt for lingvistiske og nordiske
studium
Universitetet i Oslo
Postboks 1123 Blindern
NO-0317 Oslo
c.e.s.ore@iln.uio.no

Åse Wetås
prosjektdirektør NO 2014, dr.art.
Institutt for lingvistiske og nordiske
studium
Universitetet i Oslo
Postboks 1021 Blindern
NO-0315 Oslo
ase.wetas@iln.uio.no

Stora finska ordböcker i ett historiskt perspektiv

Klaas Ruppel & Caroline Sandström

The first dictionary of the Finnish language of greater importance was Ganander's dictionary from 1787, which remained as a manuscript, and was only published facsimile between 1937 and 1940 (printed version 1997). The second great Finnish dictionary was compiled and published by Lönnrot between 1866 and 1880. The process of making and publishing dictionaries was important for the nationalistic movement in Finland in the 19th and 20th centuries. A dictionary of Finnish dialects was already started in the 19th century; a large archive of dialects was gathered by professionals, but also by what is known as *crowdsourcing*. For the contemporary Finnish language, three large descriptive and normative dictionaries can be identified. The first, *Nykysuomen sanakirja*, was published between 1951 and 1961. It was followed by *Suomen kielen perussanakirja* in the 1990s and *Kielitoimiston sanakirja* at the beginning of the 21st century; both consisting of some 100 000 entries.

1. Inledning

I vårt bidrag diskuterar vi utvecklingen av ordböcker för det finska allmänspråket och de finska dialekterna. Vi koncentrerar oss på ordböcker som antingen i fråga om omfång eller i fråga om betydelse för finsk lexikografi kan betecknas som ”stora”. Tidsmässigt tecknar vi en linje från 1600–1700-talen då de första finska ordböckerna uppstod, via tidiga tvåspråkiga ordböcker, till utgivningssituationen i dag.

2. Finskans samhälleliga utveckling under 1800–1900-talen

Inledningsvis ger vi en bakgrund till den politiska och sociala situation inom vilken de finska ordböckerna tillkom och belyser finskans väg till ett samhällsbärande språk. Finska har alltid varit modersmål för huvuddelen av befolkningen i Finland. Inom kyrkan användes finska tidigt och någon form av elementär, ofta sporadisk och temporär, undervisning på finska förekom (se t.ex. Klinge 1996:137–143; Fagerlund et al. 2001:122f.). Som ett arv från den period då Finland var en del av det svenska riket förblev svenska administrations- och bildningsspråk. Också efter 1809, då Finland tillföll Ryssland i 1808–1809-årskrig och blev ett autonomt storfurstendöme under den ryska kejsaren, förblev svenska administrationsspråk – bl.a. fortsatte 1734 års lag att gälla som grund för Finlands lagar (Klinge 1996:27, 39–40; Fagerlund et al. 2001:335).

År 1831 grundades Finska litteratursällskapet, *Suomalaisen Kirjallisuuden Seura*, som kom att spela en viktig roll för finsk lexikografi och folkligt bildningsarbete.¹ År 1856 blev finska ett läroämne i läroverk. Första finska läroverket grundades i Jyväskylä år 1858. Genom en ny skolförordning fick finskan 1865 vid sidan av svenskan ställning som skolspråk i läroverk och gymnasier. (Klinge 1996:256–269; Andersson 1999:416–419; Varpio 1999:351f.)

En stark nationell väckelse och tanken om ”ett folk, ett språk” blev ett nationellt projekt som hade som mål att utveckla finskan som kultur- och bildningsspråk. Ett finskt skriftspråk skapades och tog form under senare hälften av 1800-talet. Normeringsprocessen pågick under flera decennier. Eftersom man av historiska skäl ville

1 De dätida språkförhållandena belyses av att då Finska litteratursällskapet inledde sin verksamhet försökte man föra protokoll på finska, men avstod från det, eftersom det finska skriftspråket inte ännu fungerade för protokoll. Protokollen skrevs på svenska fram till 1860-talet (Varpio 1999:351).

frigöra sig från svensk påverkan skapades nya finska ord ofta på puristisk grund (se t.ex. Klinge 1996:40–43).

Genom en förordning som stärkte finskans ställning blev finska ett officiellt språk i Finland 1902. Att Finland har två nationalspråk, finska och svenska, slogs fast i 1919-års regeringsreform. För finskan kan detta ses som en demokratiseringsprocess som lyckades; finskan hade nått positionen som ett samhällsbärande språk.

Ordböckerna var en viktig del av det nationella finska projektet, vilket också gäller många andra nationalspråk. Anderson framhåller att verksamheten att redigera – och i synnerhet att låta trycka och sprida – ordböcker har varit ett viktigt instrument för att bygga upp och skapa en nationell samhörighetskänsla i många europeiska länder (Anderson 2003:3–7, 37–46).

3. De första ordböckerna över det finska språket

Den lexikografiska historien kan för finskans del anses ha börjat år 1637, då Ericus Schroderus finska ordlista *Lexicon Latino-Scondicum* som omfattar 2 400 ord utkom. Den följdes av en del andra liknande ordlistor och hjälpredskap som i första hand var avsedda för tolkar.

År 1745 publicerade Daniel Juslenius den första egentliga finska ordboken, *Suomalaisen Sana-Lugun Coetus/Fennici Lexici Tentamen/Finskt Orda-Boks Försök*. Men den var varken till innehåll eller omfång någon stor ordbok. (För uppgifter om de första finska ordböckerna, se Hakulinen 1967; Hovdhaugen et al. 2000:61–63.)

3.1. Gananders ordbok *Nytt Finskt Lexicon*

Den första stora finska ordboken sammanställdes 1787. Det var kaplanen Christfrid Gananders *Nytt Finskt Lexicon* (NFL) på nästan 3 000 manuskriptsidor. Lexikonet som hade över 30 000 artiklar publicerades däremot inte vid denna tidpunkt.

Gananders samtida, Henrik Gabriel Porthan, som var professor i vältalighet vid Kungliga Akademien i Åbo, framförde ändringsförslag, strykningar och kompletteringar. Han ansåg dessutom att mer material borde samlas in, vilket fördröjde publiceringen. En del forskare har antagit att Porthan genom ändringsförslagen med avsikt försökte förhindra att NFL trycktes (se Hakulinen 1967:94; NFL 1997:IX; Häkkinen 1998:635). Häkkinen (1998:635f.) påpekar dock att Porthans anmärkningar och förbättringsförslag var berättigade, eftersom det fanns svagheter i manuskriptet. Alfabetiseringen var inkonsekvent och bl.a. stod dialektala uttalsvarianter (t.ex. *ihte*, *itte* av ordet *itse* 'själv'), böjda former av ord och produktiva avledningar som egna uppslagsord.

NFL är en stor ordbok: den utgjorde den största finska ordsamlingen då den sammanställdes för cirka 230 år sedan. Senare tiders lexikografer har utnyttjat detta material när de utarbetat nya finska ordböcker. Först 150 år efter att manuskriptet blivit färdigt utkom en faksimilutgåva av lexikonet (NFL 1937–1940). Det räckte ytterligare sextio år innan NFL år 1997 utkom i tryck. Trots denna fördröjning lyfte NFL finsk lexikografi till en ny nivå. Ganander hade på ett systematiskt sätt samlat in ord och uttryck ur både litteratur och talspråk. Han gav exempel på hur orden användes i levande språkbruk och angav för det mesta källa eller insamlingsort för beläggen (Hakulinen 1967:95–98).

Efter Porthans död hamnade manuskriptet i händerna på en annan lexikograf, Kustaa Renvall. Han använde det som huvudkälla för sin egen ordbok *Suomalainen Sana-Kirja. Lexicon linguae Finnicae, cum interpretatione duplici, copiosiore Latina, brevior Germanica*. Den omfattar drygt 20 000 artiklar och kom ut under åren 1823–1826. Vid Åbo brand 1827 brann biblioteket vid Akademien ner. Gananders manuskript hade sannolikt blivit förstört ifall det inte hade varit i Renvalls förvar, men det räddades tack vare Renvalls ordboksprojekt. Efter att Renvalls ordbok hade kom-

mit ut minskade däremot behovet av att färdigställa Gananders manuskript och göra det publicerbart.

3.2. Lönnrots stora ordbok 1866–1880

Nästa milstolpe för de stora finska ordböckerna var Elias Lönnrots *Suomalais-Ruotsalainen Sanakirja/Finskt-Svenskt Lexikon* (SRSK) som utgavs häftesvis under åren 1866–1880. Ordboken innehåller cirka 240 000 artiklar och måste därför redan genom sin storlek räknas bland de stora finska ordböckerna.² Utgår man från mängden artiklar är detta verk den största finskspråkiga ordboken som hittills utgivits i sin helhet. På svenska har SRSK granskats av Romppanen (2001:163–173) och Koukkunen (2004); också Haarala (1999:150) diskuterar ordbokens betydelse för finsk lexikografi.

Lönnrot utarbetade uppslagsord och artiklar utgående från samtida tillgängliga ordlistor. Han tillfogade produktiva avledningar, av vilka en stor del aldrig använts i levande språkbruk (Rapola 1931:200; Ruoppila 1962:210; Hovdhaugen et al. 2000:252f.; Koukkunen 2004:137f.). Rapola (1932:41) förhåller sig kritiskt till det konstruerade material som ingår i ordboken, eftersom det rentav kunde vara menligt för forskningen. Det gav också intryck av att ordboken var mer omfattande än den i själva verket är.

Lönnrots metoder kan dock försvaras i ljuset av 1800-talets nationella ideologi med dess strävan att utvidga det finska ordförrådet och utveckla finskan till ett bildningsspråk. Många dåtida ordskapelser är i dag naturliga element i det finska språket; t.ex. *aate* 'idé', *henkilö* 'person' (Hakulinen 1961:301f. och 1979:459f.). Lönnrot var en aktiv ordskapare. Genom sin ordbok ville han visa på rikedomen i det finska språket. Han uppfattade bl.a. alla de avledningar som kunde skapas med hjälp av produktiva suffix som ett bevis för detta (Hakulinen 1967:103).

2 Antalet artiklar har tidigare beräknats vara 200 000. Koukkunen (2004:135f.) har dock kommit fram till en högre siffra.

3.3. Jämförelser mellan Gananders och Lönnrots ordböcker

Det finns flera faktorer som förenar SRSK och NFL. För det första överfördes en del av ordförrådet – via framför allt Renvalls ordbok – från Ganander till Lönnrot. Lönnrot verkar däremot inte ha använt Gananders manuskript som primär källa för sin ordbok (SRSK 1866:I).

Som ovan framgått bör NFL snarare betraktas som en materialsamling än som en färdig ordbok. Lönnrot tycks också själv ha uppfattat sin egen ordbok som en materialsamling, som senare borde byggas ut till en riktig ordbok. I förordet räknar han upp de ordböcker och de personer som utgör hans källor. Lönnrot påpekar att han i enlighet med Litteratursällskapets beslut ordnat ordsamlingarna alfabetiskt, ”afvikande från den etymologiska uppställningen eller stamordsordningen, som i alla hittills utkomna fullständigare finska ordböcker uteslutande begagnats” (SRSK 1866:I). I slutordet till ordboken konstaterar han fjorton år senare (SRSK 1880:1085):

Med dessa korta slutord öfverlemnar jag således detta arbete icke såsom en kritiskt utarbetad ordbok öfver finska språket, utan hellre såsom en alfabetiskt ordnad samling af alla vid dess utarbetande tillgängliga finska ord, hvilken väntar sin vetenskapliga bearbetning af någon lexikograf.

Lönnrot torde i denna bedömning vara alltför blygsam och sträng mot sig själv, men p.g.a. insamlingsmetoden har verket onekligen vissa drag som påminner om en ordsamling.

Ett gemensamt drag för NFL och SRSK är att det i bägge ingår material som ordboks författarna själva samlat in eller låtit samla in direkt ur folkligt språkbruk. Lönnrot lyckades under sina resor samla in material också från östliga finska dialekter som det dittills

hade funnits få uppgifter om. Eftersom han använde svenska som metaspråk, och gav förklaringar för folkliga ord och uttryck på svenska blev förklaringarna naturliga och levande. Renvall hade i sin ordbok huvudsakligen använt latin, som inte lämpade sig lika väl för förklaringar av finska ord (Hakulinen 1967:99).

Den viktigaste gemensamma nämnaren för NFL och SRSK är dock att de bara formellt är tvåspråkiga finsk–svenska ordböcker. Det primära syftet var att lyfta fram det finska ordförrådet. Att använda svenska som metaspråk var ett naturligt val vid denna tidpunkt, eftersom det finska språket ännu inte var moget för uppgiften. Ett exempel på detta är att det i synnerhet i SRSK ges ett brett spektrum av svenska motsvarigheter till de finska orden. De svenska orden ska alltså inte ses som ekvivalenter i en tvåspråkig ordbok, utan deras uppgift är att beskriva betydelsefälten för finska ord (Hakulinen 1967:105; Haarala 1999:150; Romppanen 2001:171; Koukkunen 2004:133, 149). Att SRSK inte är en översättningsordbok utan en ordbok som belyser det finska språket får också stöd av det faktum att ordbokens titel inte ursprungligen var *Suomalais-Ruotsalainen Sanakirja/Finskt-Svenskt Lexikon*, utan *Suomalainen ja ruotsalainen sanakirja*. Med denna titel, Finsk *och* svensk ordbok, behandlar den samtida Yrjö Koskinen (1867) det första häftet av SRSK.³ Genom namnvalet strävade Lönnrot efter att visa på skillnaderna mellan sin egen ordbok och till exempel Euréns några år tidigare utkomna *Finsk–Svensk Ordbok* (1860). (Se även Haarala 1999:150.)

Att Yrjö Koskinen redan 1867 uppfattade SRSK som en förklarande finsk ordbok framgår av det avslutande omdömet i hans recension:

När vårt språk någon gång i tiden blivit så förädlad och högt skattat, att det finska folket får en ordbok för att rätt upp-

3 De ursprungliga ordbokshäftena finns inte bevarade, och man måste därför ty sig till sekundära källor.

fatta och använda sitt språk så bör detta verk om finska naturligtvis utarbetas på finska, såsom brukligt är på andra kulturspråk.

4. 1900-talets stora finska ordböcker

Under 1800-talets allra sista år fick den finska lexikografen en ny inriktning, då språkforskaren och politikern E. N. Setälä trädde in på scenen. På ett möte på Finska litteratursällskapet 1896 introducerade han ett ordboksprogram för vetenskapliga finska ordböcker, som skulle omfatta en allmänfinsk ordbok, en ordbok över äldre finskt skriftspråk och en ordbok över folkmålen. Senare kompletterade han programmet med en etymologisk ordbok (Vilppula 1999:407f.; Onikki-Rantajääskö 2011:544–547).

Setäläs vision vann understöd och arbetet med att förverkliga ordboksprogrammet inleddes. Under åren 1951–1961 utgavs en etymologisk ordbok (SKES). Arbetet med en ordbok över det äldre finska skriftspråket (VKS 1985–) pågår fortfarande. Ordböckerna över allmänspråket och den finska dialektordboken kommer vi i det följande att beskriva mer ingående.

4.1. Den finska dialektordboken

Av de finska ordboksprojekten genom tiderna är *Suomen murteiden sanakirja* 'ordbok över finska dialekter' (SMS) den mest omfattande; ett mastodontverk och den absolut största ordboken som någonsin gjorts i Finland. Enligt de planer som uppgjordes på 1960-talet skulle den utges i tjugo band på inemot tusen sidor (Vilppula 1999:407f.).

De första tankarna på en finsk dialektordbok föddes under slutet av 1870-talet. En anspråklös insamling av finskt dialektmaterial påbörjades och ett ordregister inrättades på Finska littera-

tursällskapet. Det fanns en stark vilja att skapa en nationell ordbok över de finska folkmålen. En av förebilderna för de språkforskare som ville skapa en finsk dialektordbok var SRSK (Vilppula 1976:67f.; Onikki-Rantajääskö 2011:544f.).

Diskussionerna om en ordbok över de finska dialekterna fortsatte framför allt vid Finska litteratursällskapet som åtog sig ansvaret för projektet. Den stora frågan var, hur man bäst skulle beskriva dialekterna. Skulle man satsa på språkprov ur dialekterna, en grammatisk beskrivning eller en ordbok? Insamlingen av dialektmaterial fortgick; huvudsakligen med hjälp av ämnesvisa frågelistor som byggde på SRSK. År 1922 hade man samlat in material för ett dialektregister som omfattade en miljon registerkort, men kvaliteten på det insamlade materialet var varierande. Man bedömde att det för en stor dialektordbok behövdes en samling på cirka 2,5 miljoner registerkort (Vilppula 1976:68f.; Onikki-Rantajääskö 2011:556f.).

År 1916 grundades ett aktiebolag som år 1924 ersattes av den finska ordboksstiftelsen *Sanakirjasäätiö*. Stiftelsen organiserades från början så att den leddes gemensamt av vetenskapliga sällskap och föreningar, samt representanter från universitet. Via stiftelsen kunde man ta emot statlig finansiering skattefritt och kanalisera statliga tipsmedel. Målsättningen var att samla in ett digert finskt dialektmaterial, registrera och ordna materialet och sedan påbörja redigeringen av en dialektordbok (Vilppula 1976:68f.; Onikki-Rantajääskö 2011:555ff.).

Men ordboksprojektet kom att bli mer omfattande och långvarigt än man hade planerat. Man ville åstadkomma en så heltäckande beskrivning som möjligt genom att samla in material över variationen i uttal, betydelser och utbredning för dessa.⁴ Den fortsatta insamlingen av material skedde på olika sätt; bl.a. via studenter vid universitetet som samlades och skrev ner ord ur sina egna dia-

4 Det framfördes även alternativa lösningar, vilket kunde ha förhindrat att projektet blev så omfattande (Onikki-Rantajääskö 2011:556f.).

lekter på motsvarande sätt som de Svenska landsmålsföreningarna arbetade med mönster från Uppsala. Studenter utbildades för insamlingsarbete och sändes ut som stipendiater för att göra uppteckningar runtom i landet. Ett omfattande material samlades även in genom att kontinuerligt publicera tidskriften *Sanastaja* med frågelistor som lekmän ombads besvara utifrån sin egen dialekt. Denna form av *crowdsourcing* kom att bli omfattande: på 1930-talet hade man en stab på cirka 1 200 personer som regelbundet besvarade frågelistorna, och insamlingen pågick fram till 1990 (Länsimäki 1976:58f.). Dessutom satsade man på att göra s.k. storsamlingar i vissa socknar som var jämnt utspridda över Finland. Till en början hade man som målsättning att samla in ett heltäckande material över dialekterna i hundra socknar, men stannade sedan för 23 socknar. En avgörande fråga var om ordboken skulle vara lingvistisk eller en beskrivning av folkkulturen. Man stannade slutligen vid en lingvistisk ordbok (Onikki-Rantajääskö 2011:564).

Redigeringen av den finska dialektordboken påbörjades på 1960-talet. Kompletterande insamling av material fortsatte även. I dag omfattar arkivet för SMS omkring 8 miljoner registerkort och är därmed ett av de största dialektarkiven i världen. SMS beskriver ordförrådet i det språk som talades på den finska landsbygden i slutet av 1800-talet och början av 1900-talet. Förutom dialekterna inom Finlands nuvarande gränser innefattar ordboken även finska dialekter som har talats i en del områden i Ryssland, Sverige och Norge, t.ex. i Ingermanland och på Karelska näset i Ryssland, i Tornedalen i Sverige och i Finnmarken i Norge (Vilppula 1999:407 och 2000:457f.).

År 1986 utgavs det första bandet av SMS som då planerades som ett tryckt ordboksverk i 20 band med ca 350 000 uppslagsord (Vilppula 1999:407f.; Onikki-Rantajääskö 2011:563f.). Då Forskningscentralen för de inhemska språken inrättades hörde SMS till hörnstenarna i verksamheten och var ett flaggskepp för Forsk-

ningscentralen. Framtill 2010 utgavs 8 band, avsnittet *a-kurvottaa*, och år 2012 övergick man till elektronisk publicering. Redigeringsarbetet och utgivningen av SMS fortsätter på Institutet för de inhemska språken. Den elektroniska publiceringen har kommit väl i gång; på nätet finns redan avsnittet *kiainen-kyntsöttää* och nytt material publiceras kontinuerligt.

4.2. Den nydanande moderna finska ordboken

Det skulle dröja nästan ett sekel efter att SRSK hade utkommit innan den första helt enspråkiga finska ordboken publicerades. År 1927 beslutade Finlands riksdag att en nuspråklig finsk ordbok skulle påbörjas. Arbetet på ordboken inleddes år 1929. *Nyky-suomen sanakirja* (NSSK) publicerades i sex delar under åren 1951–1961. Haarala (1999:151–155) ger en god beskrivning av ordbokens tillkomst.

Uppdragsgivaren var liksom i fråga om SRSK Finska litteratursällskapet. Denna gång skred en stor redaktion till verket, och år 1939 var materialinsamlingen för NSSK nästan slutförd. Materialet excerperades ur samtidslitteratur, både skönlitteratur och facklitteratur utgiven under perioden 1880–1950, men också äldre språkligt och kulturellt epokgörande verk som Kalevala och Aleksis Kivis produktion beaktades (Hakulinen 1967:106). Det äldre finska skriftspråket (fram till omkring 1810) och dialekterna lämnades utanför ordboksprojektet.

En jämförelse mellan NSSK och SRSK visar att det är en stor skillnad mellan de lemman som upptas i de två ordböckerna. Detta beror på att det finska samhället under början av 1900-talet hade förändrats mycket i jämförelse med samhället under senare hälften av 1800-talet då Lönnrots ordbok utkom. Finland blev självständigt 1917, och genomgick flera krig. Den tekniska utvecklingen skedde oerhört snabbt, och Finland industrialiserades och urbaniserades.

Ruoppila (1962:213f.) illustrerar skillnaderna mellan SRSK och NSSK med hjälp av ordet *jää* 'is' och de sammansättningar och avledningar som ges utifrån *jää*- i dessa ordböcker. I SRSK finns 285 och i NSSK 373 ord på *jää*-. Men bara 70 av dem utgörs av samma ord. En förklaring till detta kan vara att Lönnrot själv konstruerade en stor mängd avledningar vilka utelämnades i NSSK. Dessutom finns ett stort antal nya lemman i NSSK, som beror på att språket hade förändrats därför att omvärlden förändrades.

Varje nytt band av NSSK uppmärksammades såväl inom fennistiken som i dagspressen. Ordboken var en nationell prestation och den fick beröm och goda omdömen (Haarala 1999:154). Hakulinen (1967:106) menar att det är "en topprestation utan like inom vår lexikografiska litteratur" och Ikola (1962:138) skriver "ett storverk när det gäller kodifieringen av vårt språk, [och ett verk] som språkanvändare och forskare inom många kulturspråk avundas oss". Haarala (1999:154) framhäver att "ordboken överträffade alla förväntningar". Men NSSK kom inte att bli fullt så normativ och vägledande som man ursprungligen fastslagit i riksdagens beslut (Haarala 1999:154f.). Ikola (1965:189) understryker att ordboken framför allt är deskriptiv. Då första delen av NSSK kom ut konstaterade Ruoppila (1951) i fråga om ortografin ett faktum som kan förvåna oss i dag: nämligen att den finska ortografin fastslogs först i och med att NSSK hade blivit publicerad. På denna punkt var ordboken synnerligen normativ.

NSSK är en milstolpe inom finsk lexikografi eftersom den är den första professionellt lingvistiska finska ordboken. Samtidigt utgör den en vändpunkt, vilket tydligt belyses av den kritik som riktats mot NSSK. För det första är språkvårdsrekommendationerna inte alltid entydiga. För det andra används ett stort antal förkortningar som bara personer med språkvetenskaplig utbildning förstår. För det tredje bygger indelningen i betydelsegrupper främst på grammatiska kriterier, inte på semantiska. Man kan konstatera att det är en utpräglad akademisk produkt. NSSK har

redigerats av språkforskare och den lämpar sig därför bäst för akademiskt utbildade användare (Haarala 1999:154f.). Eftersom man inte förväntade sig att NSSK skulle bli lönsam ansåg man att staten borde finansiera den (Ruoppila 1962:211). Bedömningen visade sig dock vara felaktig. NSSK erbjöds till förlaget Werner Söderström, som under sextio år gav ut ett stort antal nya upplagor – vissa i tre delar och vissa i sex delar, också pärmen omformades något i upplagorna. Däremot har NSSK aldrig reviderats, utan innehållet har förblivit identiskt – inklusive tryckfelen (Haarala 1999:149). Framför allt har titeln *Nykysuomen sanakirja* 'Nufinsk ordbok' varit en nyckel till den enorma försäljningssuccén – den har genom åren varit en verklig guldgruva för förlaget.

Det kan tilläggas att det vid andra privata förlag började utges konkurrenter till NSSK. Gemensamt för dem var en strävan efter att vara tillgängliga och lättare att använda än NSSK. Dessa ordböcker lämnas utanför diskussionen, eftersom de skulle kräva ett utrymme som inte står i proportion till deras betydelse. Däremot följer vi upp de två allmänfinska ordböcker som utarbetats med statliga medel, *Suomen kielen perussanakirja* 'Finsk basordbok' (SKPS) och *Kielitoimiston sanakirja* 'Språkbyråns ordbok' (KS).

4.3. De två nyaste allmänfinska ordböckerna

För det finska allmänspråket utgavs SKPS i tre band under åren 1990–1994 och som CD år 1997. SKPS var den första elektroniska finska ordboken. Nya upplagor utkom som CD och som betald nätordbok år 2004. Ordboken fick också ett nytt namn *Kielitoimiston sanakirja* 'Språkbyråns ordbok' (KS). Namnförändringen blev aktuell eftersom det hade visat sig att namnet "Basordbok" inte fungerade. En del ansåg att namnet syftade på en basordlista över ordförrådet och andra uppfattade att det var en mindre omfattande ordbok än det i själva verket var. Det mest självklara namnet – *Nykysuomen sanakirja* 'Nufinsk ordbok' var redan upp-

taget och kunde därför inte användas. Man valde därför *Kielitoimisto* (det vill säga ”Språkbyrån”) som en del av namnet. Finska språkbyrån var ett vedertaget namn som var känt som en garant för statligt finansierad språkvård och språkrådgivning. Som ovan nämnts kunde NSSK vara svår att använda för en vanlig språkligt intresserad person. I de nyare allmänfinska ordböckerna, SKPS och KS, har man medvetet försökt ta hänsyn till ordboksanvändare som saknar språkvetenskaplig utbildning (Haarala 1999:154f.).

Då SKPS planerades kunde synpunkter som kommit in från användare av den äldre NSSK beaktas. Redaktörerna för SKPS strävade efter att göra en mer tillgänglig ordbok och att ge entydiga rekommendationer. Antalet lemman var ungefär 100 000 och från och med denna ordbok har antalet lemman i nya upplagor av de allmänspråkliga finska ordböckerna hållits konstant.

När det gäller böjningen av finska nomen och verb kan det beroende på typen av ordstam ske ljudförändringar vid böjning i olika kasus, numerus, tempus osv. I jämförelse med NSSK ges i SKPS mer noggranna uppgifter om de varierande böjningstyperna, så att det blir lättare för användarna att uppfatta informationen. Man lyckades i arbetet med SKPS minska antalet böjningstyper. I NSSK anges 130 böjningsklasser; 85 för nomen, 45 för verb, men i SKPS är bara hälften av klasserna kvar. I både NSSK, SKPS och KS ges uppgifter om uppslagsordets böjning i form av en liten indexsiffra efter ordet.⁵ I NSSK är indexsiffran visserligen så liten att det nästan krävs förstoringsglas för att kunna läsa den. Det har också framgått att många läsare inte alls uppfattat vad indexsiffran betyder. I KS expliceras typen av stadieväxling för orden; en sak som i NSSK överläts till läsarens egen kunskap om reglerna för finsk stadieväxling.

5 I den tryckta KS ges en indexsiffra och böjningsklass, i den elektroniska versionen ges dessutom böjningsparadigm för orden.

4.4. Frågan kring det normativa och det deskriptiva

De äldre finska ordböckerna, NFL och SRSK, är deskriptiva, medan NSSK och dess efterföljare har såväl ett deskriptivt som ett normativt syfte. Detta framträder genom rekommendationer om både ortografi och ordens användning. Spänningen mellan det normativa och det deskriptiva återfinns både i NSSK och i dess efterföljare SKPS och KS. Som ovan framgått är NSSK svår att tolka. Haarala (1999:155) skriver:

I ordartiklarna används grammatiska termer flitigt, och de långa artiklarna är ofta strukturerade enligt grammatiska snarare än semantiska kriterier. En del artiklar är rentav så svårtolkade att det tar flera dagar för en erfaren lexikograf att göra en komplett analys.

När det gäller ortografi godkänns i NSSK t.ex. enbart *sielläpäin* 'på det hållet' och *sieltäpäin* 'från det hållet' sammanskrivet, medan däremot KS är mer liberal och även godkänner sårskrivning: *siellä päin* och *sieltä päin*.

Ett återkommande diskussionsämne i finskan har varit rektionen för verbet *alkaa* 'börja'. Det handlar om valet av infinitivform för det verb som *alkaa* står tillsammans med i verbfraser som 'börja göra (ngt)'. Båda uttrycken *alkaa tehdä* och *alkaa tekemään* har starkt fäste i finska dialekter. Men som den allmänspråkligt korrekta formen valdes strukturen *alkaa* + 1 infinitiv: *alkaa tehdä*. Formen *tekemään* är 3 infinitiv i illativ av verbet *tehdä*. Den senare strukturen markeras i NSSK som dialektal och det anges att man ska använda 1 infinitivformen med verbet *alkaa*. KS (2012) intar samma ståndpunkt, men rekommendationen uttrycks mer analytiskt och läsaren ges information som kan stå som grund för vilkendera infinitivform språkbrukaren själv vill använda. I KS nämns också explicit att båda strukturerna förekommer som vari-

anter i talat språk. I början av 2014 fattade Finska språknämnden beslutet att lyfta fram formen *alkaa tekemään* som en jämbördig parallellform. Den nya situationen avspeglas även i den nyaste upplagan av KS som publiceras under år 2014.

Redan då första delen av NSSK, avsnittet *a-i*, kom ut fäste Penttilä (1955:389) uppmärksamhet vid vissa brister i beskrivningen som varken SKPS eller KS har lyckats lösa på ett systematiskt sätt. Man borde i finska ordböcker ägna mer uppmärksamhet åt att beskriva verbens rektion, vilket är svårt att lösa, eftersom verben inte har fasta formella strukturer utan styrs av semantiska förhållanden. Verbet *syödä* 'äta' kan t.ex. styra ackusativ (ifall något helt och hållet äts upp), partitiv (då fokus är på maten som ämne) eller beroende på kontexten med elativ (för att visa hur litet av något som äts). Ur lexikografens synvinkel innebär detta att man inte kan ge fasta mönster för verbets rektion, utan i stället borde beskriva olika (möjliga) syntaktiska miljöer som verbet kan uppträda i. Det är fortfarande en av de största utmaningarna inom finsk lexikografi att hitta en lösning för hur man systematiskt kunde beskriva verbens olika potentiella syntaktiska mönster. Problemet blir allt mer aktuellt då invandringen ökar och de personer som ska lära sig finska som ett främmande språk blir allt fler.

5. Slutord

Vi har i detta bidrag sammanfattat en (huvud)linje från slutet av 1700-talet fram till 2000-talet för de stora finska allmänspråkliga ordböckerna, som går från Gananders NFL och Lönnrots SRSK till *Nykysuomen sanakirja*, NSSK, och vidare till *Suomen kielen perussanakirja*, SPSK, och *Kielitoimiston sanakirja*, KS. Dessutom finns det åtminstone en sidolinje med insamlingen av det enorma finska dialektmaterialet som i dag bildar ett dialektarkiv på 8 miljoner registerkort. Också den linjen löper delvis från NFL till

SRSK, som båda innefattade material ur det folkliga språkbruket, genom ett stort kompletterande insamlingsarbete till den pågående redigeringen av den finska dialektordboken, *Suomen murteiden sanakirja*, SMS.

År 1896 presenterade E. N. Setälä sitt ordboksprogram som medförde att Finland stannade för en lösning med flera olika nationella temaordböcker: NSSK för det moderna finska allmänspråket, SMS för de finska dialekterna, de etymologiska ordböckerna, SKES och SSA, och ordboken över äldre finskt skriftspråk, VKS. Men som vi visat i vår historiska exposé finns det också andra möjliga förklaringar till hur den finska modellen för de stora nationella ordböckerna har uppstått.

Litteratur

Ordböcker

- KS (2004) = *Kielitoimiston sanakirja*. Helsinki: Kotimaisten kielten tutkimuskeskus och Kielikone Oy. s. a. CD-ROM och betald nätordbok.
- KS (2006) = *Kielitoimiston sanakirja I–III*. Helsinki: Kotimaisten kielten tutkimuskeskus. 3 uppl., reviderad 2012.
- NFL (1937–1940) = Christfrid Ganander: *Nytt Finskt Lexicon. I–III*. Porvoo: Werner Söderström Osakeyhtiö.
- NFL (1997) = Christfrid Ganander: *Nytt Finskt Lexicon*. Alkuperäiskäsikirjoituksesta ja sen näköispainoksesta toimittanut Liisa Nuutinen. Suomalaisen Kirjallisuuden Seuran Toimituksia 676. Kotimaisten kielten tutkimuskeskuksen julkaisuja 95. Hämeenlinna: Karisto.
- NSSK (1951–1961) = *Nykysuomen sanakirja I–VI*. Helsinki: WSOY.
- SKES (1955–1981) = *Suomen kielen etymologinen sanakirja I–VII*. Helsinki: Suomalais-ugrilainen Seura.

- SKPS (1990–1994) = *Suomen kielen perussanakirja I–III*. Helsinki: Painatuskeskus.
- SMS (1985–2008) = *Suomen murteiden sanakirja I–VIII*. Helsinki: Valtion Painatuskeskus/Edita Prima Oy. Nätdorbok 2012– <kaino.kotus.fi/sms> (maj 2014).
- SRSK (1866–1880) = *Suomalais-Ruotsalainen Sanakirja/Finskt-Svenskt Lexikon I–II*. Elias Lönnrot (red.). Helsinki: Suomalaisen Kirjallisuuden Seura.
- SSA (1992–2000) = *Suomen sanojen alkuperä. Etymologinen sanakirja I–III*. Helsinki: Gummerus Kirjapaino Oy.
- VKS (1985–1994) = *Vanhan kirjasuomen sanakirja*. Helsinki: Painatuskeskus. Nätdorbok 2014– <kaino.kotus.fi/vks> (maj 2014).

Annan litteratur

- Anderson, Benedict (2003): *Imagined communities. Reflections on the origin and spread of nationalism*. London/New York: Verso.
- Andersson, Håkan (1999): Skolan, modersmålet och litteraturen. I: *Finlands svenska litteraturhistoria. Första delen: Åren 1400–1900*. Helsingfors: Svenska litteratursällskapet i Finland/Stockholm: Bokförlaget Atlantis, 416–423.
- Fagerlund, Rainer, Kurt Jern & Nils Erik Villstrand (2001): *Finlands historia 2*. Esbo: Schildts Förlags Ab.
- Haarala, Risto (1999): Vem är finska ordböckerna gjorda för? I: *Nordiska studier i lexikografi 4*. Helsingfors, 149–158.
- Hakulinen, Lauri (1961): *The Structure and Development of the Finnish Language*. Bloomington: Indiana University Publications Uralic and Altaic Series 3.
- Hakulinen, Lauri (1967): Suomen sanakirjoista. I: *Kirjoituksia Suomen kielestä*. Tietolipas 5, 83–107.
- Hakulinen, Lauri (1979): *Suomen kielen rakenne ja kehitys*. 4. omarbetade uppl. Helsinki: Kustannusosakeyhtiö Otava.

- Hovdhaugen, Even, Fred Karlsson, Carol Henriksen & Bengt Sigurd (2000): *The History of Linguistics in the Nordic Countries*. Jyväskylä: Societas Scientiarum Fennica.
- Häkkinen, Kaisa (1998): Retki menneeseen maailmaan. I: *Virittäjä* 1998, 634–640. (Recension av NFL 1997).
- Ikola, Osmo (1962): Nykysuomen sanakirja valmis. I: *Sananjalka* 4, 136–38. (Recension av NSSK).
- Ikola, Osmo (1965): Die letzten Bände des Neufinnischen Wörterbuches. I: *Studia Fennica XII*, 189–190. (Recension av NSSK).
- Klinge, Matti (1996): *Finlands historia* 3. Esbo: Schildts Förlags Ab.
- Koskinen, Yrjö (1867): Suomalainen ja Ruotsalainen Sanakirja. 1 Vihko. Helsingissä 1866. I: *Kirjallinen Kuukausilehti* 1867. 16–19. (Recension av första häftet av SRSK).
- Koukkunen, Kalevi (2004): Lexikografisk granskning av *Finskt-Svenskt Lexikon* av Elias Lönnrot. I: *LexicoNordica* 11, 133–155.
- Länsimäki, Maija-Liisa (1976): Vapaaehtoiset vastaajat. I: Tuomo Tuomi (red.): *Sanojen taivalta. Puoli vuosisataa Sanakirjasäätiön toimintaa*. Helsinki: Suomalaisen Kirjallisuuden Seura, 58–66.
- Onikki-Rantajääskö, Tiina (2011): Mestari–kisälli-mallin soveltuvuus fennistiikan historiaan. I: *Virittäjä* 2011, 542–475.
- Penttilä, Aarne (1955): Nykysuomen sanakirjan II ja III osa. I: *Virittäjä* 1955, 388–389. (Recension av NSSK).
- Rapola, Martti (1931): Muudan kansallisten tieteen peninkulmapatsas. I: *Valvoja-Aika* 1931, 199–204. Helsinki: Suomalaisen Kirjallisuuden Seura. (Recension av SRSK).
- Rapola, Martti (1932): Kansankielen sanakirjahanke ja sen aatteellinen tausta. I: *Suomalainen Suomi* 15, 39–49.
- Romppanen, Birgitta (2001): *Från målspråk till källspråk. Utvecklingen av den finsk-svenska och den enspråkiga finska ordboken*. (Acta Wasaensia 91, Språkvetenskap 18). Vasa: Universitatis Wasaensis.

- Ruoppila, Veikko (1951): Nykysuomen sanakirja. I: *Uusi Suomi* 20.12.1951. (Recension av NSSK).
- Ruoppila, Veikko (1962): Nykysuomen sanakirja I–VI. I: *Finnisch-Ugrische Forschungen* 34, 210–216. (Recension av NSSK).
- Varpio, Yrjö (1999): Den finskspråkiga litteraturen. I: *Finlands svenska litteraturhistoria. Första delen: Åren 1400–1900*. Helsingfors: Svenska litteratursällskapet i Finland/Stockholm: Bokförlaget Atlantis, 351–355.
- Vilppula, Matti (1976): Murre-sanakirjan etäistä taustaa. I: Tuomo Tuomi (red.): *Sanojen taivalta. Puoli vuosisataa Sanakirjasäätiön toimintaa*. Helsinki: Suomalaisen Kirjallisuuden Seura, 67–70.
- Vilppula, Matti (1999): Ordbok över finska dialekter och evigheten. I: *Nordiska studier i lexikografi* 4. Helsingfors, 407–411.
- Vilppula, Matti (2000): Svenskan och de finska dialekterna. I: *Folkmålsstudier* 39, 457–464.

Klaas Ruppel
avdelningsföreståndare
Institutet för de inhemska språken
Berggatan 24
FI-00100 Helsingfors
klaas.ruppel.@kotus.fi

Caroline Sandström
huvudredaktör
Institutet för de inhemska språken
Berggatan 24
FI-00100 Helsingfors
caroline.sandstrom@sprakinstitutet.fi

Blinkenberg og Høybyes *Dansk-Fransk* og *Fransk-Dansk Ordbog* – fra seddel-samling til trykte ordbøger og til online-udgave

Lene Schøsler¹

The Danish-French and French-Danish dictionaries elaborated by Andreas Blinkenberg and Margrethe Thiele, subsequently by Blinkenberg and Poul Høybye are influential and large bilingual Danish dictionaries, the 1997 editions comprising respectively 172,000 Danish lemmas and more than 1 million translations of words and expressions into French, and some 83,000 French lemmas and more than 1 million translations into Danish. The paper discusses the lexicographical principles of the dictionaries and presents a number of predictable and unpredictable problems, which arise when transferring printed dictionaries to an electronic platform.

1. Introduktion

Den første større dansk-franske ordbog er Thor Sundby & Jean Guillaume Euchaire Baruël (S&B): *Dansk-Norsk-Fransk Haand-Ordbog* 1-2 (1883-84). Forudsætningen for de senere ordbøger er Margrethe Thiele, der fungerede som oversætter til fransk for Det Kongelige Danske Videnskabernes Selskab. For at sætte sig ind i de forskellige videnskabers terminologi gjorde hun optegnelser ud fra franske specialværker. Efterhånden som det indsamlede materiale voksede i omfang (seddelkartoteket var til sidst på ca. 100.000 seder), opstod hos hende tanken om at udarbejde en dansk-fransk ordbog til afløsning af S&B. Fra 1926 indgik hun samarbejde med Andreas Blinkenberg for at forberede udgivelse af en egentlig ord-

1 Tak til Gunhild Dyrberg og Lilian Stage for hjælp ved udarbejdelsen af denne fremstilling.

bog, der i 1937 forelå trykt under titlen A. Blinkenberg & M. Thiele: *Dansk-fransk Ordbog* (B&T). De forskellige udgaver er anført i litteraturlisten.

Poul Høybye var fra 1928 medredaktør ved B&T's dansk-franske ordbog og blev i 1937 sideordnet redaktør. Det store ordbogsværks videreførelse blev sikret i 1984 ved oprettelsen af *Blinkenberg-Høybye-fondet for dansk-fransk leksikografi*, hvis midler udelukkende stammer fra salg af ordbøgerne. Den sidste trykte dansk-franske udgave (D-F), Blinkenberg & Høybye 1991 (herefter B&H), omfatter ca. 172.000 danske opslagsord og over 1 mio. franske gengivelser af ord og ordforbindelser. Udgaverne fra fransk til dansk (F-D) er udarbejdet senere (se litteraturlisten). Sidste trykte F-D-udgave er fra 1997; den indeholder ca. 83.000 opslagsord og over 1 mio. danske gengivelser. Denne forskel i antal opslagsord mellem ordbøgerne skyldes i høj grad en typologisk forskel på dansk og fransk, idet dansk hovedsageligt danner nye ord ved sammensætninger af to eller flere enkeltord, hvorimod fransk forbinder enkeltord ved hjælp af præpositioner.

Blinkenbergs nyskabende indsats inden for den tosproglige leksikografi manifesterede sig på flere måder, først og fremmest ved indførelsen af en konsekvent betydnings-opdeling og en overskuelig ordning af stoffet med opdeling i *hoved* og *hale* (se afsnit 2). Resultatet af arbejdet blev en systematisk kortlægning af ordforrådet i nyere og nyeste fransk i dets mangfoldige afskygninger. Beskrivelsen omfattede ikke blot det almindelige sprog (litterært sprog og dagligsprog), men også – som en særlig fornyelse – alle arter af fagsprog, fra håndværkssprog, juridisk og teknisk sprog til videnskabernes sprog (se afsnit 3.2). Disse principper er bevaret i alle senere udgaver. Formålet for ordbøgerne er (jf. forordet til 1991-udgaven) at dække et bredt spektrum af det danske og det franske ordforråd, så ikke kun det almene sprog er fyldigt repræsenteret, således at kun meget sjældne ord ikke er medtaget, men

også ord fra alle typer af fagsprog, såvel centrale discipliner som økonomisk, juridisk og teknisk sprog, men også ord fra håndværksfag, politik, naturvidenskab, film, massemedier og humanistiske videnskaber, og at bevare et udvalg af litterære ord der anvendes af klassiske danske forfattere, således at ordbogen kan anvendes af udlændinge til at læse dansk litteratur fra Holberg, H.C. Andersen og Kierkegaard til moderne tid.

D-F fra 1991 og F-D fra 1997 blev overført til søgbar cd-rom under ledelse af Jens Rasmussen. I den forbindelse blev den oprindeligt manuelt udarbejdede ordbog indscannet og overført til et nu forældet elektronisk format, der har vist sig ikke at kunne genbruges. Koder og opmærkninger er blevet indsat manuelt, hvilket desværre ikke er sket konsekvent. Artiklerne har undertiden uens opbygning. Disse forskellige forhold har gjort, at cd-rom-udgaven ikke umiddelbart kunne konverteres til nyt elektronisk format.

Den nuværende bestyrelse for *Blinkenberg-Høbye-fondet* (Gunhild Dyrberg, Lilian Stage og Lene Schøsler, fra 2000) har i en årrække arbejdet med revision og konvertering af ordbøgerne. De tre medlemmer af fondet har måttet bruge uventet meget tid på selve konverteringsprocessen, i stedet for – som det ellers var forudset – at kunne overføre data til en ny, forbedret artikelstruktur og koncentrere sig om opdatering af indholdet. Tankerne om en ny, ideel artikelstruktur er fremstillet i Dyrberg & Stage (2005). Problemerne i forbindelse med konverteringsprocessen vil blive illustreret i afsnit 3.1 med henvisning til illustrationen af artikelgennemgangen i afsnit 2.2.

Overførsel til elektronisk netudgave er sket i 2012 gennem aftale mellem Nyt Nordisk Forlag og ordbogen.com. B&H er tilgængelige på internetadressen www.ordbogen.com, hvor man kan abonnere på B&H (og andre ordbøger). Fondets bestyrelse arbejder videre på databasen. I denne nordiske sammenhæng vil jeg koncentrere mig om den danske side af ordbøgerne (D-F).

2. Artiklernes opbygning

De leksikografiske principper for B&H er konciperet af Blinkenberg og udgør et brud med samtidens mere intuitive stoforganisering som den f.eks. findes hos S&B.

2.1. De leksikografiske principper

Ordbøgerne er opdelt i artikler. Såfremt der findes flere opslagsord med samme stavemåde, starter artiklerne med et homografnúmer (romertal fulgt af punktum). Homografer skelnes ud fra følgende kriterier: ordklasseforskel, væsentlig konstruktionsforskel, semantisk forskel.² Først i online-udgaven fastlægges fast rækkefølge af homografer efter ordklasse: substantiv, adjektiv, talord, pronomen, verbum, adverbium, præposition, konjunktion, interjektion. Eks. *I. pode* (substantiv), *II. pode* (verbum).

Opslagsordet følges af en ækvivalent, altså oversættelse til målsproget. Ækvivalenten repræsenterer den mest dækkende gengivelse af opslagsordet i den pågældende betydning eller sammenhæng. I de tilfælde, hvor der eksisterer flere anvendelige gengivelser, er de vigtigste af disse nævnt. Den første ækvivalent er i reglen den, der viser den almindelige anvendelse. Denne del af artiklen refereres der til som *hovedet*, den sidste del som *halen*.³ I hovedet er betydningsforskelle angivet med kapitæler (A, B, C ...), hovedsagelig efter semantiske kriterier, og med henblik på relevans for målsproget. Sidstnævnte princip adskiller naturligvis tosprogede ordbøger som B&H fra etsprogede ordbøger. De to principper (betydningsforskelle på kildesprog og målsprogsrelevans) kan undertiden

2 Principper med konsekvent homografskelnen på basis af formelle kriterier er en leksikografisk nyskabelse fra Blinkenbergs side. Dette princip har efterfølgende inspireret den danske leksikografi, især den tosprogede.

3 Opdelingen i *hoved* og *hale* er en nyskabelse fra Blinkenbergs hånd, der har haft varig indflydelse på den danske leksikografi.

være i konflikt med hinanden. I større artikler er ækvivalenterne markeret med fortløbende arabertal, der bruges som henvisning mellem hoved og hale for at spare plads. Sådanne krydshenvisninger er nødvendige i tosprogede, men ikke i etsprogede ordbøger. I online-udgaven er pladsbesparelse ikke længere nødvendig, og arabertallene er erstattet af den relevante ækvivalent. Denne erstatning har været årsag til konverteringsproblemer (se afsnit 3).

Efter hovedet kommer som nævnt *halen* med ordforbindelser opbygget alfabetisk og ikke anført under de respektive hovedbetydninger, som det er tilfældet for ODS for eksempel. Dette valg skyldes bl.a., at det ikke altid er muligt at henføre en ordforbindelse til en given hovedbetydning. Den alfabetiske ordning afhænger af *det karakteristiske ord* for ordforbindelsen, der kan stå foran eller efter opslagsordet. I halen henvises som nævnt med arabertal i parentes til de nummererede underopdelinger i hovedet, hvilket giver en kompleks, men præcis information til brugeren, men det er en krævende struktur for redaktøren. Alfabetiseringsrækkefølgen er angivet med fed skrift ved første eksempel. Ved forbindelser og sammensætninger er opslagsordet i de trykte udgaver erstattet af en tilde. I onlineudgaven er opslagsordet skrevet ud. Dette har forårsaget problemer ved konverteringen (se afsnit 3).

I ordbøgerne gives der oplysninger om bl.a. ordklasse, stil, emne, grammatik, og for fransk angives lydskrift (i F-D ordbog).

I onlineudgaven er der indført farvemarkeringer: Kildesproget er angivet med blå skrift, målsproget med sort skrift. Ordklasse, stil, grammatik, sprog og forklaringer står med grøn skrift. Emnet står med rød skrift.⁴ Det valgte layout giver mulighed for hurtig forståelse af oplysningstyper. Opstillingen i onlineudgaven er mere luftig end i papirudgaven, og der er indføjet linjeskift ved hver underopdeling.

4 Ved fremmedord angives oprindelsessprog (f.eks. latin, engelsk). Beklageligvis har det ikke her været muligt at gengive farvemarkeringerne – der henvises til at konsultere onlineudgaven.

2.2. Illustration af de leksikografiske principper

I dette afsnit vil jeg illustrere de leksikografiske principper ud fra verbet *ligge*. Jeg vil først vise hvordan S&B (1883) præsenterer dette verbum, derefter B&T's fremstilling (1937). Blinkenberg's principper fra B&T følges efterfølgende i B&H-udgaverne, uden leksikografiske ændringer.

S&B opdeler i to artikler (homografer), nummereret med arabertal: 1. *ligge v.n.* (= verbe neutre, dvs. intransitivt) og 2. *ligge v.a.* (= verbe actif, dvs. transitivt). Under 2. *ligge* henvises til *likke*, som sideform til *slikke* P & *obsc. fig.* – Ens Maas lécher (*ou baiser*) le cul à qn. Betydning 2 findes ikke i de senere ordbøger af B&T og B&H. Sideformen findes ikke heller i DDO, men dog i ODS: **likke**, *v.* [ˈlegə] (*undertiden skrevet ligge, lægge, se ndf. samt u. II. Fos 1; jf. ogs. sideformerne til Kuttelikker*). Dansk er skrevet med gotisk skrift, hvilket ikke er gengivet her. I det følgende omtales kun første artikel.

Første artikel er organiseret således, at der først gives nogle oversættelsesforslag, herefter nogle hyppige forbindelser hvis rækkefølge ikke er indlysende. Siden følger ordforbindelser organiseret efter adjektiver og adverbier: *bi, brak, død, fjernere, for, fuldt, godt, hen* osv. Rækkefølgen er alfabetisk, men det er ikke indlysende, hvilket ord alfabetiseringen tager sit udgangspunkt i. F.eks. er *ligge fjernere* alfabetiseret efter bogstavet *f*, men eksemplificeret ved *Intet ligger mig fjernere*, – man kunne have forventet alfabetisering ud fra *Intet*, fordi det er en fast del af udtrykket. Slutteligt er angivet sammensætninger, alfabetisk organiserede. Uddrag af artiklen følger her:

1. *ligge être couché, reposer, être placé ou situé, se trouver, être ; (have sit Natteleje) coucher ; (om Høne) couvrir, (om Elv [norsk⁵]) être pris; lade N. – laisser; (opgive)*

5 Norsk er angivet ved hjælp af en stjerne.

laisser là, (en Procès) délaisser, (af Glemsomhed) laisser, oublier (d'emporter) qch; lade N. – fremme laisser qch en vue; lad det –! n'y touchez pas! hvor –r Skylden? à quoi cela tient-il? hos hvem –r Skylden? à qui la faute? [..] m. *adj.* & *adv.* – bi *mar.*⁶ être à la cape; – bi for Takkel og tov capéer (*ou* être cape) à sec ; – brak *agr.* être en friche ; [..] Sammensætninger : Liggedage *pl. com. mar.* jours *m.pl.* de planche *m* starie *f*; – dags-Penge [..]

B&T (1937) introducerer en klar opdeling mellem *hoved* og *hale*. Der skelnes i hovedet mellem to hovedbetydninger A og B. Krydshenvisningerne mellem *hoved* og *hale* er anført med fortløbende arabertal. Halen er alfabetisk organiseret.

A) (= ~ *ned*) (1) être couché (2) être étendu (3) être allongé ; – (= *have sit Natteleje*) (4) coucher (5) (*undert., is. om visse Dyr*) gîter; – (= *hvile*) reposer; – (*om Høne, = ~ paa Æg*) (7) couvrir – (*mar., om Dybgaaende*) (8) caler.
B) (= *være anbragt*) (9) être placé (10) se trouver (11) être (12) (*uordentligt*) traîné ; – (= *være beliggende*) (13) être situé (14) se situer.

I forsk. Forb.: ~ *paa Aarerne mar.* se tenir sur les avirons [..] ~ *imellem A og B* (10) (11), être bordé par A et B; ê. compris entre A et B [..]

I onlineversionen af B&H beholdes hovedinddelingen i A og B, men arabertallene til krydshenvisning mellem *hoved* og *hale* er fjernet og erstattet af ækvivalenten. Denne ændring har været kilde til konverteringsproblemer, se nærmere i afsnit 3.1. Der er tilføjet idiomatisk information (jf. nedenfor: *oversættes ofte ved il y a, il y avait, etc.*). Ordforbindelserne er reviderede. De forkortelser, der

6 Maritime udtryk (*mar.*) er angivet med et anker, kommercielle udtryk (*com.*) med en hermesstav.

var nødvendige i de trykte udgaver vedrørende mål- og kildesprog, er skrevet ud.

ligge (vi) (lå, ligget)

A (= ligge ned)

être couché

être allongé

(udstrakt) être étendu

(= overnatte) coucher

(= hvile) reposer

(om visse dyr) gîter

(= ligge på æg) couvrir

(mar om dybgående) caler

B (= befinde sig)

se trouver

être

(uordentligt) traîner

(= være beliggende) être situé, se situer

(oversættes ofte ved il y a, il y avait, etc.)

- ligge 10 fod **agter** (for) (mar) caler 10 pieds derrière (devant)
- ligge nord **an** (mar) avoir le cap au nord
- hvad ligger vi an? (mar) où est le cap?
- bringe et skib til at ligge misvisende øst an (mar) amener le cap du navire sur l'est magnétique
- ligge an mod porter sur, appuyer sur, [bien] reposer sur
- klæde staget, hvor sejlet ligger an (mar) fourrer l'étai au portage de la voile
- vide hvad der ligger **bag** ved connaître le dessous des cartes
- ligge **bi** (mar) être à la cape, capeyer
-
- ligge **imellem** A og B se trouver entre A et B, se situer entre A et B; être compris entre A et B
- ligge på årerne (mar) rester sur les avirons

2.3. Opsummering af principperne i Blinkenberg & Høybye

Blinkenbergs leksikografisk nyskabende principper, især for den tosprogede leksikografi, er først og fremmest den konsekvente etablering af homografiskelnen og opdelingen i *hoved* og *hale* (se afsnit 2.1). I *hovedet* gives en oversigt over opslagsordets forskellige ækvivalenter på målsproget, i *halen* gives eksempler på oversættelse af opslagsordet i forskellige ordforbindelser. Ved komplekse artikler, der har adskillige ækvivalenter, er disse nummererede med arabertal i parentes, og i halen er der henvisning til disse, hvilket giver pladsbesparelse. Fordelen ved en sådan opdeling i to klart definerede dele, er at en bruger med solid sproglig viden hurtigt kan finde en dækkende gengivelse af opslagsordet i en given betydning ved blot at kaste et blik på hovedet. Til gengæld må man konstatere, at alfabetiseringen af eksemplerne undertiden følger lidt tilfældige principper: Snart er eksemplet alfabetiseret efter et adjektiv, snart efter et substantiv, et verbum, en nægtelse eller en præposition. Ifølge B&H (1997) er alfabetiseringsprincippet i halen valgt af hensyn til brugeren, idet der siges: "I en lang eksempelsamling gør dette princip det let for brugeren at finde frem til det udtryk, der søges." Man kan dog sætte spørgsmålstegn ved, hvorvidt dette princip altid hjælper brugeren, når brugeren ikke på forhånd kan vide, hvilket ord i ordforbindelsen alfabetiseringen er foretaget efter. Men det er et generelt problem for alle ordbøger. Se videre afsnit 3.1. vedrørende overvejelser om ændret struktur. I netudgavens reviderede artikler er opstillingen overskueliggjort og simplificeret ved hjælp af linjeskift, markering foran hver ny forbindelse og valg af verbalform er harmoniseret (samme basisform er valgt). Rækkefølgen er stadigvæk alfabetisk, med fremhævelse ved hjælp af fed skrift.

3. Problemer i forbindelse med udarbejdelsen af den seneste onlineudgave

3.1. Konverteringsproblemer

B&H er blevet overført fra en trykt udgave med søgbar cd-rom til en netversion placeret under ordbogen.com. Dette arbejde startede i 2001. Den nuværende bestyrelse forventede, at der forelå en elektronisk version, der umiddelbart kunne overføres og modificeres ud fra bestyrelsens beslutning om moderniseringsprincipper. Det var hensigten, at den nye artikelstruktur skulle være mere brugervenlig såvel for dansk- som for fransksprogede. Der var tale om en mere eksplicit struktur, inspireret af den der benyttes i Den Danske Ordbog. I forhold til de gamle udgaver af B&H betød det, at der ved opslagsordet først angives en betydningsbeskrivelse i parentes, derefter ved valenskerne en valensbeskrivelse, og til sidst ækvivalenter på målsproget. Det var vores opfattelse, at ordbogsbrugerne herved kunne få de bedste muligheder for at foretage et kvalificeret valg af oversættelse; vi har nemlig erfaret, at et af problemerne ved de tidligere udgaver er, at brugerne har svært ved at erkende, hvilken betydning det danske opslagsord har. Derfor har de også svært ved at vælge den rigtige oversættelse. I en ny artikelstruktur kunne vi endvidere foretage en revision af hoved-hale-princippet, således at ordforbindelserne placeres direkte under den betydning af opslagsordet, som de naturligt hører ind under, i de tilfælde hvor det er muligt (jf. afsnit 2.1 og 2.3).⁷ Vi ønskede desuden at foretage en parallel revision af ordbøgerne (F-D og D-F) med henblik på at gøre dem indbyrdes konsistente i udformning og ækvivalensvalg. Da vi imidlertid ikke har kunnet få eksterne midler til en grundlæggende ændring af strukturen, og da det viste

⁷ Vi havde med andre ord overvejet en modifikation af hoved-hale-princippet på grund af de i afsnit 2.3. nævnte brugerrelaterede overvejelser, hvilket ville have været et brud med Blinkenberg-traditionen.

sig at en konvertering fra den oprindelige struktur til en ny ideel struktur var umulig, har vore planer ikke kunnet realiseres.

Som nævnt løb vi ind i problemer, da det viste sig at den database, der dannede grundlag for den sidste trykte udgave og for dens cd-rom, ikke kunne anvendes til en mere eller mindre automatisk konvertering til et nyt redigeringsformat. Vi besluttede derfor efter flere års frugtesløse forundersøgelser at konvertere den eksisterende base til et mere enkelt format inden en egentlig revision. Selv dette beskedne mål var særdeles vanskeligt. Udgangspunktet var jo oprindeligt en håndskreven seddelsamling, der manuelt var overført til tryk, senere til fotosats, indscannet og manuelt opmærket. De tidligere redaktørers arbejde har været forbilledligt omhyggeligt, men især i artikelopbygning og opmærkning er der sket fejl og uregelmæssigheder. Jeg vil i det følgende give tre eksempler.

3.1.1. Oplysningernes hierarki

Hovedprincippet for angivelse af oplysningernes hierarki er, at den danske side starter med romertal, herefter bogstaver og herefter arabertal. Det betyder at konverteringsprogrammer kan gå systematisk frem efter sådanne angivelser. Imidlertid har visse artikler rækkefølgen romertal, bogstaver, romertal og arabertal. Det betyder, at programmet opfatter de indlejrede romertal som starten på en ny artikel med nyt homograf-nummer. Resultatet af den automatiske overførsel betyder, at artiklerne deles i flere artikler. Principielt kan man ikke forudse, hvilke artikler har indlejrede romertal. Disse fejl skal derfor spores og rettes manuelt (over 1500 fejl er rettet).

3.1.2. Erstatning af arabertal med ækvivalenter

Som det fremgår af opslaget *ligge* i B&T, hovedbetydning B, kan *ligge* oversættes med bl.a. (10) *se trouver* og (11) *être*. I ordforbindelsen ~ *imellem A og B* (10) (11), sker der to ændringer ved konverteringen, den første er erstatningen af tilden med opslagsordet

(se mere herom næste afsnit), den anden er erstatningen af arabertallene med ækvivalenterne. Den automatiske erstatning vil da for det første ændre ~ *imellem A og B* (10) (11) til *ligge imellem A og B* og for det andet anføre ækvivalenten *se trouver, être*, hvilket bliver nonsens, fordi det er mangelfuldt, da oversættelsen af *imellem A og B* så mangler ved ækvivalenterne. Der skal stå *se trouver entre A et B, être entre A et B* som anført i onlineversionen (afsnit 2.2.). Det gælder her som for fejltypen anført i afsnit 3.1.1, at man principielt ikke kan forudse i hvilke artikler sådanne fejl findes. Disse fejl skal derfor spores og rettes manuelt (1032 artikler er blevet kontrolleret).

3.1.3. Erstatning af tildemarkering med opslagsordet

Problemet med erstatning af tildemarkeringen med opslagsordet opstår ved ord med ændring af ordets form. Opslagsordets grundform indsættes automatisk i stedet for tilde, hvilket giver et tilfredsstillende resultat ved ord uden ændring af grundformen (se *ligge*-eksemplet afsnit 3.1.2.), men ved ord som *egen – eget – egne* og *eksempel – eksemplet – eksempler*, giver indsættelse af grundformen en fejl, når grundformen indsættes i stedet for en bøjet form. F.eks. blev *stå på egne ben* konverteret fra *stå på ~ ben*, hvilket uden manuel korrektion ville give det fejlagtige resultat: *stå på egen ben* (ca. 6000 fejl i forbindelse med tildeerstatning er rettet).

3.2. Detaljeringsgrad

Såvel B&T som B&H er imponerende i deres dækning af det danske og franske ordforråd, så det er et værk med høj detaljeringsgrad i artiklerne. Det kan man let overbevise sig om ved at konsultere de tilsvarende ordbøger, selv den sidste udgave af Gyldendals ordbøger, som er relativt omfattende (Hansen et al. 2001). Jeg minder om, at de mange ord oprindeligt er indsamlet på sedler af et antal redaktører og mange interesserede franskkyndige per-

soner. De forskellige personer har haft forskellige fagsproglige interesser, hvilket afspejler sig i varierende dækning af emner, og de forskellige ansvarshavende redaktører har ikke ønsket at afbalancere emnefordelingen,⁸ imidlertid er alment sprog er langt den dominerende kategori. Det er umuligt at vide, hvordan fordelingen mellem de forskellige emneområder ideelt set burde være, men på grund af B&H's oprindelseshistorie er der opstået skævheder. Jeg anfører nedenfor (figur 1) nogle eksempler på fordeling mellem fagområderne, dels for at vise bredden i ordbøgernes dækning af emner, dels for at vise, hvor store forskelle der er i antal artikler inden for de forskellige emner. Særligt påfaldende er det, hvor stort det teknisk-naturvidenskabelige ordforråd er i forhold til det humanistiske. Det er indlysende, at i de tilfælde, hvor der er mange artikler inden for et emne, opnår ordbogen en højere detaljeringsgrad. Konsekvent detaljeringsgrad er formodentlig kun mulig for (nystartede) ordbogsprojekter, der bygger på konsekvent udnyttelse af et afbalanceret korpus.

tech = teknik = 3747	sport = 667
zoo= zoologi = 3595	relig = religion = 286
mar = maritim = 3640	herald = heraldik = 99
bot = botanik = 3432	philat = filateli = 18
mil = militær = 1517	alim = fødevareindustri = 7
com = handel = 1369	arts = bildende kunst = 6
écon = økonomi = 779	

Figur 1: Fordeling på et antal emner, onlineudgaven.

3.3. Udvidelse af ordforrådet

Tidligere, i trykte udgaver, krævede en fornyelse af ordforrådet, at man måtte overveje at fjerne ældre ord, hvilket kunne være i strid med ønsket om at bevare det ældre litterære ordforråd. I en net-

⁸ Den nuværende bestyrelse arbejder aktivt med at revidere og især at opdatere emnerne.

version er der ikke umiddelbart problemer med plads – se dog afsnit 3.4. I princippet kan man derfor i netordbøger indføre nye ord uden at være nødt til at slette gamle ord. Men visse nye ord er rene modefænomener – hvor meget skal man tage med? Og hvor meget skal man modificere ord der har ændret betydning? Lad os skelne mellem to forhold, for det første indførelse af nye ord på grund af nye objekter eller begreber. Det gælder f.eks. ord som *alkolås* og *havmøllepark*, det gælder også nye ord inden for regnskab og økonomisk sprog. For det andet er der nye betydninger af kendte ord eller nye ord eller sammensætninger der opstår i forbindelse med at samfundet ændrer sig. Det gælder f.eks. *nabohjælp*, *whistleblower* og *flyverskjul*. I det sidste tilfælde drejer det sig måske blot om kreativ sprogbrug og ikke om egentlig sprogændring, der bør registreres? Det er en udfordring for alle ordbøger, men måske især for de trykte. Et andet problem forbundet med sprogets udvikling er stilmarkeringer: Hvad der tidligere var lav stil eller injurierende, kan senere blive acceptabelt – tænk blot på ordet *racist* i dansk, der i forbindelse med en sag om Pia Kjærsgaard fra Dansk Folkeparti blev dømt som ikke længere at være et injurierende ord. Samtidig opstår der hele tiden nye lavstilsudtryk og kraftudtryk.

Et interessant eksempel på at sproget ændrer sig i takt med historisk betingede samfundsændringer, er ordet *nødhjælpsarbejde*. I ordbogen fra 1937 var begreberne *nød* og *nødhjælpsarbejde* tæt forbundet med den økonomiske situation i 1930'erne med voldsom arbejdsløshed. Det fremgår også direkte af definitionen i ODS:

Nød-hjælps-arbejde, *et.* (jf. Nødarbejde 2; især *emb.*) arbejde, der iværksættes til afhjælpning af arbejdsløshed.
LovNr.529²²/₁₂ 1921. §32.

I overensstemmelse med ODS blev den danske term oprindeligt oversat til det meget lidt mundrette: *travail public [décrété d'intérêt général et] entrepris pour (el. destiné à) combattre (el. pallier) le chô-*

mage. Med tilføjelsen: *Mk entreprendre de grands travaux d'utilité publique pour employer les chômeurs.*

I vore dage dækker termen *nødhjælpsarbejde* en anden virkelighed, og i overensstemmelse hermed oversættes den i netudgaven med *action humanitaire d'urgence*. Da B&H som anført ønsker at muliggøre forståelse og oversættelse af ældre sprog, bevarer netudgaven den oprindelige oversættelse ved siden af den moderne.

3.4. Pladsproblemer før og nu

De trykte ordbøger har indlysende problemer med plads. Det betyder for det første, at redaktører er fristede til at fjerne ældre stof for at få nye ord ind – hvilket **ikke** har været B&H's politik, eftersom denne ordbogs målsætning som nævnt har været at kunne bruges til at forstå også ældre værker. Behovet for at begrænse omfanget mest muligt af økonomiske årsager har medført, at trykte artiklers opbygning er blevet fortættede. Det fremgår af afsnit 2, at der er indført mange forkortelser og symboler, både af opslagsordet og af den videre kontekst, hvilket undertiden gør det vanskeligt at konsultere de mere komplekse artikler.

Med etablering af elektroniske platforme, der principielt ingen pladsbegrænsninger har, kan de mange forkortelser af opslagsord opløses og skrives fuldt ud. Det har vi gjort i de reviderede artikler i den elektroniske base.

Eftersom B&H nu ligger på en kommerciel platform, er vi imidlertid impliceret i overvejelser angående den teknologiske udvikling og brugernes ønske om hurtig adgang til information. På længere sigt kan det betyde, at ordbogen.com vil satse på nye applikationer, der kan downloades og læses på små skærme (smartphones eller iPhones). Hvor B&H's "varemærke" tidligere var at kunne tilbyde brugerne den korrekte og fyldigst mulige information, kan det blive et handicap at være så omfattende, når teksten skal ned på en lille skærm. Dette er en udfordring for en

stor ordbog som B&H. Ordbogen.com arbejder med forskellige løsningsmodeller. Disse omfatter bl.a. at spore tidligere opslag af samme bruger og indsnævre interessefeltet i overensstemmelse hermed eller at reducere antallet af ækvivalenter.

Som et resultat af B&H's høje kvalitet på såvel dansk- som fransksiden har den kommercielle udbyder i øvrigt ønsket at benytte dansksiden som udgangspunkt (nemlig som *pivot*) for nye ordbøger til og fra dansk. Det har bestyrelsen imidlertid ikke ønsket at involvere sig i.

4. Indflydelse fra Blinkenberg & Høybye

B&T og B&H's leksikografiske principper har haft afgørende indflydelse på den danske tosprogsleksikografi, der i stor udstrækning har overtaget disse principper. Ordbøgernes artikler er ofte blevet brugt som model for andre, mindre dansk-fransk-danske ordbøger. Typisk har redaktørerne for de mindre ordbøger nedkortet B&H-artiklerne og udeladt stof. Den konkrete indflydelse fra B&H kan undertiden afsløres ved utilsigtede fejl der ukritisk er overtaget af de mindre ordbøger. Et eksempel er substantivet *paradisæble*, der i B&H (1991) oversættes med *pomme de paradis* eller *pomme d'api*, hvilket ikke er korrekt. I Gyldendals ordbog overtages den fejlagtige oversættelse *pomme d'api*. Den rigtige ækvivalent findes i onlineudgaven af B&H: (træet) *pommier à fleurs*, *pommier d'ornement* (*Malus baccata*); (frugten) *fruit du pommier à fleurs*, *fruit du pommier d'ornement*.

5. Inspirationskilder

Vi har anført ovenfor, at B&T, og derefter B&H leksikografisk set er fornyende i forhold til forgængerne. Oplysningsmæssigt har

ordbøgerne trukket på dedikerede redaktører og informanter. I den aktuelle revisionsfase sørger vi for at revidere løbende begge sider, både F-D og D-F for at sikre konsekvens. Dette kan gøres så meget lettere, som det nu drejer sig om én elektronisk base.

Vedrørende D-F reviderer vi ud fra de eksisterende ordbøger over dansk. Vi konsulterer ODS for ældre ordforråd for at sikre os at oplysningerne er korrekte og at ord vi ikke selv kender, faktisk eksisterer eller har eksisteret. Når det gælder moderne dansk, konsulterer vi DDO og Dansk Sprognævn via sproget.dk, og vi konsulterer internettet (Google) og diverse elektroniske databaser (se referencer i litteraturlisten). I kommentarfeltet angiver vi kilder og eksempler. Tilsvarende angående F-D konsulterer vi *Petit Robert's* netudgave, *Trésor de la langue française, Dictionnaire de l'Académie française* og internettet, hvilket ligeledes angives i kommentarfeltet.

6. Konklusion

B&T og senere B&H er store ordbøger, der har været fornyende i forhold til de forudgående ordbøger og inspirerende for den senere leksikografi i Danmark. Ordbøgerne er under stadig revision, dels for at rette op på skævheder og fejl, dels på grund af nødvendige opdateringer. B&H var ikke planlagt som en emnemæssigt afbalanceret, korpusbaseret ordbog, selvom den har en imponerende emnemæssig og stilmæssig bredde. Bestyrelsen for B&H satser på ansvarsfuldt at forvalte den rige arv fra generationer af utroligt flittige redaktører og tilknyttede interesserede seddelskrivende bi-dragydere.

Ligesom for de andre store, videnskabeligt baserede ordbøger er det relevant at stille spørgsmålet om B&H's fremtid og finansiering. Bestyrelsen har besluttet, at en tilpasning til de nye brugeres behov med bl.a. nye krav til formatet er en nødvendighed, og vi har derfor skiftet fra det trykte til det elektroniske medium med

de problemer og de fordele, der er beskrevet ovenfor. Derimod har vi ikke ønsket at gå på kompromis med kvalitet og med fuldstændighed som ideal. Et ideal der kræver fortsat opdatering og bearbejdelse af den elektroniske base. Så længe vedligeholdelsen af B&H foregår uden ekstern finansiering, er det svært, for ikke at sige umuligt, at nå de høje idealer. Det er ønskeligt, at staten eller andre ansvarlige sponsorer forstår, at store videnskabeligt funderede ordbøger, ordbøger over modersmålet såvel som tosprogsordbøger, er nødvendige, ikke mindst i små sprogområder som de nordiske, og at der derfor bør afsættes midler til udvikling og vedligeholdelse af de store ordbøger.

Litteratur

Trykte ordbøger

- B&T = Andreas Blinkenberg & Margrethe Thiele (1937): *Dansk-fransk ordbog*. København: H. Hagerup.
- Blinkenberg, Andreas & Poul Høybye (1953): *Dansk-fransk Suppleringsordbog, Tillæg*. København: H. Hagerup.
- Blinkenberg, Andreas & Poul Høybye (1969): *Dansk-fransk suppleringsordbog, Tillæg, 2. forøgede udgave*. København: H. Hagerup.
- Blinkenberg, Andreas & Poul Høybye (1975-76): *Dansk-fransk ordbog I-II, 3. reviderede og forøgede udgave*. København: Nyt Nordisk Forlag Arnold Busck.
- B&H = forkortelsen dækker principielt over såvel den dansk-franske som over den fransk-danske ordbog fra 1997, se reference nedenfor. Imidlertid refereres der fortrinsvist til den dansk-franske udgave i artiklen: Andreas Blinkenberg & Poul Høybye (1991): *Dansk-fransk ordbog I-II, 4. reviderede udgave*. Ved Jens Rasmussen et al. København: Nyt Nordisk Forlag Arnold Busck.

- Blinkenberg, Andreas & Poul Høybye (1964-1966): *Fransk-dansk ordbog I-II*. København: H. Hagerup.
- Blinkenberg, Andreas & Poul Høybye (1984): *Fransk-dansk ordbog I-II. 2. reviderede udgave*. København: Nyt Nordisk Forlag Arnold Busck.
- Blinkenberg, Andreas & Poul Høybye (1997): *Fransk-dansk ordbog. 3. reviderede udgave*. Ved Jens Rasmussen et al. København: Nyt Nordisk Forlag Arnold Busck.
- DDO = *Den Danske Ordbog* (2003-). Det Danske Sprog- og Litteraturselskab. København: Gyldendal.
- Hansen, Else Juul et al. (2001): *Dansk-Fransk Ordbog*. København: Gyldendal.
- ODS = *Ordbog over det danske Sprog* (1918-2005) Det Danske Sprog- og Litteraturselskab. København: Gyldendal.
- Rey-Debove, Josette & Alain Rey (2012): *Le Petit Robert*. Paris: Dictionnaires Le Robert.
- S&B = Thor Sundby & Jean Guillaume Euchaire Barué (1883-84): *Dansk-Norsk-Fransk Haand-Ordbog 1-2*. København: Gyldendal.

Elektroniske ordbøger og kilder

- Blinkenberg & Høybye dansk-fransk, fransk-dansk ordbog <www.ordbogen.com> (december 2013).
- Gyldendals netudgave (2010) <<http://gyldendal.dk/digitale-boger/ordboeger>> (netudgaven er let afvigende fra den trykte udgave 2001) (december 2013).
- <<http://ordnet.dk>> (med ODS, DDO, korpus.dk, ved Det Danske Sprog- og Litteraturselskab) (december 2013).
- <<http://sproget.dk>> (sproget.dk redigeres af Dansk Sprognævn og Det Danske Sprog- og Litteraturselskab) (december 2013).
- Le Petit Robert Electronique* <<http://www.lerobert.com>> (december 2013).

Trésor de la langue française informatisé <<http://atilf.atilf.fr>> (december 2013).

Dictionnaire de l'Académie française <<http://www.academie-francaise.fr/le-dictionnaire/consultation-en-ligne>> (december 2013).

Anden litteratur

Dyrberg, Gunhild & Lilian Stage (2005): Kan man bruge en ord-bogsartikel uden hoved og hale? Overvejelser over en ny artikelstruktur, I: *Sprint*, 43-61.

Lene Schøsler
professor emeritus, dr.phil.
Københavns Universitet
Njalsgade 128-30
DK-2300 København S
schoesl@hum.ku.dk

Behovet for store ordbøker i en samlet norsk ordbokpolitikk

Arnfinn Muruvik Vonen¹

In this article, the concept of ‘large scientific dictionary’ is approached, and two Norwegian projects are presented and discussed – one almost completed, the other still at the planning stage. All kinds of dictionaries make up an important part of the total written-language infrastructure of a language community. Large scholarly dictionaries represent fundamental documentation of a language and often function as works of source and reference for smaller dictionaries. One current and one planned large scholarly dictionary project of Norwegian are presented and discussed: *Norsk Ordbok 2014* (‘Norwegian Dictionary 2014’), covering Norwegian spoken dialects and the Nynorsk written tradition, and *Bokmålets og riksmålets ordbokverk* (‘The Dictionary of Bokmål and Riksmål’), covering that segment of written Norwegian that is not Nynorsk.

1. Innledning

I denne artikkelen vil jeg drøfte behovet for store vitenskapelige ordbøker, og to norske prosjekter blir omtalt – det ene nesten ferdig, det andre fortsatt på planleggingsstadiet. Alle slags ordbøker utgjør en viktig del av den totale skriftspråklige infrastrukturen i et språksamfunn. Store vitenskapelige ordbøker representerer grunnleggende dokumentasjon av et språk og fungerer ofte som kilde- og henvisningsverker for mindre ordbøker.

¹ Takk til Marit Hovdenak og andre kolleger i Språkrådet for tips til denne artikkelen. Ansvar for det som står her, er selvsagt likevel mitt alene.

2. Ordbøker og en samlet ordbokpolitikk

Det overordnede styringsdokumentet i norsk språkpolitikk er KUD (2008), ofte omtalt som «språkmeldingen», som Stortinget sluttet seg til i 2009. I dette dokumentet kommer det tydelig fram at ordbøker er viktige for et språksamfunn:

Alle former for ordbøker og leksikalske oppslagsverk utgjør ein viktig del av den totale skriftspråkelege infrastrukturen i eit språksamfunn. Alt dette er på sitt vis med på å styrkja skriftkulturen spesielt og nasjonalspråket generelt. (KUD, kap. 8.3.4.1)

Med «skriftspråklig infrastruktur» kan vi forstå ressurser som er utviklet og direkte eller indirekte gjort tilgjengelig for brukerne av skriftspråket, og som utgjør forutsetninger for en mest mulig effektiv og god bruk av det.

Språkmeldingen (kap. 8.3.4.1) trekker et overordnet, men ikke skarpt, skille mellom *ordbøker* og *leksikon* (leksikalske oppslagsverk). Mens ordboken «inneheld meir eller mindre omfattande kunnskap om orda sjøve», er leksikonet «ei kjelde til kunnskap om dei tinga, fenomena eller saksforholda som orda refererer til».

Ordbøker deles i språkmeldingen inn i *enspråklige* og *to- eller flerspråklige* (KUD, kap. 8.3.4.3), og enspråklige ordbøker deles videre inn i enspråklige *spesialordbøker* og enspråklige *allmennordbøker*. Av enspråklige allmennordbøker identifiserer KUD (kap. 8.3.4.4 og 8.3.4.5) tre kategorier: *rettskrivningsordbøker* (skoleordbøker), *ettbinds definisjonsordbøker* (også kalt håndordbøker) og *store, vitenskapelige ordbøker* (også kalt dokumentasjonsordbøker).

Ordbøker trengs for en rekke ulike formål, og i en samlet ordbokpolitikk bør alle disse ordbøkkategoriene inngå. For eksempel er tospråklige ordbøker nær uunnværlige for den som skal lære

et språk, og for den som skal oversette en tekst fra ett språk til et annet. På samfunnsnivå er de av stor betydning for å stimulere befolkningen til å lære flere språk. Enspråklige ordbøker av ulike slag er nyttige for å lære rettskrivning, for å lære terminologien i et fag, eller mer generelt for å bli bedre kjent med ordforrådet i et språk. I mangfoldet av ordbokbehov spiller behovet for de store, vitenskapelige ordbøkene en spesielt grunnleggende rolle, som vi skal se nedenfor. Generelt kan man si at ordbøker i alle skriftkulturer er en forutsetning for språklig demokratisk deltakelse. I gode ordbøker kan alle finne igjen sitt eget språk, og med ordbøker kan alle få hjelp til å bli enda bedre og tryggere språkbrukere.

Å produsere ordbøker, især store ordbøker, er meget ressurskrevende, og i Norge er det usikkerhet om finansieringssituasjonen på feltet i årene framover. Ikke minst har universitetene de senere årene valgt å nedprioritere det leksikografiske feltet spesielt og språklig dokumentasjonsarbeid generelt i sin rekrutteringspolitikk. Språkrådet har derfor det siste året tatt initiativ til en samlet ordbokpolitikk, der behovet for både store vitenskapelige ordbøker og andre ordbokkategorier blir ivarettatt. Våren 2013 arrangerte vi et seminar der ulike aktører ga uttrykk for sine perspektiver. I august 2013 sendte vi en rapport fra seminaret til Kulturdepartementet sammen med et brev med anbefalinger (Språkrådet 2013a, 2013b).

I rapporten kommer det blant annet fram at det offentlige er finansieringskilde for mye av ordbokarbeidet i de andre nordiske landene, som vi ofte sammenligner oss med (Hannesdóttir 2013). Dette gjelder de rettskrivningsordbøkene der den offisielle skriftnormen formidles, det gjelder oppbygging av korpus, digitalisering av gamle seddelarkiv etc., og det gjelder støtte til terminologiarbeid og dialektordbøker. Spesielt i Sverige og Finland gjelder dette også eldre og nyere minoritetsspråk. Det er imidlertid store forskjeller mellom landene hva angår finansiering av de store nasjonale ordbokverkene. I Sverige er det således en frittstående

institusjon, Svenska Akademien, som helt siden den fikk sitt kongelige oppdrag i 1786, har hatt ansvaret for både dokumenterende og normformidlende enspråklig leksikografi. Norge har ingen ikke-statlig institusjon med tilsvarende historiske røtter og tilsvarende finansiell stabilitet.

Bidragstyperne til rapporten (Språkrådet 2013b) representerer ulike aktører og interesser, og rapporten gjenspeiler disse ulikhetene både gjennom ulik vektlegging av offentlige institusjoners og private forlags rolle i ordbokarbeidet og gjennom ulike leksikografiske tradisjoner.

Etter tilbakemeldinger fra Kulturdepartementet sendte vi i mars 2014 et oppfølgende brev og notat (Språkrådet 2014a, 2014b), som for tiden (juli 2014) er til behandling i departementet. I brevet og notatet anbefaler Språkrådet at det blir opprettet et offentlig finansiert senter for vitenskapelig dokumentasjon av de språkene Norge har ansvar for, for vedlikehold og utvikling av norske språksamlinger fra norrønt til moderne norsk, og for bevaring og bruk av dokumenterende språksamlinger. Av hensyn til den faglige kvaliteten og rekrutteringen av medarbeidere anbefaler Språkrådet at universitets- og høyskolesektoren vurderes som en mulig tilknytning for senteret, mens en tilknytning til en virksomhet under Kulturdepartementet kan være et alternativ.

I det offentlige ordskiftet etter Språkrådets brev og notat er det kommet fram et interessant allment universitetspolitisk spørsmål om hvorvidt tidligere tildelte samfunnsoppdrag – som at Universitetet i Oslo har fått tildelt ressurser til å ivareta norsk leksikografi fra 1970-tallet av – skal være bindende for en akademisk institusjon, eller om institusjonen selv skal kunne omdefinere slike oppdrag (se for eksempel Toft 2014).

3. Store ordbøker

Det som skiller de store vitenskapelige ordbøkene fra de andre ordbokkategoriene, er at de har en grunnleggende dokumentasjonsfunksjon. Selv om også rettskrivningsordbøkene og de enkelte definisjonsordbøkene til en viss grad har en slik funksjon, er det de store, vitenskapelige ordbøkene som her er grunnleggende, og som ofte fungerer som kilde- og referanseverk for de mindre ordbøkene. KUD (kap. 8.3.4.5) understreker:

Dokumentasjonsordbøkene er ikkje primært rettskrivningsordbøker, men har først og fremst til formål å forklara og eksemplifisera kva orda tyder, og korleis dei blir brukte.

Dokumentasjonsordbøkene er altså dermed et særlig viktig og allment tilgjengelig resultat av språklig dokumentasjonsarbeid, og 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. (KUD, kap. 8.3.4.5)

Et viktig stikkord her er *vitenskapelig*. Det kan stilles spørsmål om hva som kjennetegner et vitenskapelig ordbokarbeid. I en utredning fra Det historisk-filosofiske fakultet ved Universitetet i Oslo (UiO 1989) argumenteres det for at kriteriene for vitenskapelighet ikke er grunnleggende annerledes i ordbokarbeid enn i annet faglig arbeid. Dette utvalgets mandat var

å vurdere om og eventuelt under hvilke forutsetninger ordbøker, feltarbeidsrapporter, registre, arkivregistranter, tekstutgaver, kataloger og annen relevant aktivitet ved institutter og samlinger, som ikke blir publisert som vitenskapelige artikler eller avhandlinger, er å betrakte som vitenskapelig meritterende.

I utvalgets konklusjon heter det:

Utvalget mener at de arbeidsoppgaver som nevnes i mandatet, er vitenskapelig meritterende dersom de oppfyller de samme kriterier som vanligvis stilles til forskning og vitenskap. Således vil deler av ordboks- og arkivarbeid kunne klassifiseres som grunnforskning, anvendt forskning eller utviklingsarbeid. Andre oppgaver vil være av mer rutinemessig karakter som riktignok krever vitenskapelig kompetanse, men som ikke nødvendigvis gir ny viten eller innsikt.

Utredningen knytter vurderingen av vitenskapelighet til «vanlig aksepterte trinn i en forskningsprosess»:

- målsetting og problemstilling
- teori- og begrepsdiskusjon
- metodebruk og modellutvikling
- datainnsamling, databehandling og datasystematisering
- hypoteseprøving og tolking
- konklusjon og resultater

I et stort ordbokprosjekt er det viktig med et utredningsarbeid som legger det teoretiske og metodiske grunnlaget for prosjektet. Metoder og problemstillinger bør vurderes underveis og om nødvendig endres når ny innsikt tilsier det. Innsamling, behandling og

systematisering av data er en krevende virksomhet som forutsetter vitenskapelig kompetanse. Erfaring fra slikt arbeid og fortrolighet med et omfattende kildemateriale kan gi gode innfallsvinkler til ulike forskningsfelt og faglige problemstillinger. Det å knytte krav til vitenskapelighet til et leksikografisk arbeid kan sees som å kreve at det skal være høy kvalitet i alle disse arbeidsprosessene.

4. To prosjekter

Ordbokhistorien for det norske språket er rimelig nok preget av de ulike tradisjonene for utvikling og normering av norsk. Tradisjonene har sine røtter i 1800-tallets ulike responser på spørsmålet om hva slags språk som burde erstatte det danske, som hadde vært offisielt språk i Norge i lang tid. Denne artikkelen gir ikke rom for noen nærmere analyse av denne historien, og her vil jeg bare påpeke at de to prosjektene som skal omtales her, er forankret i hver sin av disse tradisjonene, ofte kalt henholdsvis *nynorsktradisjonen* og *bokmålstradisjonen*. Det ene bygger på det arbeidet som Ivar Aasen begynte på 1800-tallet (se Aasen 1873), og andre videreførte, med å dokumentere de norske dialektene og bygge og videreutvikle et nytt norsk skriftspråk på basis av dem og det norrøne språkhistoriske utgangspunktet. Det andre bygger på den litterære tradisjonen som har et dansk utgangspunkt og med tiden har utviklet en egen språklig identitet nærmere norsk talespråk.

Norsk Ordbok 2014 (NO2014) har som mål å produsere *Norsk Ordbok*, en ordbok som dokumenterer ordforrådet i norske dialekter og nynorsk skriftspråk. Arbeidet med *Norsk Ordbok* tok til på slutten av 1920-tallet, og det redaksjonelle arbeidet med siste bind planlegges slutført i 2014. Verket vil da bestå av tolv trykte bind med ca. 300 000 ordartikler. Arbeidet med verket tar utgangspunkt i et omfattende innsamlet materiale, blant annet et nynorsk-korpus som nå er på vel 100 millioner ord, og et seddelarkiv med

3,2 millioner sedler. Selve prosjektet NO2014 ble organisert i 2002 i et samarbeid mellom Universitetet i Oslo og Kulturdepartementet. Bakken (2005) beskriver hvordan prosjektorganiseringen har virket inn på arbeidet med ordboken, og nevner at det at tilfanget er tilgjengelig på Internett, gjør ordboken etterprøvable i vitenskapelig forstand. Ore & Wetås (2014) beskriver hvordan prosjektet er blitt digitalisert, og hvordan digitaliseringen på ulike måter har gjort produktet til en ressurs med et stort flerbrukspotensial.

Det planlagte prosjektet Bokmålets og riksmålets ordbokverk (BRO) har som mål å dokumentere ordforrådet i den delen av skriftlig norsk som ikke er nynorsk. Prosjektet er planlagt som et samarbeid mellom Det Norske Akademi for Språk og Litteratur og Universitetet i Oslo. En viktig komponent i verket vil være *Det Norske Akademis Store Ordbok* (NAOB), en revidert utgave av *Norsk Riksmålsordbok*, som kom ut i årene 1937–1995. NAOB skal kun utgis digitalt og forventes ferdig ca. 2017 med rundt 250 000 artikler. Korpusgrunnlaget for BRO vil ta utgangspunkt i det eksisterende Leksikografisk bokmålskorpus ved Universitetet i Oslo, som består av ca. 100 millioner ord i tekster fra 1985 til i dag. For mer om revisjonen av *Norsk Riksmålsordbok* og NAOB-prosjektet, se Nilstun (2012) og Henriksen (2012).

5. Noen utfordringer

De to store prosjektene som er trukket fram her, er meget ulike og er i meget forskjellige faser. NO2014 er for tiden inne i sin slutfase. Prosjektet avsluttes tidlig i 2015, og det foreligger ingen vedtatte planer for faglig og teknisk vedlikehold. Siden prosjektet har pågått over så lang tid, blir det endelige resultatet preget av at de tekniske mulighetene er helt andre nå i produksjonen av de siste bindene enn de var i produksjonen av de første. De fem første bindene av *Norsk Ordbok* er ikke søkbare i digital form, og det vil være

en ressurskrevende prosess å gjøre hele verket tilgjengelig gjennom digitale medier. En slik fulldigitalisering av verket inngikk heller ikke i prosjektplanen fra 2002. I lys av den digitale utviklingen som har vært i samfunnet siden den gangen, må det likevel betraktes som en hemsko for utbredelsen og bruken av verket at bare sju av tolv bind vil være digitalt søkbare når verket står ferdig.

Når det gjelder BRO-prosjektet, er ikke alle detaljer i arbeidsmåter etc. gjort eksplisitte ennå. Jeg imøteser en mer eksplisitt plan for dette prosjektet og regner med at utviklingen av de nye delene i bokmålsordbokverket vil skje i samsvar med vitenskapelige prinsipper. Jeg legger også til grunn at erfaringene fra NO2014 kan være særlig relevante å trekke inn også for BRO når det gjelder vitenskapelighet, ikke minst erfaringene med strenge krav til empirisk belegg og etterprøvbarehet.

NO2014 og BRO, de to prosjektene som til sammen har ambisjoner om å dokumentere ordforrådet i norsk, både skriftlig og muntlig, er altså på mange måter forskjellige fra hverandre og står overfor nokså forskjellige utfordringer. Samtidig er det enighet om at det trengs grundig og tilgjengelig dokumentasjon av alle former for skriftlig og muntlig norsk. Håpet er at erfaringene fra prosjektene i tiden framover kan komme til nytte på tvers av de tradisjonene de representerer.

I Norge er det vanskelig å tenke seg noen annen hovedfinansieringskilde for store ordbokverker enn det offentlige. Det blir dermed et viktig spørsmål om politiske myndigheter i årene framover ønsker å prioritere ordbokarbeid like høyt som i de årene prosjektet NO2014 har løpt. Språkrådet skrev i sitt brev til Kulturdepartementet i 2013:

Dersom statsbudsjettene også etter 2014 legger opp til en samlet bevilgning til ordbokarbeid som er omtrent på samme nivå som dagens, vil vi raskt kunne merke forbedringer i dokumentasjonen av språk i Norge. (Språkrådet 2013a)

Som nevnt ovenfor er det også et uavklart politisk spørsmål hva slags institusjoner som bør drive vitenskapelig ordbokarbeid – universitets- og høyskolesektoren, kultursektoren eller en helt annen del av samfunnslivet. Det haster å få avklart disse viktige spørsmålene og kommet videre i dokumentasjonen av norsk språk og språk i Norge.

Litteratur

- Bakken, Kristin (2005): Norsk Ordbok 2014 – rammer og utfordringer. I: Ruth Vatvedt Fjeld & Dagfinn Worren (red.): *Nordiske studier i leksikografi 8. Rapport frå Konferanse om leksikografi i Norden. Volda 20.–24. mai 2003*. Oslo, 29–35.
- Hannedóttir, Anna Helga (2013): Ordbokspolitik i Sverige, Danmark, Island och Finland – en översikt. I: Språkrådet (2013b), 13–17.
- Henriksen, Petter (2012): Massive ordbokspalter i transparent presentasjon – NAOB på nett: samspillet mellom datastruktur, innhold og brukergrensesnitt. I: Birgit Eaker, Lennart Larsson & Anki Mattisson (red.): *Nordiska studier i leksikografi 11. Rapport från Konferensen om leksikografi i Norden. Lund 24–27 maj 2011*. Lund, 251–262.
- KUD (2008) = *St.meld. nr. 35 (2007–2008). Mål og mening. Ein heilskapleg norsk språkpolitikk*. Kulturdepartementet. <<http://bit.ly/UAGxaV>> (juni 2014).
- Nilstun, Carina (2012): Fra dype betydningshierarkier til flattere struktur – et innblikk i revisjonen av Norsk Riksmålsordbok. I: *LexicoNordica* 19, 123–140.
- Norsk Ordbok* (1966–) = *Norsk Ordbok. Ordbok over det norske folkemålet og det nynorske skriftmålet*. Oslo: Det Norske Samlaget. Nettutgave: <no2014.uio.no> (juni 2014).

- Norsk Riksmålsordbok* (1937–1995). 6 bind. Oslo: Riksmålsvernet / Det Norske Akademi for Sprog og Litteratur. (Under revisjon under navnet *Det Norske Akademis Store Ordbok*, planlegges utgitt digitalt i 2017 av Det Norske Akademi for Språk og Litteratur i samarbeid med Kunnskapsforlaget.)
- Ore, Christian-Emil & Åse Wetås (2014): *Norsk Ordbok* i den digitale tidsalderen. I: *LexicoNordica* 21 (dette nummeret).
- Språkrådet (2013a) = *En samlet ordbokpolitikk etter 2014*. (Brev fra Språkrådet til Kulturdepartementet.) <<http://bit.ly/iq45u0>> (juni 2014).
- Språkrådet (2013b) = *En samlet ordbokpolitikk etter 2014. Rapport fra et seminar 20. mars 2013*. <<http://bit.ly/1pqJKEt>> (juni 2014).
- Språkrådet (2014a) = *Norsk ordbokpolitikk*. (Brev fra Språkrådet til Kulturdepartementet.) <<http://bit.ly/1lXAKTs>> (juni 2014).
- Språkrådet (2014b) = *Norsk ordbokpolitikk. Notat fra Språkrådet til Kulturdepartementet 3. mars 2014*. <<http://bit.ly/1qnWmLK>> (juni 2014).
- Toft, Martin (2014): Kulturdepartementet krev klare svar frå UiO om språksamlingane. I: *Uniforum* 28.5.2014. <<http://bit.ly/1l3P-KxD>> (juni 2014).
- UiO (1989) = *Vurdering av ordboks- og arkivarbeid i forhold til vitenskapelige kvalifikasjoner*. (Utrykt innstilling.) Utvalget for forskningsspørsmål ved Det historisk-filosofiske fakultet, Universitetet i Oslo.
- Aasen, Ivar (1873): *Norsk Ordbog med dansk Forklaring*. Christiania: P.T. Mallings Boghandel.

Arnfinn Muruvik Vonen
 direktør, professor dr.art.
 Språkrådet
 Postboks 8107 Dep
 NO-0032 OSLO
 amv@sprakradet.no

ICKE-TEMATISKA BIDRAG

Kommentarer til Sven Tarp: Retskrivningsordbog på godt og ondt

*Margrethe Heidemann Andersen, Jørgen Nørby Jensen &
Anita Ågerup Jervelund*

This article is a comment on some of the criticisms which Sven Tarp puts forward in his article *Retskrivningsordbog på godt og ondt* in *LexicoNordica* 20. Sven Tarp's criticism has been accepted on a number of points, for example, that it would have been more practical if a user of the digital version of *Retskrivningsordbogen* (the official dictionary of Danish Orthography) could be redirected from the archaic spelling to the new and correct spelling. With regard to other criticisms, however, we are somewhat more hesitant. Thus, we are looking for clear guidelines on how "the lexicographical principles on proscription" are to be realized efficiently in dictionaries.

1. Indledning

I sin artikel *Retskrivningsordbog på godt og ondt* i *LexicoNordica* 20 (Tarp 2013a) kritiserer Sven Tarp den seneste udgave af *Retskrivningsordbogen* på en række punkter. Eksempelvis mener Tarp at der er for få opslagsord i ordbogen, at det ligger udenfor Dansk Sprognævn's opdrag at angive ordforbindelser og betydningsoplysninger, og at Sprognævnet bør bruge "det leksikografiske princip om proskription" i udarbejdelsen af *Retskrivningsordbogen*.

I denne artikel kommenterer vi en række af Sven Tarps kritikpunkter.

2. Ordudvalget i Retskrivningsordbogen

Sven Tarps første kritikpunkt går på antallet af opslagsord i Retskrivningsordbogen. Efter Tarps mening er der slet og ret for få. Kritikken genfindes i Theilgaard (2013) og i Trap-Jensen (2013).

Retskrivningsordbogen indeholder i alt ca. 64.000 opslagsord samt ca. 10.000 sammensætningseksempler (fx *sødmælkskalv* under opslagsordet *sødmælk*). Dette ordstof skal ifølge ordbogen dække "rigssprogets almindelige ordforråd, herunder en del fagord med en vis almen udbredelse" (RO 2012:11). Man kan selvfølgelig diskutere hvor mange ord der hører med til rigssprogets almindelige ordforråd, og nogen klar grænse mellem almensprog på den ene side og fagsprog, gruppesprog, slang m.m. på den anden er det næppe muligt at trække. Så længe Retskrivningsordbogen kun forelå i en trykt udgave, var der en naturlig grænse for hvor mange ord den kunne indeholde, men denne begrænsning er der naturligvis ikke i den digitale verden. Sven Tarp – og andre – kan således glæde sig over at den elektroniske udgave af Retskrivningsordbogen fremover løbende vil blive forsynet med nye opslagsord (se Jervelund 2014; jf. også afsnit 9).

Lige meget hvor mange ord man fylder i ordbogen, vil man dog aldrig kunne dække hele ordforrådet. I dansk kan der dannes et utal af sammensætninger, og kun et fåtal af dem kan naturligvis være med i en ordbog, det være sig en trykt eller en elektronisk. Når Sven Tarp (2013a:187) hævder at "en underliggende database kan indeholde hele det danske ordforråd ud i dets yderste afkroge", er der således tale om en sandhed med endog meget store modifikationer. Formålet med Retskrivningsordbogen er da heller ikke at dokumentere alle eksisterende ord i dansk (det kan ingen ordbog gøre), men at oplyse om den officielle retskrivning til dem der skriver eller retter en tekst. Og disse oplysninger skulle brugerne gerne kunne finde frem til vha. ordudvalget og retskrivningsreglerne i ordbogen. Når Sven Tarp ydermere påstår at der udenfor

Retskrivningsordbogens opslagsord råder “nærmest Klondykelignende tilstande med hensyn til retskrivning” (Tarp 2013a:186), er det heller ikke rigtigt. At ord som fx *computerskærm*, *grillspyd* og *radioforhandler* ikke står i Retskrivningsordbogen, betyder ikke at man kan skrive dem som man vil. De skal skrives på netop den anførte måde. Eller som Erik Hansen udtrykker det: “Er man i tvivl om hvordan *villahave* staves, kan man slå *villa* og *have* op og så skrive dem sammen til ét ord” (Hansen 2000:12).

3. Sammensætningsoplysninger

I RO 2012 er der indført sammensætningsoplysninger ved langt flere opslagsord end tidligere. Sådanne oplysninger er ifølge Tarp “guld værd, fordi dansk er kendetegnet ved mange komposita, hvor netop sammenkædningen ofte volder problemer” (Tarp 2013a:194). Derimod er det ifølge Tarp et problem at sammensætningseksempler som *strandstol* og *strandplante* (under opslagsordet *strand*) ikke også er at finde på alfabetisk plads: “En bruger, som har retskrivningsproblemer i forhold til et af disse ord, vil derfor ikke umiddelbart kunne finde dem i den trykte udgave”. Hertil kan siges at når man under opslagsordet *strand* i Retskrivningsordbogen får at vide at det hedder *strand-* i sammensætninger, så gælder det for *alle* de sammensætninger med *strand* som førsteled som man overhovedet kan finde på. Hvis det blot gælder om at have så mange sammensætninger med som muligt, kan man da godt – som Den Danske Ordbog – nævne langt over 30 opslagsord med *strand* som førsteled (Retskrivningsordbogen har 18). Men brugerne vil jo stadig kunne søge forgæves efter lige netop deres “strandord”. Hverken Retskrivningsordbogen, Den Danske Ordbog, Den Danske Netordbog, Nudansk Ordbog eller nogen andre ordbøger nævner fx sammensætningerne *strandsild*, *strandsex* og *strandhygge*, til trods for at ingen af ordene kan siges at være

ualmindelige. En Googlesøgning 22.4.2014 gav således flere tusinde belæg. Den slags ord danner man når man har brug for det, og i Retskrivningsordbogen finder man hjælp til orddannelsen.

Ingen ordbøger er uden huller, heller ikke Retskrivningsordbogen. Men at ordbogen skulle være præget af en “usystematisk lem-maselektion” (Tarp 2013a:187), er efter vores mening ikke korrekt.

4. Ordforbindelser

Om ordforbindelser i Retskrivningsordbogen skriver Tarp at de kan være “uhyre nyttige”, men at det “ligger uden for nævnets opdrag” at angive sådanne (Tarp 2013a:193). Hvorfor det skulle være forbudt for Sprognævnet at angive ordforbindelser under forskellige opslagsord, giver Sven Tarp imidlertid ingen forklaring på. Samme synspunkt fremfører Tarp (2002) om Retskrivningsordbogen fra 2001 hvor han spørger om “Retskrivningsordbogen nu har vundet hævde for, at kollokationer er en del af retskrivningen” (Tarp 2002:201). Og i Tarp (2013b) hedder det videre om kollokationer at de vel næppe direkte har “noget med retskrivningen at gøre, da de enkelte ord kan findes retskrevet andre steder i ordbogen”. Som påpeget af Christian Becker-Christensen (2013:21) synes Tarp at tro at “eksempler er med for deres egen skyld ligesom i en receptions- eller tekstproduktionsordbog og ikke som ordidentifikatorer i en retskrivningsordbog”. Ordforbindelsen *virre med hovedet* er fx kun med for at identificere opslagsordet *virre*. Brugeren skal med andre ord kunne se at han har slået op på det rigtige ord. Så når Sven Tarp (2002:206) gerne vil vide om det hedder *der er løbet meget vand i havet/i stranden/på stranden siden sidst*, må svaret være at en sådan oplysning ikke hører hjemme i Retskrivningsordbogen, men i stedet må søges andre steder, fx i de gængse betydnings- og konstruktionsordbøger.

5. Bøjningsangivelser

Udsagnsord er efter Tarps mening “problembarnet over alle problembørn” (Tarp 2013a:193). Især savner han i Retskrivningsordbogen en eksplicit angivelse af datid passiv og bydemåde. Kritikken kunne rettes mod alle andre dansksprogede ordbøger, for der er ikke tradition for at angive sådanne bøjningsformer på alfabetisk plads. Hvis man fx gerne vil vide hvad *cykle*, *sagtne* og *pierce* hedder i bydeform, leder man således også forgæves i både Nudansk Ordbog, Den Danske Ordbog og Den Danske Netordbog. De korrekte former (*cykl*, *sagtn* og *pierce*) fremgår af § 30 i Retskrivningsordbogen, og der er ingen der har hindret fx Den Danske Netordbog i at videreformidle dem. Bortset fra det kunne det være udmærket at nævne bydemåde på alfabetisk plads i netudgaven af Retskrivningsordbogen, og det er der da også overvejelser om at gøre.

Mere tvivlsomt er det om angivelse af datid passiv hører hjemme i en retskrivningsordbog, for valget mellem fx *begikkes* og *blev begået* kan næppe siges at være et retskrivningsspørgsmål. Former som *begikkes*, *vandtes* og *genoptoges* kan naturligvis dannes, men de fleste nutidige sprogbrugere foretrækker formuleringer som *mordet blev begået i tirsdags*, *kampen blev vundet af FCK* og *mødet blev genoptaget kl. 20* fremfor *mordet begikkes ...*, *kampen vandtes ...* og *mødet genoptoges ...*

Når formen *begås* (nutid passiv) nævnes på alfabetisk plads, er det for at vise at stavemåden ikke er *begåes* (se i øvrigt RO 2012:19).

6. Dobbeltformer

Ifølge Tarp forsvinder dobbeltformer ofte ud af Retskrivningsordbogen inden ændringen er slået igennem i befolkningen, eksempelvis “hvis en ny form er blevet introduceret i foregående oplag,

og de to former kun har rivaliseret i denne korte periode” (Tarp 2013a:189). Her må det præciseres at nye former aldrig bliver introduceret i nye oplag, men kun i nye udgaver. Dertil kommer at det kun er relativt sjældent at staveformer og ordformer ændres uden en vis overgangsperiode, der i øvrigt gerne strækker sig over en temmelig lang årrække, i reglen mindst 10 år. I RO 2012 drejer det sig primært om en række konsekvensændringer, fx når opslagsordet *e-mail-adresse* ændres til *e-mailadresse* som følge af at reglen om bindestreg i sammensætninger med forkortelser ændres (§ 57.3), eller når *T-bone-steak* ændres til *T-bonesteak*. Dertil kommer opslagsordet *accessoires*, der nu er ændret til *accessoires*, og *columnist*, der nu skal skrives *kolumnist*. Det førstnævnte eksempel er ændret således at stavemåden er bragt i overensstemmelse med stavemåden i det långivende sprog, og det sidstnævnte eksempel er ændret af hensyn til stavemåden i ordene *kolumne* og *kolumnetitel*. Ændringerne fra *spin-off* til *spinoff* og fra *tagselv-bord* til *tag selv-bord*, der også nævnes af Tarp, er først og fremmest indført af systematiske hensyn. Alle disse ændringer og argumenterne bag dem er i øvrigt udførligt beskrevet i Nørby Jensen (2013). En grundigere diskussion af skrivemåden af engelske låneord på *-in*, *-off*, *-on*, *-out* og *-up* findes i Nørby Jensen (2006).

Tarp nævner også ændringerne af *remiks* til *remix* og *nobelpris* til *Nobelpris*. Om ændringen af *remiks* til *remix* kan det siges at vi her har afskaffet en form hvor hverken stavemåden eller den dertilhørende danske udtale viste sig at være i tråd med den faktiske sprogbrug, og hvor det derfor ikke gav mening at beholde formen *remiks*. Endelig er der faktisk slet ikke sket nogen ændring i stavemåden af *Nobelpris*, som stadigvæk også må skrives *nobelpris* (jf. Retskrivningsordbogens § 12.10.b). Vi har dog ændret stavemåden af *Nobelpris* på alfabetisk plads da vi i Retskrivningsordbogens § 12.13.d skriver at i de tilfælde hvor der indgår et proprium i betegnelser for “fortjensttegn, ordenstegn, priser og udmærkelser”, er det som regel mest nærliggende at skrive betegnelsen med stort,

altså *Nobelpris*. Det er derfor mest naturligt at denne skrivemåde også findes på alfabetisk plads, men nogen egentlig ændring er der altså ikke tale om. Også priserne *Oscar* og *Bodil* er i øvrigt på samme baggrund ændret på alfabetisk plads fra lille begyndelsesbogstav til stort begyndelsesbogstav.

Konklusionen er således at det kun er en lille gruppe ord der har ændret staveform eller ordform med den nye udgave af Retskrivningsordbogen uden en overgangsperiode med dobbeltformer. Alle retskrivningsændringerne kan findes på de (i øvrigt fyldestgørende) lister over ændrede ord og staveformer der ligger på Sprognævnets hjemmeside, og man vil til hver en tid kunne orientere sig om ændringerne på disse ordlister eller læse om argumenterne bag i artikler i *Nyt fra Sprognævnet*. Når det er sagt, vil vi dog gerne give Sven Tarp ret i at det ville have været hensigtsmæssigt hvis der fx i forordet til den trykte udgave havde været en henvisning til listerne på Sprognævnets hjemmeside (Tarp 2013b:129). På samme måde ville det også have været til gavn for sprogbrugerne hvis de i den digitale udgave af Retskrivningsordbogen kunne blive henvist fra forældede stavemåder som *majonæse*, *columnist* og *spin-off* til de korrekte stavemåder *mayonnaise*, *kolumnist* og *spinoff*. Begge kritikpunkter har vi taget til os, og vi vil derfor bl.a. arbejde på at gøre søgemulighederne i den digitale version af RO 2012 bedre.

7. Dobbeltformer og anbefalet form

Sven Tarp kritiserer den måde Sprognævnet behandler dobbeltformer på, fordi vi ikke følger “det leksikografiske princip om proskription” og således ikke tager hensyn til brugerne og deres behov. Tarp mener at vi skal anbefale én form og samtidig gøre opmærksom på at der også findes en anden tilladt form (i nogle tilfælde flere tilladte former), således at brugerne ikke skal bruge tid på at vælge mellem former som de måske ikke har nogen forudsæt-

ninger for at tage stilling til (Tarp 2013a:188). Tarp står ikke alene med dette synspunkt. Udover Bergenholtz, som Tarp selv nævner (se også Bergenholtz & Bøgelund 2002), har bl.a. Trap-Jensen (2013:12) givet udtryk for at Sprognævnet bør vægte dobbeltformerne – med SAOL (den svenske pendant til RO) som forbillede.

Også Galberg Jacobsen (1999) har diskuteret og opstillet mulige løsningsmodeller for vægtning af de valgfrie former. Han undlader dog at give retningslinjer eller kriterier for hvordan man skal vælge mellem dobbeltformerne og afgøre hvilke der skal anbefales, og hvilke der blot skal tillades. Han er klar over at det er en alvorlig svaghed ved forslaget, og han erkender at “der vil være meget store problemer med at foretage disse valg på en kvalificeret måde” (Galberg Jacobsen 1999:10).

For mere end 100 år siden og en del år frem fandtes der faktisk anbefalede former i den danske retskrivning, men det var ikke uproblematisk. I retskrivningsbekendtgørelsen fra 1892 var der ved visse dobbeltformer en markering med en stjerne der viste hvilken af formerne ministeriet anbefalede (fx **fejg/fej* og **tidt/tit*). Ofte anbefalede ministeriet de traditionelle former fremfor de nyere former, og det må siges at være et udvælgelseskriterie der lod den officielle retskrivning stå i stampe. Anbefalingsstjernen blev kritiseret af bl.a. Viggo Saaby (se Galberg Jacobsen 2010:138-139), der mente at ministeriet i en del tilfælde anbefalede den forkerte form, og det viste sig da også at denne stjerne gradvis forsvandt igen, men den kunne stadig ses i retskrivningsordbøgerne ved fx *dejj* helt frem til 1946 (Galberg Jacobsen 1999:8).

Selvom man dengang gik bort fra de anbefalede former, er det muligt man burde genindføre en sådan anbefaling hvis det kan gøres bedre i dag. Og flere, heriblandt Tarp, foreslår jo så også at Sprognævnet indfører princippet om proskription af hensyn til brugerne og deres behov. Der er dog ingen der kommer nærmere ind på hvilke retningslinjer og udvælgelseskriterier vi burde følge i dette arbejde i forbindelse med RO. Men der er ordbøger der

praktiserer proskription, bl.a. Den Danske Netordbog (DDNO), som ifølge Tarp har anvendt princippet med gode resultater (Tarp 2013a:188). Det er derfor nærliggende at se nærmere på hvordan det bliver gjort i DDNO.

7.1. Anbefalede former i DDNO

I DDNO er en dobbeltform meget ofte placeret under synonymmer blandt de almindelige synonymmer. Hvis man fx søger på *ressource*, er dobbeltformen *resurse* placeret sammen med synonymmer som *energikilde* og *reserve*. Ordet *resurse* hører ikke til blandt synonymmer, for det er jo blot en anden staveform af ordet *ressource*, ikke et helt andet ord med nogenlunde samme betydning som *ressource*. Under *Feltforklaringer* (punktet *info*) finder man en beskrivelse af synonymmer, men det nævnes ikke at der også kan stå dobbeltformer under synonymmer.

Det er muligt at denne sortering fungerer fint for ordbogsbrugerne, men at det også skulle være en god måde at vise anbefalede former på, er svært at se. Hvis man overhovedet får øje på *resurse* under synonymmerne, ser man ikke andet end at dette ord står på en liste sammen med de almindelige synonymmer. Der står intet om at *resurse* ikke anbefales. Der er heller ingen tydelig anbefaling af *ressource*. Først når man klikker på *resurse* og kommer hen til *resurse* som opslagsord, står der “Denne skrivemåde anbefales ikke, brug i stedet → *ressource*”. I øvrigt fremgår det ikke hvorfor *ressource* anbefales fremfor *resurse*. Det er ikke beskrevet i de to ordbogsartikler, og der er heller ingen generel beskrivelse af anbefalede former under fx punkt *info*.

En anden fremgangsmåde er helt at udelade den ikkeanbefalede form under den anbefalede form. I artiklen *ressourcespild* er dobbeltformen *resursespild* slet ikke nævnt, hverken som synonym eller på anden måde. Man får altså ikke oplyst den alternative skrivemåde i denne artikel, og man opdager den kun hvis man skriver

resursespild i søgefeltet. Så kommer man til artiklen *resursespild* med bemærkningen “Denne skrivemåde anbefales ikke, brug i stedet → *ressourcespild*”. Noget tilsvarende gør sig gældende i artikler som fx *bulldog*, *ion* og *vue*, hvor de alternative former *bulldog*, *jon* og *vy* udelades.

Proskription kan åbenbart betyde at man holder den ikke-anbefalede form skjult i de tilfælde hvor redaktionen mener at den ikke er værd at nævne – til trods for at den er fuldt korrekt ifølge den officielle retskrivning.

Nogle gange er der en kortere eller længere tekst der begrundes valget af den anbefalede form. Det ser ud til at den anbefalede form altid er den der er mest frekvent ifølge redaktionens undersøgelser. Under opslagsordet *højtaler* anbefales formen med ét *t*, *højtaler*, og de to formers frekvens står beskrevet i en meget lang tekst med overskriften *Grammatisk anmærkning*:

Grammatisk anmærkning

I reklametekster eller tekster skrevet af firmaer, der sælger eller producerer apparater, som omdanner elektriske svingninger til lydbølger, bruges ofte skrivemåde med to *t*'er: “højtaler”. Det gør sig især gældende ved alle former for komposita, som bruges mindre i faglige tekster end i almensproget. En Google-søgning den 16.1.2011 gav 2.220.000 med “højtaler” mod 425.000 belæg med “højtaler”, altså en klar overvægt af skrivemåde med et *t*. Vi vælger derfor helt generelt at anbefale denne skrivemåde. Denne anbefaling gælder også for komposita, selvom sprogbru- gen her har en anden tendens. Google-søgningen gav her

- 6.840 belæg med *baghøjtaler*
- 12.600 belæg med *baghøjtaler*
- 518 belæg med *bashøjtaler*
- 6.340 belæg med *bashøjtaler*
- 3.560 belæg med *centerhøjtaler*

- 17.200 belæg med centerhøjtaler
- 428 belæg med diskanthøjtaler
- 2.000 belæg med diskanthøjtaler
- Denne skrivemåde anbefales ikke, brug i stedet → højtaler

Den meget lange tekst indeholder endda fagtermen *komposita*. Den sidste sætning ("Denne skrivemåde anbefales ikke, brug i stedet → højtaler") står efter *Grammatisk anmærkning* og er svær at få øje på hvis man ønsker at springe den lange tekst over og gå direkte til anbefalingen.

Det største problem er dog at frekvensundersøgelserne udelukkende er baseret på Googlesøgninger. DDNO har altså valgt at anbefale formen *højtaler* fordi de ved hjælp af Google har fundet frem til at det er den mest udbredte af de to former. Det er af mange årsager meget farligt at stole så meget på en intelligent søgemaskine som Google. Brugere kan i øvrigt selv lave Googlesøgninger, hvilket de i høj grad også gør. Hvad skal de så med alle de tal over Googlesøgninger i en ordbog?

Den Danske Ordbog på ordnet.dk når i øvrigt frem til det modsatte resultat, nemlig at *højtaler* er den mest almindelige form. Det ses ved at formen *højtaler* står først (*højtaler* eller *højtaler*), men også ved at der længere nede i artiklen står *især i formen højtaler*. En sådan bemærkning står der kun når den ene form er meget mere udbredt end den anden (jf. <<http://ordnet.dk/ddo/artiklernes-opbygning/opslagsord>>).

At både *højtaler* og *højtaler* er anerkendte former i retskrivningen, skyldes andet end frekvens. Formen *højtaler* svarer bedst til udtalen, som er uden stød i *høj-* ligesom fx *højbane*, *højhus* og *højgravid*, og orddannelsen uden *t* på *høj* svarer til andre ord der består af et adjektiv + en betegnelse for en person eller ting som udfører en aktivitet, fx *hurtigløber*, *fritænker* og *rentegner* (Hansen 1997:11-12). Der er altså gode lingvistiske argumenter for at vælge *højtaler*. Ikke desto mindre er der mange der foretrækker for-

men *højtaler*, som har stået i RO siden 1955 og er den traditionelle form. Formen *højtaler* kom først i RO i 1986.

Under opslagsordet *ministerie* skriver DDNO:

Denne skrivemåde anbefales ikke af denne ordbog. Ved en Google-søgning i marts 2011 fandtes i alt 5.240 belæg med “ministerie” mod 320.000 med “ministerium”.

Denne skrivemåde anbefales ikke, brug i stedet → ministerium

Under opslagsordet *ministerium* nævnes formen *ministerie* ikke. Artiklen *ministerium* har altså samme opbygning som førnævnte *ressourcespild*, *bulldog*, *ion* og *vue*, hvor den alternative form skjules, og det gælder også for andre ord på *-ium/-ie*, fx *kollegie/kollegium*. Dobbeltformen *ministerie/ministerium* er dog af en anden karakter, og det er ikke rimeligt at anbefalingen er styret af frekvens alene. Det er i hvert fald diskutabelt om det altid er bedst at anbefale *ministerium* hvis man bruger formen *ministerie* i det talte sprog. Det bryder med det fonematiske princip, og vi får en meget stor afstand mellem tale og skrift (se Schack 2009:22).

I øvrigt er det uheldigt at DDNO's anbefaling er baseret på en 3 år gammel Googlesøgning. På det tidspunkt var formen *ministerie* endnu ikke en del af den officielle retskrivning og var således en ukorrekt form. Fordelingen af de to former kan derfor godt være noget anderledes i dag (prøv selv at google) – og igen om fx 2 år. Det siger noget om det store vedligeholdelsesarbejde som DDNO's proskriptive metode kræver hvis ordbogens indhold ikke skal være misvisende.

7.2. Anbefalede former i RO?

Hvis Sprognævnet skulle anbefale den ene dobbeltform fremfor den anden, ville det ikke være rimeligt kun at have frekvens som

kriterie. Der er dobbeltformer af forskellig slags, og de er der af forskellige årsager. I DDNO fokuseres der på frekvensundersøgelser i dobbeltformsanbefalingen, men der er andre retskrivningsprincipper at tage hensyn til, sådan som vi har set det ved *højtaler/højtaler* og *ministerium/ministerie* (se også Jervelund 2007:11-23). Det står naturligvis en ordbog som DDNO frit for altid at anbefale den mest frekvente form baseret på Googlesøgninger og at skjule den alternative form efter forgodtbefindende. Så er der truffet et valg for brugerne af ordbogen, men er de ikke blot blevet ført bag lyset? Vi er ikke afvisende overfor anbefalede former, men det er stadig tvivlsomt om man overhovedet kan opstille retningslinjer og udvælgelseskriterier som en institution som Sprognævnet kan arbejde med når det gælder den officielle danske retskrivning.

8. Betydningsoplysninger

I RO 2012 er der betydningsoplysninger ved lidt mere end en tredjedel af opslagsordene, nærmere bestemt de opslagsord "som skønnes at kunne give anledning til tvivl hos især de yngste af ordbogens brugere" (RO 2012:13). Betydningsoplysningerne er som hovedregel ganske kortfattede og tjener kun til identifikation af opslagsordene, og de skal således hverken opfattes som udtømmende eller normative: Retskrivningsordbogen er en retskrivningsordbog, ikke en betydningsordbog. Ikke desto mindre er betydningsoplysningerne en af de ændringer i RO 2012 der har mødt størst kritik fra leksikografisk side (se fx Theilgaard 2013:199; Trap-Jensen 2013:15), også fra Sven Tarp, der mener at selvom de mange betydningsoplysninger har nytteværdi for den tilsigtede brugergruppe (altså de yngste brugere af Retskrivningsordbogen), vil de af andre sprogbrugere blive opfattet som "forstyrrende dataoverlæs" (2013b:134). Tarp kritiserer også betydningsoplysningerne for at være tilfældigt distribueret, en kritik der også frem-

sættes af Henrik Lorentzen i Jervelund & Lorentzen (2013:9ff.) og i Rask (2013:89). Endelig udtrykkes der bekymring over om betydningsoplysningerne vil blive opfattet som normative, og der sættes spørgsmålstegn ved om at det i det hele taget er en del af Sprognævnets arbejdsopgaver at lave betydningsoplysninger i en retskrivningsordbog. Således skriver Henrik Lorentzen at det er “kontroversielt når en retskrivningsordbog, der har til formål at formidle normative oplysninger om retstavning, begiver sig ind på det semantiske område” (Jervelund & Lorentzen 2013:10), mens Tarp skriver at det ser ud til at “Sprognævnet er gået ud over det opdrag, som det oprindeligt har fået af Folketinget” (2013b:133). Til denne kritik må det først og fremmest nævnes at Sprognævnet har understreget og fortsat understreger at Retskrivningsordbogen ikke er en betydningsoordbog, og at de betydningsoplysninger der angives i ordbogen, ikke kan bruges som facitliste for ordenes betydning. Derudover er det vores indtryk at sprogbrugerne – som jo er dem ordbogen er til for – er glade for betydningsoplysningerne, og at der ikke er noget der umiddelbart tyder på at Retskrivningsordbogen rent faktisk bliver opfattet som en betydningsoordbog (hvorimod vi ofte oplever at betydningsoordbøgerne bliver opfattet som retskrivningsordbøger).

Kritikken af at betydningsoplysningerne er spredt tilfældigt ud, beror til dels på en misfortolkning af betydningsoplysningernes funktion (Becker-Christensen 2013:19ff.). Således er en betydningsoplysning unødvendig ved vendingen *gråd og tænders gnidsel* (der identificerer ordet *gnidsel*) fordi Retskrivningsordbogens betydningsoplysninger netop kun skal identificere opslagsordene og ikke skal give en udtømmende beskrivelse. På samme måde er det nok for ordidentifikationen at *guppy* defineres som “en fisk”, mens det må være op til betydningsoordbøgerne at angive en mere specifik definition af fiskearten (jf. Becker-Christensen 2013 *ibid.*). Det er netop den sondring der gør at Retskrivningsordbogen *ikke* er en betydningsoordbog og heller ikke skal opfattes som en betydningso-

ordbog, og som gør at vi holder os indenfor det opdrag som vi har fået af Folketinget. Det er dog ganske givet at der er betydningsoplysninger som man kan diskutere rimeligheden af, fx ved sammensætningseksempler som *godsbanegård*, *godsterminal* og *godsbesidder* (jf. Jervelund & Lorentzen 2013:11), ligesom der ganske givet er betydningsoplysninger der enten er for detaljerede eller for korte. Derfor vil vi da også fremover med Christian Becker-Christensens ord “findyrke og skærpe kriterierne for hvilke ord der behøver ordidentifikatorer, og i hvilken form det skal ske” (2013:23).

9. Netudgaven og den løbende opdatering af RO

Tarp opstiller 11 leksikografiske principper som han mener Sprog-nævnet kan følge i det digitale ordbogsarbejde med RO (Tarp 2013a:197). Blandt de 11 punkter nævnes muligheden for at søge på forældede stave- og bøjningsformer samt stavfejl. Som allerede nævnt, er det netop sådanne forbedringer af den digitale RO som vi planlægger at indføre. Vi er desuden enige med Tarp i at bøjningsformer bør vises i deres fuldform i netudgaven af RO. Tarp foreslår også at alle metaleksikografiske forkortelser skrives i deres fuldform. Det mener vi dog ikke er nødvendigt, for de bliver vist i deres fuldform når man kører musen hen over dem. Det gælder også *i sms.*, som Tarp betegner som en ny kryptisk forkortelse vi har indført ad bagevejen (Tarp 2013a:192). Det er nu ikke rigtigt. Den har eksisteret i RO i mere end 25 år, dvs. siden 1986-udgaven af RO.

Fire af de punkter som Tarp nævner, har været drøftet i Sprog-nævnet i løbet af 2013 i forbindelse med udarbejdelsen af strategien for den løbende opdatering af den digitale RO (se Jervelund 2014:3-7). Helt på linje med Tarps forslag er det blevet besluttet at der skal tilføjes nye opslagsord løbende, og dette arbejde er godt i gang. Normændringer (dvs. ændringer af stave-, bøjnings- og

sammensætningsformer eller ændringer i retskrivningsreglerne) vil derimod ikke blive indført i den løbende opdatering, men først efter en årrække, sandsynligvis efter ca. 10 år, og det svarer vist ganske godt til Tarps forventninger på dette område.

Tarp skriver desuden at vi ikke skal slette “gamle” ord som er gledet ud af den trykte ordbog. Det kan vi nu ikke love. Fjernelse af forældede ord er ganske vist ikke en del af den årlige opdatering ifølge strategien, men hvis redaktørerne under redigeringen støder på ord som ikke længere hører til i RO, kan ordene få et særligt forældelsesmærke, og efter 5-10 år regner vi så med at tage en lille portion forældede ord ud af ordbogen. RO skal nemlig ikke udvikle sig til en historisk ordbog, hverken i trykt eller digital form, hvor vi angiver nye retskrivningsændringer i forældede ord. Hvis ikke vi havde slettet det gamle ord *kultusministerium* i RO 2012, ville vi have haft formen *kultusministerie* ved siden af *kultusministerium*, og det virker besynderligt (Jervelund & Lorentzen 2013:8-9).

10. Afslutning

Vi har i denne artikel kommenteret en række af Sven Tarps kritikpunkter, og vi har gjort rede for de valg vi har truffet i arbejdet med RO 2012. I sin artikel giver Tarp også udtryk for at der er en række forbedringer at spore i RO 2012. Han nævner bl.a. at endelserne (især datidsendelser og bestemt form pluralis af substantiver) bliver vist på en mere brugervenlig måde end tidligere (Tarp 2013a:191ff.), og at der er kommet orddelingsangivelser i alle delelige opslagsord.

Som vi har vist her i artiklen, vil vi give Tarp ret i at der stadig er plads til forbedringer på visse områder. Der er bl.a. behov for at tilføje nye opslagsord løbende, hvilket vi er i gang med, der skal arbejdes med kriterierne for hvilke ord der behøver ordidentifikatorer, og den digitale ordbogs søgefaciliteter skal forbedres.

I Dansk Sprognævn er vi en god blanding af erfarne ordbogsredaktører, lingvister og datalingvister, og selvom den digitale fremtid vel er en udfordring for os alle i samfundet, føler vi os ganske godt rustet til at gå den i møde på ordbogsfronten.

Litteratur

Ordbøger

DDO = *Den Danske Ordbog* (2003-2005). Online: <ordnet.dk/ddo> (april 2014).

DDNO = Henning Bergenholtz i samarbejde med Heidi Agerbo Pedersen m.fl.: *Den Danske Netordbog*. <www.ordbogen.com> (april 2014).

Nudansk Ordbog = Politikens Nudansk Ordbog med etymologi. <www.ordbogen.com> (april 2014).

RO 2012 = *Retskrivningsordbogen*. 4. udg. Udgivet af Dansk Sprognævn. Online: <www.dsn.dk/ro> (april 2014).

SAOL = *Svenska Akademiens ordlista över svenska språket* (2006). 13 uppl. <http://www.svenskaakademien.se/svenska_spraket/svenska_akademiens_ordlista/saol_pa_natet/ordlista> (april 2014).

Anden litteratur

Becker-Christensen, Christian (2013): Betydningsangivelser i Retskrivningsordbogen 2012 – replik til Pingpong om Retskrivningsordbogen 2012, LEDA-Nyt nr. 55, marts 2013. I: *LEDA-Nyt nr. 56*, december 2013, 18-24.

Bergenholtz, Henning & Christina Bøgelund (2002): Hvor præskriptiv er en deskriptiv ordbog? Hvor deskriptiv er en præskriptiv ordbog? I: *LexicoNordica* 9, 79-108.

- Galberg Jacobsen, Henrik (1999): Ud med dobbeltformerne? I: *Mål og Mæle*, 22. årgang, nr. 1, 4-10.
- Galberg Jacobsen, Henrik (2010): *Ret og skrift. Officiel dansk retskrivning 1739-2005*. Bind 1-2. Dansk Sprognævns skrifter 42. Odense: Syddansk Universitetsforlag.
- Hansen, Erik (1997): Høj(t)taler. I: *Nyt fra Sprognævnet* 1997/1, 11-12.
- Hansen, Erik (2000): Normering og sprognævn. I: Anne Holmen & J. Normann Jørgensen (red.): *Sprogs status i Danmark år 2011*. Københavnerstudier i tosprogethed 32. Danmarks pædagogiske Universitet, 7-20.
- Jervelund, Anita Ågerup (2007): *Sådan staver vi – om ortografi og stavfejl*. Dansk lærerforeningens Forlag og Dansk Sprognævn.
- Jervelund, Anita Ågerup & Henrik Lorentzen (2013): Pingpong om Retskrivningsordbogen 2012. I: *LEDA-Nyt nr. 55*, marts 2013, 6-15.
- Jervelund, Anita Ågerup (2014): 100 nye ord om året – tilføjelse af nye ord i RO. I: *Nyt fra Sprognævnet* 2014/2, 3-7.
- Nørby Jensen, Jørgen (2006): Bindestrøg eller ej? Om skrivemåden af nogle engelske låneord. I: Anita Ågerup Jervelund, Marianne Rathje & Jørgen Schack (red.): *Vi skriver dig til. Festskrift til Vibeke Sandersen i anledning af 70-års-dagen*. Dansk Sprognævns skrifter 36. Dansk Sprognævn 2006, 127-144.
- Nørby Jensen, Jørgen (2013): Ændrede staveformer og ordformer i Retskrivningsordbogen 2012. I: *Nyt fra Sprognævnet* 2013/1, 13-17.
- Rask, Kirsten (2013): Anmeldelse. RO – Rettidig omhu? I: *NyS* 44, 81-92.
- Schack, Jørgen (2009): Skal vi have former af typen *gymnasie* i Retskrivningsordbogen? I: *Nyt fra Sprognævnet* 2009/4, 20-23.
- Tarp, Sven (2002): Kritiske bemærkninger til den officielle danske retskrivningsordbog. I: *LexicoNordica* 9, 195-208.

- Tarp, Sven (2013a): Retskrivningsordbog på godt og ondt. I: *LexicoNordica* 20, 185-200.
- Tarp, Sven (2013b): Retskrivningsordbog mellem to stole. I: *Hermes* 50, 127-144.
- Theilgaard, Liisa (2013): Anmeldelse. Retskrivningsordbogen, 4. udgave, 2012. I: *Danske Studier* 2013, 197-202.
- Trap-Jensen, Lars (2013): Retskrivningsordbogen 2012 – en bedre ordbog? I: *Mål og Mæle*, 35. årgang, nr. 1, 9-16.

Margrethe Heidemann
Andersen
seniorforsker
Dansk Sprognævn
Worsaaesvej 19, 4.
DK-1972 Frederiksberg C
heidemann@dsn.dk

Jørgen Nørby Jensen
informationsmedarbejder
Dansk Sprognævn
Worsaaesvej 19, 4.
DK-1972 Frederiksberg C
jn.jensen@dsn.dk

Anita Ågerup Jervelund
seniorkonsulent
Dansk Sprognævn
Worsaaesvej 19, 4.
DK-1972 Frederiksberg C
aagerup@dsn.dk

Små ordbøker – mindre og annerledes

Jan Engh

In lexicographic literature, little attention has been paid to small dictionaries, oddly enough, since they are probably the more frequently used. Further, they are not necessarily easier to edit than extensive ones: there is the selection problem, and the size entails constraints on the compilation. This becomes clear when small dictionaries are automatically analysed for conversion to database format. In this article, several aspects of small dictionaries will be discussed on the basis of the author's experience with the compilation and subsequent conversion of a pocket-size Portuguese-Norwegian-Portuguese dictionary.

1. Innledning

På leksikografikonferanser og i fagtidsskrifter er emnet påfallende ofte noe som gjelder store ordboksprosjekt, trykte og/eller i elektronisk form, og alltid med analytikerens krav til rasjonalitet, gjengitt med og styrt av datamaskinens pedanteri. Lite eller ikke noe blir nevnt om små ordbøker. Ikke noe om små enkeltordbøker og ikke noe allment om formatet i små ordbøker eller om hvordan en går fram for å redigere dem. Jeg trur heller ikke at virkelig små ordbøker kommer i betraktning når en deler ut priser for gode ordbøker. Det er underlig, for folk flest bruker nok små ordbøker vel så ofte som store. Små ordbøker er viktige. Og det er ikke gitt at det er ukomplisert å lage dem fordi de er små. En ting er omfanget og utvalgsproblemet: Hvordan skal en velge ut (bare) viktige oppslagsord? Små ordbøker er også *annerledes*. De er tenkt på en annen måte enn større ordbøker. Ikke bare fordi de har et sammentrengt og sparsomt format nettopp fordi de er små, de skal

brukes på en spesiell måte også. I denne artikkelen skal jeg drøfte enkelte sider ved små ordbøker og hvordan de skiller seg fra større ordbøker og ordboksverk i mer enn det ytre.

2. Utgangspunktet

Å lage ordbøker er på sett og vis filologi som skjønn kunst. Vel er det snakk om filologi med vitenskapelige krav til kjeldekritikk og språklig nøyaktighet og representativitet i språksamfunnet, men det er også en god del skjønn og kreativ skriving med i bildet. Det gjelder å skrive godt, og det av flere grunner: En skriver for folk med andre forutsetninger enn en sjøl har – ikke bare språklig, men også intellektuelt (og analytisk). Og her er det ikke (bare) snakk om popularisering: I de fleste ordbøker er det om å gjøre å få leserne til å skjønne, slik at de kan *bruke* opplysningene de leser. I små ordbøker, som er lagd for at en skal kunne bære dem på seg i hverdagen og på ferie, gjelder det først og fremst å lede leseren raskt og effektivt til det han/hun er ute etter – ved hjelp av språket, siden søk er umulig i trykte bøker – og uten å fortape seg i irrelevante opplysninger, som kan se nødvendige ut for systematikkens skyld. Måten dette skjer på, trenger ikke være ens for alle ordbøker av samme slag.

På slutten av 1980-tallet hørte jeg hvor mange titusen uregelmessigheter en hadde funnet da IBM gjorde *Diccionario de la lengua española* (RA) om til en database. En konvertering av *Bokmålsordboka* (BOB) (Baustad 1992) pekte i samme retning. Defekter og dårlig kvalitet kom til syne i ellers anerkjente ordbøker. Det var det nok også som oftest, men spørsmålet er om en kan snakke om defekter og dårlig kvalitet uten referanse til størrelse og format.

3. Ei ny ordbok¹

I 1992 begynte jeg å skrive på det som seinere blei til Engh (2004 og seinere 2008). Den dekker europeisk portugisisk, og trass i liten størrelse (11 x 8 x 2,5 cm) og ørlita skrift er det ei forholdsvis omfattende ordbok på rundt 33 000 oppslagsord begge veger, der artiklene har relativt fylldige definisjonsdeler med en god del fraseologi. Likevel, slik jeg skal vise, er boka tenkt som ei praktisk ordbok til bruk i dagliglivet.

Til min store overraskelse solgte ordboka bra. Og til min store tilfredshet har ingen funnet alvorlige mangler i den. Tvert imot. Alle med greie på portugisisk og norsk har sagt at ordboka er god, slik at brukerne kan skjønne hvordan de kan bruke det som står i den riktig.

I 2012 ville så forlaget lage ei elektronisk utgave, både som grunnlag for oppslag på skjermen og for å lette vedlikeholdet av ordboka seinere, uansett publiseringsform. Fra før av forelå ordboka bare som en fil med tekst. Nå brukte forlaget et analyseprogram for å lage ei XML-utgave. En slik analyse går ut på at programmet kjenner igjen enheter i satsfila på reint formelt grunnlag, og legger dem i egne felt. Resultatet var hårreisende.

Den første elektroniske utgava av Engh (2008) inneholdt ei rekke systemgenererte feil. Plutselig dukka det opp et komma fulgt av et semikolon, et enkeltord i en frase stod på ei egen linje med nummer som innholdsvariant osv. Det er for så vidt trivielt. Men i tillegg var resultatet prega av feilanalyser i alle andre enn helt enkle ordboksartikler. Fra en leksikografisk analytikers synsvinkel dreide det seg om feil eller inkonsekvenser, her som i RA og i BOB. Jeg hadde på sett og vis ikke vært konsekvent i bruken av de grafiske midlene jeg hadde til rådighet for å markere syntaksen i ordboksartiklene. Eller hadde jeg det? Meg bekjent hadde ingen

1 Takk til Diana Santos. Uten henne, verken ordbok eller artikkel.

misforstått. Fantes det en analytikerlogikk i bruken av syntaksen til forskjell fra en brukerlogikk? Slik jeg ser det, ligger forklaringa i hvordan jeg hadde brukt en fattig ordbokssyntaks kreativt for at brukerne skulle få et riktig inntrykk av hva ord betyr, og hvordan de brukes. Slik at de skulle finne de opplysningene de trengte raskt og greit. Jeg skal forklare hvordan, og for å gjøre det enklere vil jeg bare se på artikler fra portugisisk til norsk.

4. Bakgrunn og forbilde

Da jeg begynte arbeidet med ordboka i 1992, var det som et ledd i en læreprosess. Jeg lærte meg portugisisk – til de grader at nesten hele mitt privatliv kom til å foregå på portugisisk – samtidig som jeg merka meg nøye hva jeg trengte og oppdaga om det nye språket mitt. Det som gjorde meg til en slik «overgangsperson» med stadig bedre forutsetninger for å lage ei ordbok til og fra et fremmedspråk, lærte meg også hva som måtte være viktig for andre og potensielle ordboksbrukere i samme situasjon.

Det er slik en er nødt til å lage ordbøker når det ikke finnes andre ordbøker for språka det gjelder, fra før av. Og det var tilfellet for portugisisk og norsk – med ett unntak: Nilsson (1994), som er både omfattende og god, men bare *fra* norsk til portugisisk. Når jeg mente at jeg kunne gi meg i kast med en slik oppgave, var det også fordi jeg hadde en ganske omfattende *allmenn* leksikografisk erfaring, ikke bare fra redigeringa av BOB, men også fra arbeid med elektroniske ordbøker i flere format (se Engh 2009).

Den gangen fantes det ingen elektroniske språkressurser å tære på, slik det fins nå, for eksempel *Linguateca*, og selve arbeidet med ordboka skjedde på tradisjonelt håndverksmessig vis (sjøl om det var vekselvis på MAC og PC). Forbildet var i første rekke Blichfeldt & Aasen (1990). Den var temmelig tynn både i valg av oppslagsord (i alt mellom 16 og 17 000) og i ordboksartiklene, og framfor alt var

den svært enkel når det gjaldt formatet og syntaksen: Ingen romertall eller bokstaver ble brukt for å skille homonymer, som ‘agitator, oppvigler; visp’ for **agitador** eller de etter alt å dømme urelaterete betydningene av **abombado**: ‘konveks; bedrevet’. Innholdsvarianter av samme betydning ble heller ikke skilt ut med tall, jf.²

abordar nærme seg, borde.

Ikke fraseologi av ulike slag heller, slik disse artiklene viser:

abajo nede, nedenunder, ned;
j~ los terroristas! ned med
 terroristene!

acabar (av)slutte; ~ *de + inf*
 nettopp ha/være + *perf.*
part.; *acabo de comer* jeg har
 nettopp spist; -**se** være slutt.

-**se** i artikkelen **acabar** er tenkt som tillegg til oppslagsordet, og må forstås som **acabar-se**. Videre er «~» et tegn på at en kan sette inn oppslagsordet i nettopp den formen akkurat der. Blichfeldt & Aasen (1990) lot dessuten en vertikal strek skille ut rota i oppslagsordet, slik at en kunne nøye seg med å skrive inn (avlednings)endelser i artikkelen, som i **abort**|ar ... -**o**, som representerer henholdsvis **abortar** og **aborto**.

Slik det går fram, var de eneste grafiske virkemidlene som ble brukt ellers, kursivering av fraseologi, komma og semikolon i tillegg til kolon og vanlige parenteser. Og det er typisk at en stor del av artiklene bare inneholdt én enkelt ekvivalent til oppslagsordet, «**abotonar** knappe». Likevel var dette ikke noen dårlig lommeordbok.

² Ordboksartiklene slik de stod i den trykte ordboka, står i ramme.

5. Format og redigering

Jeg overtok ikke opplegget i den spanske ordboka uten endringer, men la til ordklasseopplysninger og genus for substantiv, og fjerna punktum etter hver ordboksartikkel. Jeg lot også være å ta i bruk virkemiddel av typen vertikal strek. Dette for å unngå misforståelser og for å gjøre ordboka lettere å bruke. Hadde jeg gjort som i Blichfeldt & Aasen (1990), ville en bruker ha vært nødt til å huske noe som det jo er unaturlig å huske, og som det derfor er vanskeligere å huske: deler av ord, som framstår som vilkårlige bokstavsekvenser (særlig de «tekniske røttene», som ikke er røtter i språkvitenskapelig forstand), og tilfeldige homonymer. For å vise noen portugisiske eksempler av det slaget jeg ikke ville ha i ordboka: **bomb** er en slik vilkårlig sekvens, mens dette ville være en opphoping av «avledninger»: **bomb|a** ‘pumpe’; ‘bombe’, ~**ardeamento** ‘bombing’, ~**ardear** ‘bombe’, ~**ardeiro** ‘bombefly’, ~**ear** ‘pumpe’, ~**eiro** ‘brannmann’. Økonomi måtte ikke gå utover hvor lett det skulle være å forstå og huske opplysningene.

Men etter mønster av den spanske ordboka bestemte jeg meg for at komma skulle skille mellom (tilnærmet) likeverdige ord og uttrykk, fullstendige eller delvise synonymer, mens semikolon skulle skille ulike betydninger. Siden det ikke var flere skilletegn å ta av, måtte jeg også bruke semikolon for å vise at det som kom etter i kursiv, var fraseologi. Det ble mye fraseologi – av flere grunner: For det første fordi det som framstår som fraseologi på portugisisk, kan svare til sammensatte ord i norsk (for eksempel *porto de abrigo* ‘nødhavn’). For det andre for å markere vanlige eller interessante bruksmåter (ofte kollokasjoner). Og for det tredje for å vise særskilte betydninger og betydningsnyanser av oppslagsordet i bestemte sammenhenger.

I stedet for at alle underbetydninger og fraseologi var ordna etter et rett fram logisk prinsipp, bygde det på assosiativitet. En

betydning hadde noe med neste betydning å gjøre: en presisering eller en spesifisering (Næss 1975:23, 28). Til grunn kunne det ligge allmenne prinsipper som kausalitet, men ikke nødvendigvis. Noen ganger kunne en felles assosiasjon til noe tredje, for eksempel til 'politikk', være forbindelsen mellom to innholdsvarianter. Og endelig var konkret til forskjell fra overført betydning også ofte viktig.

En vanlig bruker av ordboka er ikke i tvil om hva som er hva, som regel uten å tenke nærmere over det. Det henger sammen med at en *forstår* hva som står, i det minste ett av språka, og når en har funnet det en leiter etter, legger en ordboka fra seg igjen. Men en leksikograf som *analyserer* boka systematisk, artikkel for artikkel, ser det som isolert sett er inkonsekvenser i den logiske oppbygningen – og det gjør i hvert fall en datamaskin når den kjører et analyseprogram som analytikeren har lagd. Det var her det gikk galt da forlaget skulle analysere ordboka automatisk.

Alt gikk greit så lenge ordboksartiklene var enkle på den rette måten:³

<p>alternadamente <i>adv</i> vekselvis</p>

alternadamente *adv*
vekselvis

<p>achar <i>v</i> finne, oppdage; finne opp; oppleve; ~ <i>dificuldades</i> støte på vansker; synes, mene, tru; <i>Não acha?</i> Ikke sant?</p>
--

3 Resultatet av den første automatiske analysen står uten ramme her og i fortsettelsen.

achar *v*

1. finne, oppdage
2. finne opp
3. oppleve

*achar dificuldades støte på vansker*⁴

4. synes, mene, tru

Não acha? Ikke sant?

Ved mer kompliserte artikler ble det feil.

6. Mislykka forsøk på automatisk analyse

I den trykte ordboka stod det slik:

abaixo *adv* ned; nedover;
morro ~ nedoverbakke,
utfor; *rio* ~ nedover elva;
Abaixo a ditadura! Ned
med diktaturet!; ~ *assinado*
undertegnede; ~ *de*
nedenfor; av lavere grad
enn; ~ *dos três anos* under
tre år(s) alder)

Den elektroniske utgava så slik ut:⁵

4 I den elektroniske utgava stod både fraseologi og oversettelsen av den i kursiv, men oversettelsen var skilt ut ved hjelp av understreking.

5 Alle eksemplene her er fra det første og mislykka forsøket på omgjøring til nytt elektronisk format. Den elektroniske ordboka som ligger på nettet nå, har blitt retta. Jf. Ordnett.

abaixo *adv*

1. ned

2. nedover

*morro abaixo nedoverbakke, utfor**rio abaixo nedover elva**Abaixo a ditadura! Ned med diktaturet!**abaixo assinado undertegnede**abaixo de nedenfor;*

3. av lavere grad enn

abaixo dos três anos under tre år(s) alder

Leser en den norske oversettelsen linje for linje, virker det forvirrende. Her er det klart at ganske ulike betydninger er samla under ett og samme innholdsnummer. Dessuten er det grunnleggende prinsippet om substituerbarhet (Knudsen & Sommerfelt 1958) ikke fulgt: En kan ikke sette inn *av lavere grad enn* som oversettelse av bare ordet *abaixo* i en portugisisk utgangstekst. Det kan bare settes inn i stedet for *abaixo de*.

Hva hadde skjedd? Semikolon var helt klart tolka som linjeskift, og hvis ikke det som stod etter, kom i kursiv (og derfor ble oppfatta som fraseologi), ble det generert et nytt innholdsnummer, og linja begynte ytterst i venstremargen. Dette må en ha lagt inn slik i programmet fordi en mente at all fraseologi var «illustrasjoner», og at semikolon var en slags sammenkjeding av uttrykk som trakk veksler på en og samme underbetydning av oppslagsordet. Men resultatet av en slik analyse er språklig sett galt, og det var naturligvis ikke slik denne artikkelen var tenkt fra mi side. Leser en den ovenfra og ned (sic!) som bruker, så virker ikke artikkelen påfallende ulogisk, langt mindre feil. Den er bygd opp som ei rekke assosiative relasjoner for å kompensere for dårlig plass og manglende syntaktiske virkemidler: Fra grunnleggende betydning 'ned', så 'nedover' (på veg ned) og deretter forskjellige måter å dra nedover på (*rio abaixo* og *morro abaixo*, som attpåtil svarer til et eget ord,

nedoverbakke), før metaforisk ‘nedover’ i uttrykket *Abaixo a ditadura!* ‘Ned med diktaturet!’. Her er det et slags brudd, i det en kan assosiere til politisk virksomhet og opprop (*abaixo assinado*), før jeg nevner hva ‘nedenfor’ svarer til i et preposisjonsuttrykk (med *de*). Til sist den spesielle betydningen ‘av lavere grad enn’ og *abaixo* brukt som preposisjon med *de* igjen i denne betydningen.

På tilsvarende vis er definisjonsdelen av **abrir** tenkt som en kjede av assosiative relasjoner.

abrir *v* åpne; ~ *para dentro*
 åpnes innover, slå
 inn(over); ~ *para fora*
 åpnes utover, slå ut(over);
 ~ *para vende ut mot*; låse
 opp; løsne; sette opp; slå
 opp; ~ *as pernas* skreve

Likevel kunne den ved en tilfældighet analyseres mekanisk på samme måte som *abaixo*, men med et noenlunde riktig resultat (*abrir para* ‘vende ut mot’ er riktignok litt på kanten) – helt til en kommer til slutten. Men da går det galt, for *abrir as pernas* ‘skreve’ har fint lite med ‘slå opp’ å gjøre:

abrir *v*

1. åpne

abrir para dentro åpnes innover, slå inn(over)

abrir para fora åpnes utover, slå ut(over)

abrir para vende ut mot

2. låse opp

3. løsne

4. sette opp

5. slå opp

abrir as pernas skreve

I tilfellet **abrir** var en spesiell betydning degradert til fraseologisk eksempel. I **afligir-se** er en av betydningene til verbet med preposisjon gjort til innholdsvariant av verbet uten.

afligir-se *v* ~ *com* engste/
uroe seg for; grue seg for

afligir-se *v*
afligir-se com engste/uroe seg for;
grue seg for

Her hadde jeg også satsa på at en bruker skulle skjønne hva *afligir-se* betyr, når han/hun ser det uten *com*. Slik en vanligvis skjønner at transitive verb kan bli brukt intransitivt, uten noe direkte objekt. Skulle jeg heller ha skrevet definisjonen slik?

afligir-se *v* engste seg; ~ *com*
engste seg for; grue seg; ~ *com*
grue seg for

Når jeg ikke gjorde det, var det for å spare plass. Ei linje ekstra kan virke lite, men det utgjør en økning på 50 % i dette tilfellet, og slike artikler var det mange av i ordboka. Slikt blir det fort mange sider av, og det blir kritisk om boka skal være i lommeformat. Ikke hadde det blitt enklere å forstå heller, noe som er vel så viktig.

Behandlinga av det intransitive verbet og den preposisjonsutfyllinga det kan ta, peker mot et annet språklig trekk som ikke er behandla på en logisk måte sett fra en analytikers synsvinkel: transitivitet. Jeg hadde valgt å ikke angi om et verb var transitivt eller intransitivt. Ikke bare fordi det spanske forbildet ikke gjorde det. Det hadde en viss språkanalytisk forankring også, siden skillet mellom transitive og intransitive verb ikke er så skarpt som en gjerne liker å tru (se for eksempel Sveen 1990 og Faarlund, Lie &

Vannebo 1997:663–666). Men det var også fordi redigeringa var styrt av et assosiativt prinsipp, der transitiv og intransitiv bruk av verb ble gruppert sammen og ikke definert hver for seg – slik at en til dels hadde blitt nødt til å føre dem opp dobbelt. På den måten ble det også spart plass.

Acabar er nettopp et eksempel der transitive og intransitive betydninger kommer tilsynelatende hulter til bulter.⁶

acabar *v* avslutte; fullføre, gjøre fullstendig; oppfylle; få unnagjort; greie; ikke gå lenger; ikke være mer igjen; ende, havne; dø; opphøre, slutte; *Acabou!* Slutt (med det der)!, Kutt ut! (dagl); ~ *com* slutte med; gjøre slutt på, bruke opp; ~ *o café* drikke opp kaffen; rydde av vegen, kverke, ta knekken på (dagl); ... *ter acabado* ... er omme; ~ *de* nettopp ha; nettopp ha blitt; ~ *em* ende ved/med/i; ~ *por* ende med (å)

Lenge gikk det bra i den automatiske analysen. Først fem innholdsnummer (1–5) med det transitive *acabar*, så fem (6–0⁷) med det intransitive, inntil *acabar com*, som burde ha hatt et eget innholdsnummer, og som både kan bety ‘slutte med’, ‘gjøre slutt på’ og

6 En klar mangel som ikke angår resonnementet her, og som vil bli retta i neste utgave, er at den aspektuelle sida av *acabar* ikke kommer tydelig fram: *acabar de* ‘slutte å’.

7 Dette er naturligvis en triviell feil. Det skal stå 10.

‘bruke opp’. *Acabar* i *acabar o café*, derimot, er transitivt, men med en liknende betydning. ‘Rydde av vegen, kverke, ta knekken på’ er en annen betydning av *acabar*. Ikke bare er oppdelinga i innholdsnummer misvisende med hensyn til betydning, den roter også til forskjellen mellom den transitive og intransitive bruken av verbet.

acabar v

1. avslutte
2. fullføre, gjøre fullstendig
3. oppfylle
4. få unnagjort
5. greie
6. ikke gå lenger
7. ikke være mer igjen
8. ende, havne
9. dø
0. opphøre, slutte
Acabou! Slutt (med det der)!, Kutt ut! (dagl)
acabar com slutte med
1. gjøre slutt på, bruke opp
acabar o café drikke opp kaffen
2. rydde av vegen, kverke, ta knekken på (dagl)
... ter acabado ... er omme
acabar de nettopp ha
3. nettopp ha blitt
acabar em ende ved/med/i
acabar por ende med (å)

Tilfellet **acabar** viser også noe annet. Det lar seg ikke alltid gjøre å stille opp ubrutte kjeder av assosiative relasjoner. Noen ganger oppstår det brudd. I artikkelen **acabar** kommer det et slikt brudd mellom ‘greie’ og ‘ikke gå lenger’. I ei ordbok med en rikere syntaks kunne en ha skilt mellom slike hovedbetydninger, for eksempel

ved hjelp av en bokstav: «**acabar** *v* A avslutte; fullføre, [...] **B** ikke gå lenger; ikke være mer [...]». Her var ikke dette noen mulighet.

Til slutt et eksempel som omfatter et annet gjennomgående fenomen:

andar *m* etasje *m*; *primeiro*
 ~ andre etasje; det å gå;
 gange *m*, ganglag *n*; *v* gå;
 ~ *a pé* gå (til fots); ~ *de*
gatas kripe; *Vamos*
andando! La oss gå!; *Anda*
cá! Kom hit!; dra, reise,
 fare; ~ *de avião* reise med
 fly; ~ *com* være sammen
 med, ha fast følge med;
 ~ *de mal com* være uenig
 med; ~ *de pé atrás* være
 mistenksom; ~ *sempre*
atrás de alltid være ute
 etter

Her også er definisjonsdelen i store trekk bygd opp på samme måte som de andre vi har sett på,⁸ men til sist kommer det et par eksempler på fraseologi, der det er snakk om en overført bruk av betydningen av verbet i idiomatiske uttrykk. Jeg har rett og slett samla dem sist. De utgjør en slags restkategori, uten at jeg hadde grafiske virkemidler til rådighet som kunne markere dette. Det er ingen grunn til å analysere hele artikkelen, men den siste delen er interessant:

8 Og som i tilfellet *acabar* mangler det dessverre klare opplysninger om det aspektuelle *andar*, *andar a* 'drive å', 'gå rundt å' osv. Jf. en setning som *Ele anda a dizer muitos disparates* 'Han sier mye rart (bestandig)'.

andar *m*

[...]

5. dra, reise, fare

andar de avião reise med fly

andar com være sammen med, ha fast følge med

andar de mal com være uenig med

andar de pé atrás være mistenksom

andar sempre atrás de alltid være ute etter

Her er *andar com* ‘være sammen med, ha fast følge med’, *andar de mal com* ‘være uenig med’, *andar de pé atrás* ‘være mistenksom’ og *andar sempre atrás de* ‘alltid være ute etter’ representert som fraseologi under ‘dra, reise, fare’. Det er feil.

7. Årsak og mulig alternativ

Ovenfor har jeg nevnt en del eksempel på hva som ble feil og hvordan. Det var språklige feil, som skyldtes det som fra en strengt formell synsvinkel var feil i satsen til ordboka. Men hva var det som lå til grunn når programmet som analyserte ordboksteksten og fordelte innholdet i forskjellige felt, ikke gav et riktig resultat? Etter alt å dømme var det fordi *forståelsen* manglet. Programmet var lagd av nordmenn som nok hadde god erfaring fra ordboksarbeid og grep om de praktiske sidene ved databehandling av språk de kunne. Men de kunne ikke portugisisk, og de skulle heller ikke bruke ordboka som ordbok. Analysen tok verken hensyn til hvilken betydning ord og uttrykk hadde, eller til intensjonene bak redigeringa.

Reint teoretisk ville det kanskje vært mulig å gjøre om ordboka fra sats til elektronisk utgave automatisk, for eksempel ved hjelp av teknikker som ser etter hvordan ord og uttrykk forekommer i samband med hverandre i store korpus eller ordbøker. Men

når det gjelder portugisisk og norsk, er det ikke mye å søke i: Den eneste ordboka verdt navnet, er Nilsson (1994). Men den går som nevnt bare én veg, fra norsk til portugisisk, og den er ikke tilgjengelig i digital form. I praksis hadde da Kunnskapsforlaget heller verken ressurser eller den rette kompetansen til å prøve noe slikt. Manglende forståelse gjorde også at en heller ikke skjønnte at en burde ha satset på en løsning der det meste analyseres automatisk, men der systemet stopper og spør når det er grunn til det.

8. Avslutning

Slik jeg har vist, kan et enkelt analyseprogram ta feil når det skal gjøre satsen til ei trykt lita ordbok om til ei elektronisk ordbok. I tillegg til reint trivielle feil, dreier det seg om feil som skyldes inkonsekvenser og unøyaktigheter, dersom en ser resultatet fra en analytikers synsvinkel. Slike feil forbinder en gjerne med dårlig kvalitet. Men slik jeg ser det, skyldes de feila jeg har nevnt, ikke at den trykte ordboka er av dårlig kvalitet som lita ordbok.⁹ I stedet har det med formatet i små ordbøker å gjøre. I praksis *kan* små ordbøker være redigert etter helt andre retningslinjer enn store ordbøker.

Både plassmangelen og hensikten med små trykte ordbøker stiller andre krav til redigeringa.¹⁰ Små ordbøker er ment for raske oppslag i det praktiske liv, der en er fornøyd når en finner en betydning som passer i situasjonen der og da. Gjerne det vi med en ikke-vitenskapelig term kaller hovedbetydning – en betydning som

9 At ordboka inneholdt andre feil, er helt klart. De manglende aspektuelle betydningene av *acabar* og *andar* er alt nevnt. Formelle feil fantes også, men det er irrelevant her.

10 For elektroniske ordbøker blir størrelsen – antall ord, linjer eller bytes – mindre viktig. Det er en oppgave for seg å undersøke hvordan små trykte og elektroniske ordbøker skiller seg fra hverandre, eller om store ordbøker lar seg bruke i «små» sammenhenger, for eksempel ved oppslag over telefon.

forekommer oftest i praksis, og som alle kan være enige om uansett etymologi, typologiske forhold som avledning eller annen systematikk. Uansett hva analytikere måtte mene (jf. Kilgarriff 1997), så er det menneskelig å tenke i hovedbetydninger. Derfor kan ikke ordboksredaktøren gå systematisk og rasjonelt til verks, slik en venter av omfattende ordboksverk – og slik en rasjonell analyse krever. Dessuten er det sjølsagt ikke noe alternativ å bare ramse opp ord og betydninger, om det så var i alfabetisk orden. I stedet må en utnytte mulige assosiasjoner mellom ord og betydninger, og skrive kreativt. Redaksjonen får et visst preg av retorikk – for ikke å si skjønnlitteratur: Ordboksredaktøren skal nevne det viktigste først, det som det er mest sannsynlig at leseren er på jakt etter, og så «drive» leseren fram ned igjennom definisjonsdelen uten at leseren mister interessen og gir opp. Helt til leseren forhåpentligvis finner den rette betydningen. Det kan skje ved hjelp av assosiative sprang i en slags handling, som brukeren forstår, og som gjør det lett å lese: En indre konsekvens eller koherens som må gjelde hele ordboka, men som kan variere fra ordbok til ordbok i måten den karrige syntaksen blir utnyttet på. Filologi som skjønn kunst, med andre ord.

Det er dette som kommer til syne når en underkaster små trykte ordbøker den samme overflatiske, reinte formelle analysen som større og langt mer omfattende ordbøker med et format og en syntaks deretter. Derfor blir det feil å gjøre mislykka formell analyse av små ordbøker til et spørsmål om (bare) kvalitet. Det er klart at feil kan oppstå fordi kvaliteten ikke er god nok – men også fordi det skal være sånn. Små ordbøker har *andre* kvaliteter.

Litteratur

Ordbøker

- Blichfeldt, Signe Flydal & María Luisa Villanueva Aasen (1990): *Spansk lommeordbok: spansk-norsk, norsk-spansk*. Oslo: Kunnskapsforlaget.
- BOB (1986) = Marit Ingebjørg Landrø & Boye Wangensteen (red.): *Bokmålsordboka*. Oslo: Universitetsforlaget.
- Eng, Jan (2004): *Portugisisk lommeordbok: portugûês-norueguês, norueguês-portugûês*. (Lommeordbøker fra Kunnskapsforlaget.) [2. utgave 2008] Oslo: Kunnskapsforlaget.
- Nilsson, Kåre (1994): *Norsk-portugisisk ordbok*. [2. utgave] Oslo: Universitetsforlaget.
- Ordnett = <<http://www.ordnett.no/>> (februar 2014).
- RA = *Diccionario de la lengua española*. (1984) [20. utgave] Madrid: Real Academia Española.

Annen litteratur

- Baustad, Jostein (1992): Automatisk analyse av maskinleselige ordbøker til bruk i en orddatabase. I: Ruth Vatvedt Fjeld (red.): *Nordiske studier i leksikografi. Rapport fra Nordisk konferanse i leksikografi i Oslo, mai 1991*. Oslo, 423–431.
- Eng, Jan (2009): Lexicography for IBM. Developing Norwegian linguistic resources in the 1980s. I: John Impagliazzo, Timo Järvi & Petri Paju (eds.): *History of Nordic Computing 2. Second IFIP WG 9.7 Conference, HINC2, Turku, Finland, August 21–23, 2007. Revised Selected Papers*. (IFIP advances in information and communication technology 303.) Boston: Springer, 258–270.

- Faarlund, Jan Terje, Svein Lie & Kjell Ivar Vannebo (1997): *Norsk referansegrammatikk*. Oslo: Universitetsforlaget.
- Kilgarriff, Adam (1997): I don't believe in word senses. I: *Computers and the Humanities* 31/2, 91–113.
- Knudsen, Trygve & Alf Sommerfelt (1958): Principles of unilingual dictionary definitions. I: Eva Sivertsen et al. (red.): *Proceedings of the Eighth International Congress of Linguists*. Oslo: Universitetsforlaget, 92–101.
- Linguateca = <<http://www.linguateca.pt>> (februar 2014).
- Næss, Arne (1975): *En del elementære logiske emner*. [11. utgave] Oslo: Universitetsforlaget.
- Sveen, Andreas (1990): *Transitivitetsvekslinger i norsk: pseudointransitiver og ergativer*. Upublisert magistergradsavhandling, Universitetet i Oslo.

Jan Engh
universitetsbibliotekar
Universitetet i Oslo
Postboks 1009 Blindern
NO-0315 Oslo
jan.ENGH@ub.uio.no

Appifiering till allas lycka? Om danska ordboksappar med särskilt fokus på DDO

Louise Holmer & Emma Sköldberg

In recent years, most dictionary publishers have developed dictionary apps. However, in comparison with the comprehensive scientific study of printed and web-based dictionaries, conducted research on this kind of lexicographical resource is so far limited. This article discusses four monolingual dictionary apps developed by the Society for Danish Language and Literature. The focus is on the contemporary dictionary *The Danish Dictionary*. Different search functions are discussed, as well as how data is presented. The article also deals with what kinds of dictionaries are best suited for the app format. A key issue is whether this format really does render all dictionaries justice, and if app development is always advisable from a user's perspective.

1. Inledning

I takt med att antalet smarta telefoner och surfplattor ökat har utvecklingen av ordboksappar tagit fart. Enligt Gao (2013) och Rundell (2013) har ordböcker som är tillgängliga via mobiltelefonen stora fördelar i förhållande till sina analoga föregångare. I likhet med e-ordböcker har ordboksapparna i teorin obegränsat med utrymme. Vidare tillhandahåller apparna multimediala presentationssätt och de tillåter mer avancerade sökningar och länkning mellan ordboksartiklar och externa sidor. Apparna kan även uppdateras på ett enkelt sätt vilket är en fördel för både producenter och användare.

Gao (2013:213, 217) konstaterar att allt fler engelskspråkiga ordboks förlag har ”hoppat på appifieringståget” och samma utveckling

kan skönjas i Norden; Norstedts, Gyldendals och Kunnskapsforlaget har samtliga tagit fram ett flertal ordboksappar. Många appar från internationella ordboksförlag och producenter har också blivit mycket populära. Ett tydligt exempel är Dictionary.com, en engelskspråkig gratisapp, som har laddats ner mer än 60 miljoner gånger (Gao 2013:214). Vetenskapliga bidrag om ordboksappars innehåll och funktioner är emellertid fortfarande få till antalet. Några undantag utgör Holmer (2011), som recenserar ordboksappar från Norstedts, och Hoel (2012) som jämför appversionerna av *Retskrivningsordbogen* och *Svenska Akademiens ordlista*, SAOL. Se även Gao (2013), Rundell (2013), Svarverud (2014) och Køhler Simonsen (2014). Den sistnämnde presenterar resultaten av en mindre användarundersökning. Redogörelser av lexikograferna själva för ställningstaganden i samband med utveckling av ordboksappar är också få, vilket är en nackdel för dem som vill dra lärdomar av andras lexikografiska arbeten.

Ordboksredaktionen vid Institutionen för svenska språket, Göteborgs universitet (där artikelförfattarna är verksamma), har varit med och tagit fram SAOL-appen med, i skrivande stund, mer än en halv miljon nerladdningar. Just nu utarbetar vi, på uppdrag av Svenska Akademien, en app för den tryckta definitionsordboken *Svensk ordbok utgiven av Svenska Akademien* (SO 2009). Arbetet med SO-appen har väckt en hel del frågor som rör de specifika förutsättningar som gäller för ordboksappar i förhållande till tryckta ordböcker och nätversioner av desamma. En central fråga är om appformatet verkligen är lämpat för alla sorters ordböcker, t.ex. såväl moderna som historiska, och om det tillgodoser användarnas behov. Eller, som i vårt fall, hur ska vi kunna säkerställa att SO-appen blir en god lexikografisk produkt?

I samband med det redaktionella arbetet med SO-appen har vi studerat ett antal befintliga ordboksappar. I den här artikeln diskuteras hur appformatet fungerar i fyra danska ordböcker, *Den Danske Ordbog*, *Ordbog over det danske Sprog*, *Holberg-Ordbog* och

Moths Ordbog, dvs. fyra tämligen olikartade verk, där den förstnämnda är den som mest liknar SO. De aktuella ordboksapparna, som är gratis att ladda ner, är utvecklade av Det Danske Sprog- og Litteraturselskab (DSL) i samarbete med Dviance och Miracle A/S. Inför denna artikel är apparna i första hand testade på en iPhone 5 (iOS 6.1.4) och en Samsung Galaxy S4 (Android 4.3). Även iPad-versionen av ovan nämnda appar har studerats, i första hand för att ge jämförelsematerial. Testperioden löpte mellan mars och maj 2014. Testade versioner har varit 1.2.6 (iOS) och 1.1 (Android). Samtliga appar är onlineappar vilket innebär att den mobila enheten måste ha internetanslutning för att apparna ska fungera.

I det följande förhåller vi oss framför allt till apparnas allmänna egenskaper, sökfunktioner och hur (artikel)innehållet presenteras. *Den Danske Ordbog*-appen är mest avancerad och behandlas därför mer ingående. För att möjliggöra jämförelser mellan apparna och andra versioner av ordböckerna (där samtliga nätversioner är tillgängliga via ordnet.dk) inleder vi med en kort beskrivning av respektive förlaga.

2. Förlagor – i tryck och på nätet

Den Danske Ordbog (hädanefter DDO) utkom i sex band 2003–05. Denna samtidsordbok redogör för ordförrådet i modern danska från mitten av 1950-talet till idag. DDO är korpusbaserad och innehåller drygt 62 000 lemman. Ordboken är polyfunktionell och vänder sig till såväl modersmålstalare som inlärare (se vidare bl.a. Lorentzen & Trap-Jensen 2006). Nätversionen av DDO blev allmänt tillgänglig 2009 och den uppdateras fortlöpande (Trap-Jensen 2010, Lorentzen & Theilgaard 2012).

Ordbog over det danske Sprog (ODS) är en historisk ordbok som behandlar det danska språket från år 1700 till 1950. Verket gavs ut 1918–1956 i 28 band. Ett supplement på fem band har senare till-

kommit. ODS innehåller totalt ca 188 000 uppslagsord. Den finns på nätet sedan 2005. ODS vänder sig i första hand till alla som är intresserade av danska språket ur ett historiskt perspektiv (DSL 2014).

Holberg-Ordbog (Holberg) redogör för ordförrådet i arbeten av den dansk-norske författaren Ludvig Holberg (1684–1754). Arbetet med ordboken påbörjades redan på 1940-talet, men verket gavs ut först 1981–88. Ordboken, som innehåller drygt 26 000 lemmor fördelade på fem band, tillhandahåller utöver högfrekventa ord, betydelser och fraser i Holbergs verk också hänvisningar och citat. Ordboken ska dels fungera som hjälpmedel för läsare av Holbergs arbeten, dels tjäna som underlag för en beskrivning av språket i författarens tryckta verk. Den vänder sig i första hand till specialister, översättare och språkhistoriker (Eegholm-Pedersen 2006:147–148, 165). I nätversionen av Holberg är ordboks innehållet mer tillgängligt än i böckerna, bl.a. tack vare stavningshjälp och mer användarvänlig layout.

Moths Ordbog (Moth) är ett historiskt verk. Manuskriptet förfärdigades av ämbetsmannen Mathias Moth, som i slutet av 1600-talet påbörjade en insamling av danska ord i syfte att färdigställa en ordbok över det samtida danska språket. När Moth var färdig omfattade handskrifterna hela 9 000 folioark, men tyvärr kom de aldrig att tryckas. Ordboken är tvåspråkig med en större dansk-latinsk del och en mer ordlistartad latinsk-dansk del. Vidare är den indelad i en verbalordbok och ett lexikon, med bl.a. ortnamn och zoologiska ord (se Akhøj Nielsen 2014). Moth gjordes tillgänglig online 2013 och därefter följde den app som granskas här. De digitala utgåvorna omfattar dock bara den dansk-latinska delen. Nätversionen är i nuläget relativt enkelt uppbyggd men förbättras och utökas kontinuerligt.

3. De fyra apparna med fokus på DDO

I de följande diskuteras de fyra apparna mer ingående med DDO som utgångspunkt.

3.1. DDO-appen

DDO-appen är stilren; färgerna är snygga och layouten luftig. Startsidan har ett sökfält och ett antal ikoner, som i iPhoneversionen är försedda med rubrikerna *Home*, *Search*, *Favourites* och *Settings*.¹ Under *Favourites* listas uppslagsord som man valt att spara i kronologisk ordning. Det verkar inte finnas någon övre gräns för antal ord som kan favoritmarkeras. I enlighet med Hoel (2012:266–267) efterlyser vi även alfabetisk ordning. Listan blir annars snart svåröverskådlig. Under *Settings* finns bl.a. information om appens utgivare och utvecklare. Användarna kan också recensera produkten i App Store/Google Play. Vidare kan användarna föreslå nya uppslagsord, ge feedback till DDO-redaktionen och dela artiklar via mejl eller sms, vilket är positivt. Dock hade vi i samband med ingångssidan förväntat oss mer information om appen och om DDO. Tanken är nog att användarna ska gå från appen till ordnet.dk där mer information ges, men åtminstone vissa uppgifter borde få en framhävd position, inte minst sökinstruktioner.

Det mest iögonfallande på appens ingångssida är det lemma (eller den ordförbindelse) som lanseras som *Dagens ord/Dagens uttryk*. Den 1 april 2014, exempelvis, var dagens uttryck *drive gæk med* ’lave sjov med; holde for nar’. Extra information av encyklopedisk karaktär sätter in det utvalda ordet/uttrycket i ett sammanhang. Den aktuella dagen stod det: ”Tag dig i agt for aprilsnarren i dag!”. En viktig funktion hos dessa ord och uttryck torde

1 Flera rubriker är på engelska, vilket verkar ha med den mobila enhetens språkställning att göra. Om man laddar ner apparna utanför Danmark måste man i nuläget ändra hela enhetens språk till danska för att även få apprubriker på detta språk.

vara att ge ett smakprov på vad ordboken innehåller och locka till vidare användning. Orden och uttrycken är högaktuella (i stället för att bara slumpas fram) och de bidrar till att DDO-appen upplevs som en dynamisk produkt. Tack vare att appen bygger på en online-lösning kan redaktionen med kort varsel publicera ett ord eller uttryck som blivit aktuellt. Redaktionen ska ha en eloge för att de lyckas så väl med urvalet.

Enkel sökning utförs genom att man börjar skriva in sitt tänkta uppslagsord. Vartefter bokstäverna matas in sker en successiv filtrering av lemmalistan. Man kommer till artikelvisningen genom att klicka på det eftersökta ordet i lemmalistan eller skriva in hela ordet och klicka på *Search*. Lemmat visas överst i skrollistan följt av de 29 nästkommande orden. Maxgränsen för visade lemman är i nuläget alltså begränsad till 30, och det kan tyckas lite snålt. I de flesta andra appar, som t.ex. *Norstedts Engelska Ordbok Pro 2.0*, OED och SAOL, kan man dessutom skrolla både uppåt och neråt i lemmalistan. Denna funktion är nödvändig för dem som vill bilda sig en uppfattning om lemman som återfinns i anslutning till det eftersökta ordet.

I DDO-appen kan man söka på grundform och böjda former. Appen tillhandahåller också, som de flesta ordboksappar, stavningshjälp. Om man exempelvis vill slå upp substantivet *donor* ('donator') men skriver *donnor*, visas åtta förslag varav det första alternativet är korrekt, de sista lite mer långsökta. För att aktivera stavningshjälpen måste man klicka på *Search* eftersom det annars står "No result" under sökrutan. Detta meddelande är default om man söker på en sträng som inte ingår i lemmalistan.

DDO-appen stöder trunkerad sökning. Söker man t.ex. på *bog%* får man svaret att 98 lemman (bl.a. *bogaktuel*) matchar sökningen. Vidare kan man också söka på *%bog* och *b%g*. I appen visas de 10 översta träffarna av sökresultaten; vill användarna ta del av samtliga träffar får de gå via en länk till nätversionen. Den här typen av sökningar kan givetvis vara intressant för bl.a. korsords-

lösare. Funktionen hade med fördel kunnat nämnas i någon av de instruktionstexter som i nuläget alltså saknas. Ännu ett minus är att appen inte stöder samma jokertecken (wild cards) som nätversionen, dvs. * och ?.

I DDO-appen kan man också söka fram flerordskombinationer som utgör sublemman i artiklarna. Skriver användarna in ordförbindelsen *se rødt* ('blive rasende') kommer de direkt till idiomet. Men appen klarar betydligt mer avancerade sökningar än så. Om användarna är ute efter idiomet *binde på hænder og fødder* ('gøre fuldstændig afhængig') men bara skriver in substantiven i grundform (dvs. ”*hånd fod*”) får de ändå träff på uttrycket. Detta är imponerande.

Nedan ges ett exempel, lemmat *isvaffel*, på hur en artikel ser ut i DDO-appen.



Figur 1: Artikeln *isvaffel* i DDO-appen (iPhone).

Lemmat återges i rött, ljudfilsikonen i gult och övrig artikeltext i svart. Språkprovet är placerat i en ruta med grå bakgrund. Textstorleken förefaller lagom stor för att användarna ska träffa rätt när de touchar – även då de befinner sig i rörelse, vilket inte är ovanligt vid appanvändning (jfr s.k. ”fat finger syndrome”; se Køhler Simonsen 2014). Det vore emellertid en fördel om textstorleken kunde ändras (via *Settings* eller zoomning), allt för att tillmötesgå olika användare och olika brukssituationer.

När man skriver in ett ord i sökrutan och klickar på *Search* kommer man direkt till själva artikeln, såvida inte ordet ifråga är homograft. I det senare fallet sker en disambiguering genom en överskådlig sammanställning av aktuella artiklar överst på skärmen.

Artiklarnas innehåll presenteras på ett mycket tilltalande sätt. Indelningen i huvudbetydelser och underbetydelser framgår klart och tydligt och ordförbindelserna får generöst med utrymme. En hel del ord i artiklarna, bl.a. synonymer och antonymer, utgör länkar så att man kan klicka sig vidare till andra artiklar i ordboken, vilket är mycket användarvänligt.

Vad gäller visningsläge stöder iPad och Android såväl porträttvisning som landskapsvisning medan iPhone-versionen endast stöder porträttvisning. Att man kan vrida enheten för att utnyttja landskapsvisningsläget är klart gynnsamt vid läsning av längre artiklar.

DDO-appen återger bara delar av de uppgifter som finns i nätversionen av ordboken. Vill användarna se hela artiklarna kan de gå via länken till ordnet.dk. En relevant fråga är vilka informationskategorier redaktionen valt att ta med. I appen återfinns lemma, ordklass, böjning, kärnbetydelse, underbetydelse, synonym, antonym, ”se också”, exempelmeningar från korpus (citat), idiom och historiska uppgifter. Däremot saknas bl.a. nätversionens uttalsnotationer, grammatiska uppgifter, sammansättningar och avledning. En klar fördel med såväl nät- som appversionen i för-

hållande till den tryckta boken är det inlästa uttal som finns vid en majoritet av uppslagsorden. Själva appformatet och telefonernas gränssnitt passar bättre vid kortare artiklar. Längre artiklar kan vara svåra att överblicka, trots att bara valda delar av innehållet återges. Användarna får då skrolla en hel del för att nå länken till nätversionen, eftersom länken alltid ligger längst ner på sidan.

En tydlig skillnad mellan appen och pappers- och nätversionen av DDO är informationskategoriernas ordning i artiklarna. Som exempel kan placeringen av historiska upplysningar nämnas. I appen tycks de historiska uppgifterna betraktas som mycket relevanta eftersom de för det första har fått vara kvar när andra uppgifter har utelämnats och för det andra har en mycket framhävd position initialt i artikeln, i förhållande till den tryckta versionen där de placerats i slutet. Enligt vår bedömning är såväl daturval som informationskategoriernas inbördes ordning i appen lämpliga med tanke på ordbokens bredare målgrupp (jfr ovan). Till skillnad från ”Kort visning” av DDO på nätet, som visar en kortare version av den fulla ordboksartikeln och, som enligt Trap-Jensen (2010:1139) i första hand ska stödja reception, kan appen stödja flera funktioner, men kanske inte någon fullt ut. Det hade varit intressant att ta del av app-redaktionens lexikografiska tankegångar kring tänkta målgrupper och brukssituationer, men vi har inte funnit några sådana i skrift (jfr Trap-Jensen 2010). Vi efterlyser också fler användarundersökningar så att apputveckling blir mer förankrad i empiri (se t.ex. Holmer 2011; jfr dock Køhler Simonsen 2014).

Sammanfattningsvis är DDO-appen en lexikografisk produkt som imponerar på en rad olika sätt. Inte förvånande utsågs den till ”årets app” i kategorin Nyheter & Reference vid Danish App Awards 2013. Vid utvecklingen av appen torde redaktionen haft mycket stor nytta av allt det arbete som lagts ner på nätversionen av ordboken. Samtidigt finns det detaljer i appen som kan förbättras ytterligare; vi återkommer till dem i artikelns sista del.

3.2. ODS, Holberg och Moth i appformat

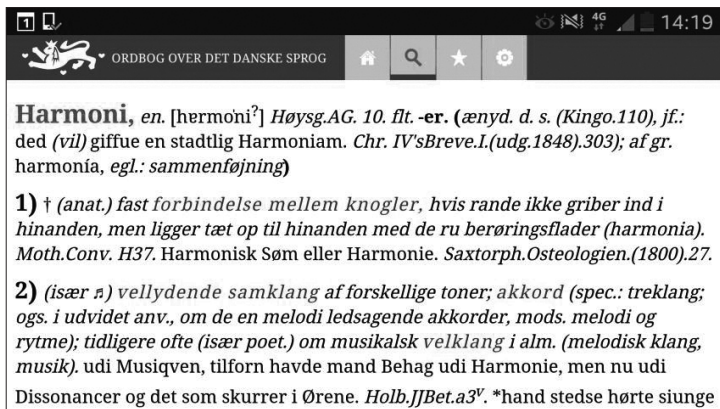
Startsidorna hos de historiska ordboksapparna är uppbyggda på snarlikt sätt som i DDO-appen. Färger och ikoner är desamma och favoritmarkering och delning av artiklar följer också DDO-mönstret. De historiska ordboksapparna innehåller dock kortfattad information om verken redan på ingångssidan. Holberg tillhandahåller även en kort text om samarbetet bakom ordboken med länkar till projektets hemsida. Informationstexterna fungerar tillfredsställande i ett första skede och sätter in apparna i sitt sammanhang. En annan skillnad är de fem exempellemman som, under rubriken *Tilfældige opslagsord*, växlar varje gång man laddar om sidorna. Till skillnad från *Dagens ord/Dagens udtryk* i DDO verkar dessa lemman helt slumpas fram och vissa av orden torde vara klart mer intressanta för användarna än andra. Moth har dessutom en fast lista på elva uppslagsord under rubriken *Redaktionen anbefaler*, t.ex. *Skole* och *Ny Jork*. Såväl de slumpvis valda som de rekommenderade orden fungerar som goda ingångar för i synnerhet nya användare som vill bekanta sig med ordböckernas innehåll.

Sökningen går till på liknande sätt som i DDO; vartefter man skriver filtreras förslagen, och även här är listan begränsad till 30 ord framåt i alfabetet. Stavningshjälpen fungerar överlag väl. Vad gäller normalisering av ålderdomliga stavningssätt verkar man med ODS ha kommit längst; där lotsas användarna från t.ex. *gå* till artikeln *gaa*. I Moth får man än så länge söka med äldre stavning.

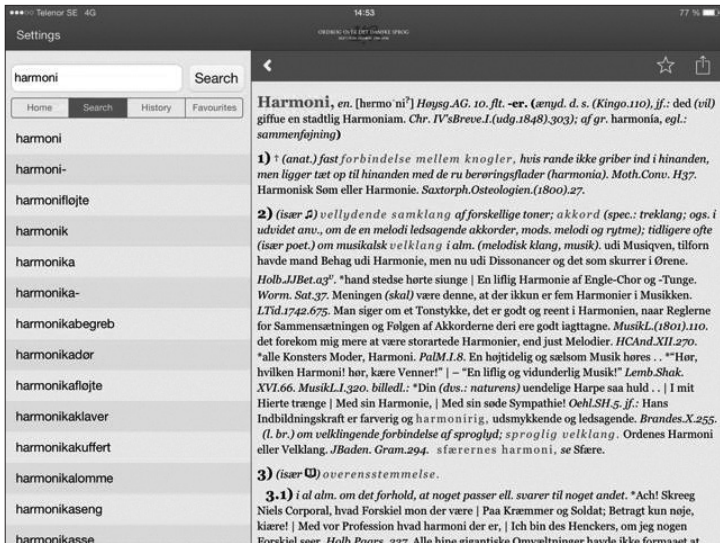
I ODS och Holberg kan man söka på alternativformer. Till skillnad från i de tryckta ordböckerna har sammansättningar fått tydligare lemmastatus och är därmed enklare att hitta. I de historiska apparna stöds också trunkerad sökning, men återigen noterar vi att denna möjlighet riskerar att förbli okänd för användarna eftersom söksinstruktioner saknas. Sökning på flerordsuttryck (av

slaget *se rødt*) inne i artiklarna fungerar däremot inte och man kan inte heller söka fram ordförbindelser genom att skriva in enskilda konstituenten (jfr *binde* på *hænder* og *fødder* ovan). Men den här typen av sökningar går heller inte att utföra i nätversionerna, så just dessa begränsningar har inte med själva appformatet att göra. Snarare kan de troligen kopplas till att de historiska grundmaterialen inte är fullt så digitalt uppmärkta som de data som ingår i DDO. Söker användarna på ord som inte finns upptagna i ordböckerna står det ”No result”. I nätversionen av Moth upplyses användarna däremot av texten ”Ordet findes i en anden ordbog. Slå ’xxx’ op i Ordbog over det danske Sprog.” Det vore mycket användbart för användarna om det förekom fler hänvisningar av detta slag mellan apparnas lemmalistor.

Av figur 2 och 3 framgår hur en artikel kan se ut i landskapsläge i de historiska ordboksapparna, dels i en telefon, dels i iPad. Exemplet gäller substantivartikeln *Harmoni* i ODS.



Figur 2: Artikeln *Harmoni* i ODS (något beskuren) (Android).



Figur 3: Artikeln *Harmoni* i ODS (nägot beskuren) (iPad).

I artiklarna är stickordet, precis som i DDO, rött och i övrigt är huvuddelen av artikeltexten svart. Teckenstorleken ligger fast. I nuläget finns det ingen hyperlänkning inom verken. Under artiklarna anges vilket band eller vilken ordboksdel som artiklarna är hämtade från. Längst ner finns, som i DDO, en länk till verkens nätversioner. Dessa tycks inte vara mobilanpassade, men apparna och dess länkar kan ändå fungera som en naturlig ingång till ordböckerna i sin helhet.

Gemensamt för de historiska apparna är att de ger ett mer komprimerat intryck, även i den lite generösare iPad-visningen. Den luftiga layout som kännetecknar DDO-appen saknas. Ordboksartiklarna återges också i sin helhet oavsett hur omfattande de är. Riktigt stora artiklar, t.ex. *gaa* med inte mindre än 40 huvudbetydelser i ODS, är i princip omöjliga att överblicka. De dubbla lodstreck som markerar inledning på en underbetydelse i såväl ODS som Holberg (i tryck och på nät) har dessutom fallit bort i

appversionerna, vilket ytterligare försvårar för användarna. Vidare har några av fördelarna med nätversionerna i förhållande till de tryckta förlagorna, t.ex. de lättillgängliga förkortningslistorna, tyvärr gått förlorade i apparna. ODS-appen innehåller ett flertal ikoner inne i artiklarna (se t.ex. not- och boksymbol i figur 2 och 3) men dessa förklaras heller inte. Tack vare länken mellan app och nätversion kan användarna förstås med lätthet gå in på nätsidan och det är utmärkt, för enligt vår mening gör trots allt nätsidorna – särskilt när dessa visas på stor skärm – de historiska ordböckerna större rättvisa.

4. Sammanfattande diskussion

Ordboksapparna är här för att stanna – åtminstone för ett bra tag framöver. Men frågan är om de kan betraktas som självständiga lexikografiska produkter. I den här artikeln granskas fyra ordboksappar som är utgivna av Det Danske Sprog- og Litteraturselskab.

DDO-appen, som enligt vårt förmenande är den mest genomarbetade av de fyra, har många förtjänster. Den är snygg och överskådlig, ger ett dynamiskt intryck och presenterar på ett intresseväckande sätt *Dagens ord/Dagens uttryk*. Vidare är sökmöjligheterna mycket goda, i synnerhet vad gäller ordförbindelser. DDO-appen bör kunna attrahera många olika användare, såväl spontan-användare som mer avancerade brukare. De informationskategorier som tagits med i appversionen stöder såväl reception som produktion. Enligt vår bedömning är informationen tillräcklig för att täcka grundläggande behov. Vill användarna ha mer kan de enkelt gå vidare till nätversionen av ordboken, som emellertid inte är mobilanpassad. En sympatisk detalj är den potentiella kommunikationen mellan användare och utvecklare. Vid icke-träff har användare möjlighet att sända in den aktuella söksträngen som förslag på nyord till redaktionen.

Det faktum att appen kräver nätåtkomst kan betraktas som en nackdel eftersom uppkopplingen kan vara långsam, instabil, icke-existerande eller dyr (Rundell 2013:5). Samtidigt har online-formatet mycket klara fördelar, inte minst med tanke på länken till nätversionen, vid uppdatering av innehållet och vid funktionen *Dagens ord*.

De historiska apparna är betydligt enklare uppbyggda. I dem återges ordboksartiklarna i sin helhet, oavsett hur omfattande de är. Dessa appar's största förtjänst är att de fungerar som reklamplare för verken *per se*, och det är inte att förakta. I synnerhet Moth har nu chansen att få den uppmärksamhet den så väl förtjänar via denna nya distributionskanal. Vill användare ta del av innehållet i de aktuella verken rekommenderar vi i dagsläget nätversionerna av ordböckerna. En önskan från vår sida är att apparna utvecklas så att de i högre grad kan fungera som självständiga produkter än vad de gör nu.

En bra app som DDO kan förstås bli ännu bättre. För det första efterlyser vi några korta rader om själva ordboken och om sökfunktioner i appen. För det andra skulle anpassning till röststyrning kunna gagna många användare, dvs. att de kan uttala det uppslagsord de är intresserade av och att ordet då visas på skärmen (se Køhler Simonsen 2014:428). Denna funktion stöds redan av många engelska ordboksappar (som t.ex. Merriam-Webster och Dictionary.com) och gynnar användare som snabbt och lätt vill slå upp ett ord, men som tycker att det är omständligt att knappa in det. Sådana sökmöjligheter verkar ännu ligga en bit bort i nordiska sammanhang. En mindre förändring rör t.ex. hanteringen av bl.a. svenska och isländska bokstäver i sökfunktionen. Sökning på **köbe* ger förslag som *koben* och *kæbe* med det eftersökta ordet, *købe*, på 13:e och sista plats. En enkel översättningsmekanism i programmet (ö = ø och liknande) skulle alltså kunna främja, i första hand kanske, nordiska användare. För det tredje skulle DDO-appen vinna på att fler ord i artiklarna vore hyperlänkar.

Också denna funktion finns redan i flera engelska ordboksappar. Det är emellertid inte ovanligt att mer avancerade appar från de stora internationella förlagen kostar pengar. Att DSL-apparna är gratis är oerhört positivt, och det ska naturligtvis vägas in när man tar funktionaliteten i beaktande.

Men ovanstående förslag handlar om framtiden. DSL visar redan nu på riktigt intressanta möjligheter för övriga lexikografiska aktörer i Norden. DDO-appen av idag är en mycket viktig förebild för dem som vill att ordböcker ska presenteras på ett modernt sätt som dessutom kan tilltala nya användargrupper.

Onekligen passar appformatet bättre för, som Køhler Simonsen (2014:424) uttrycker det, ”kognitivt set relativt ’lette opgaver’”. Det är en stor utmaning att presentera en historisk ordbok som ODS i portioner som användarna kan tillgodogöra sig. Køhler Simonsen (2014) har genomfört en mindre användarundersökning av appar och visar bl.a. att hans informanter föredrar att använda webbversionen av ett mer omfattande medicinskt uppslagsverk framför samma verk i appformat. Køhler Simonsens konklusion är att appar, precis som andra lexikografiska produkter, bör utvecklas med hänsyn till faktorerna: ”bruger, situation, tilgang, opgave, data og behov” (2014:428, fet stil i originalet). Vi är övertygade om att apputvecklingen skulle tjäna på att i högre grad vila på användarundersökningar än vad den gör idag.

Avslutningsvis hoppas vi att den här artikeln bidrar till en mer övergripande diskussion om ordboksappar och på vilka grunder de utarbetas. Appformatet tycks vara bättre lämpat för mer dynamiska samtidsordböcker än för omfattande, och ofta mer statiska, historiska verk. Samtidigt vill vi också lyfta frågan om vad alternativet till ordboksappar skulle vara – är det måhända mer mobilanpassade nätdböcker?

Litteratur

Ordböcker

DDO = *Den Danske Ordbog* (2003–2005). København: Det Danske Sprog- og Litteraturselskab & Gyldendal. Tillgänglig via <<http://ordnet.dk/ddo>>. Appversioner (2014): 1.2.5 (iOS) och 1.1-13 (Android). © Det Danske Sprog- og Litteraturselskab. Utvecklade av Dviance och Miracle A/S.

Dictionary.com (2014). Version 4.4 (iPhone) © Dictionary.com, LLC.

Holberg = *Holberg-Ordbog. Ordbog over Ludvig Holbergs Sprog. I–V* (1981–1988). København: Det Danske Sprog- og Litteraturselskab. Tillgänglig via <<http://holbergordbog.dk/>>. Appversioner (2014): 1.2.4 (iOS) och 1.1 (Android). © Det Danske Sprog- og Litteraturselskab. Utvecklade av Dviance och Miracle A/S.

Merriam-Webster Dictionary (2014). Version 3.0.3. (Android) Merriam-Webster, Incorporated.

Moth = *Moths Ordbog*. <www.mothsordbog.dk> (maj 2014). Appversioner (2014): 1.2.4 (iOS) och 1.1 (Android). © Det Danske Sprog- og Litteraturselskab. Utvecklade av Dviance och Miracle A/S.

Norstedts engelska ordbok Pro 2.0 (2014). Version 2.0.0. (iPhone).

ODS = *Ordbog over det danske Sprog*. Udg. af Det Danske Sprog- og Litteraturselskab. I–XXVIII (1919–1956). Supplement I–V (1991–2005). København: Gyldendal. Tillgänglig via <<http://ordnet.dk/ods>>. Appversioner (2014): 1.2.4 (iOS) och 1.1-13 (Android). © Det Danske Sprog- og Litteraturselskab. Utvecklade av Dviance och Miracle A/S.

OED = *Oxford English Dictionary* (2014). © Oxford University Press, 2010, 2012. Version 4.3.069 (Android).

ordnet.dk = <<http://ordnet.dk/>> (maj 2014).

- SAOL = *Svenska Akademiens ordlista* (2014). © Svenska Akademien & Isolve AB. Version 1.1.3 (Android).
- SO = *Svensk ordbok utgiven av Svenska Akademien* (2009). Stockholm: Norstedts (i distribution).

Annan litteratur

- Akhøj Nielsen, Marita (2014): Moths ordbog og dens virkningshistorie. I: *LexicoNordica* 21 (detta nummer).
- DSL (2014): Baggrund og tilblivelse. Om tilblivelsen av ODS. <<http://ordnet.dk/ods/fakta-om-ods>> (maj 2014).
- Eegholm-Pedersen, Svend (2006): Holbergordbogen – en nordisk forfatterskabsordbog. I: *LexicoNordica* 13, 147–168.
- Gao, Yongwei (2013): The Appification of Dictionaries: From a Chinese Perspective. I: I. Kosem et al. (eds.): *Electronic lexicography in the 21st century: thinking outside the paper. Proceedings of the eLex 2013 conference*. Tallinn: Trojina, Institute for Applied Slovene Studies/Eesti Keele Instituut, 213–224.
- Hoel, Jan (2012): Appsolutt fingerferdig! En anmeldelse av ordbok-appene RO og SAOL. I: *LexicoNordica* 19, 255–271.
- Holmer, Louise (2011): Norstedts ordboksapp. I: *LexicoNordica* 18, 307–322.
- Køhler Simonsen, Henrik (2014): Brukerne er allerede mobile! I: R. Vatvedt Fjeld & M. Hovdenak (red.): *Nordiske studier i leksikografi* 12. Oslo: Novus, 416–429.
- Lorentzen, Henrik & Liisa Theilgaard (2012): Online dictionaries – how do users find them and what do they do once they have? I: R. Vatvedt Fjeld & J. M. Torjusen (eds.): *Proceedings of the 15th EURALEX International Congress*. Oslo: University of Oslo, 654–660.
- Lorentzen, Henrik & Lars Trap-Jensen (2006): Den Danske Ordbog: en ordbog for eksperter, godtfolk och almindeligt interesserede. I: *LexicoNordica* 13, 197–208.

- Rundell, Michael (2013): Redefining the dictionary: From print to digital. I: *Kernerman Dictionary News* 21. <<http://kictionaries.com/kdn/kdn21.pdf>> (maj 2014).
- Svarverud, Rune (2014): Nye kvalitetsverktøy for brukere av kinesisk i Skandinavia. I: *LexicoNordica* 21 (detta nummer).
- Trap-Jensen, Lars (2010): One, Two, Many: Customization and User Profiles in Internet Dictionaries. I: A. Dykstra & T. Schoonheim (eds.): *Proceedings of the XIV Euralex International Congress*. Leeuwarden: Fryske Akademy, 1133–1143.

Louise Holmer
biträdande forskare, doktorand
Lexikaliska institutet
Institutionen för svenska språket
Box 200, Göteborgs universitet
SE-405 30 Göteborg
louise.holmer@svenska.gu.se

Emma Sköldberg
universitetslektor, docent
Lexikaliska institutet
Institutionen för svenska språket
Box 200, Göteborgs universitet
SE-405 30 Göteborg
emma.skoeldberg@svenska.gu.se

RECENSIONER

Bloomsburys vägledning till lexikografin

Anna Helga Hannesdóttir

Howard Jackson (ed.): *The Bloomsbury Companion to Lexicography*. London, New Delhi, New York & Sydney: Bloomsbury 2013. 420 sidor. Pris: 190 USD för den tryckta boken, 155 USD för en digital version.

1. Inledning

The Bloomsbury Companion är den serie akademiska handböcker som förlaget Bloomsbury Publishing nu ger ut. På det språkvetenskapliga området har bland annat syntax, fonologi och fonetik behandlats. Nyligen kom så *The Bloomsbury Companion to Lexicography* (2013); en antologi över det lexikografiska forskningsfältet. De senaste åren har flera sådana verk publicerats, ett par av dem har också recenserats i tidigare nummer av *LexicoNordica* (Grønvik 2012, Eide 2013). Volymen innehåller bidrag från ett tjugotal forskare varav flera också medverkar i en del av de tidigare antologierna. En fråga som därför inställer sig redan vid en första hastig genombläddring är om detta arbete främst kompletterar eller överlappar de tidigare verken.

Boken marknadsförs på förlagets hemsida som ”the definitive guide to a key area of linguistic study” och som garanti för kvaliteten citeras recensioner av två lingvister. Det aktuella fältet preciseras som lexikografi och bokens syfte anges vara: ”to reflect on the research that has been and is being done in lexicography and to point the way forward”. Samma information ges i baksidestexten.

I det följande presenteras bokens uppläggning och innehåll. De enskilda kapitlen behandlas olika utförligt; allt innehåll ägnas inte

lika stor uppmärksamhet. Avslutningsvis diskuteras vissa centrala frågeställningar samt det helhetsintryck som verket ger av lexikografien.

2. Bokens innehåll och uppläggning

Volymens utanförtexter består av sedvanlig förlagsinformation, innehållsförteckning, en presentation (s. vii–xiii) av de 23 forskare som bidragit till volymen samt ett avslutande sakregister. Innehållstextens åtta tematiska och numrerade delar bärs upp av totalt 21 kapitel, olika många i de olika delarna. Del 1 utgörs av en introduktion av bokens redaktör, Howard Jackson, och del 2 av en översikt av Paul Bogaards över den lexikografiska forskningens historia. Tyngdpunkten ligger i delarna 3–5. I del 3 diskuteras *Research Methods and Problems* i tre kapitel, del 4 samlar 10 kapitel på temat *Current Research and Issues* och del 5, *New Directions in Lexicography*, innehåller tre kapitel. De avslutande delarna, 6–8, består av ett kapitel var: del 6, *Resources* av Reinhard Hartmann, del 7, *Glossary of Lexicographic Terms* av Barbara Ann Kipfer och del 8, *Annotated Bibliography* sammanställd av bokens redaktör.

Kapitlen numreras inte löpande i volymen utan separat inom varje enskild del; en bagatell som faktiskt bidrar till ett något rörigt intryck. Ett stort plus är däremot att varje kapitel inleds med en innehållsförteckning i form av en ruta där underrubrikerna listas. De olika kapitlen är fristående från varandra på så vis att noterna återfinns sist i kapitlet liksom bibliografin över all litteratur som åberopas.

3. Del 1: Arbetets utgångspunkter

Redaktörens introduktion inleds med en begreppsutredning

beträffande den praktiska och den teoretiska sidan av fältet och relationen mellan dem. För den teoretiskt inriktade lexikografin används termen *metalexikografi* och i boken är det denna aspekt snarare än den praktiska som står i fokus. Som målgrupp för boken utpekas främst studenter, närmare bestämt blivande metalexikografer, men även andra som har ett teoretiskt intresse för ordböcker och ordboksproduktion.

De två övergripande frågor som diskuteras bland lexikografer och metalexikografer skisseras också: lexikografins akademiska hemvist och lexikografins akademiska status generellt samt frågan om själva existensen av en lexikografisk teori. Där varierar åsikterna från att det inte finns någon sådan alls till att all lexikografisk verksamhet är tydligt teoretiskt förankrad.

Efter den korta introduktionen till det lexikografiska fältet följer en översikt över bokens innehåll. De olika kapitlen refereras kortfattat men ändå tillräckligt detaljerat för att ge läsaren en tydlig bild av innehållet (s. 3–18). Introduktionskapitlet fungerar därigenom som en utmärkt översikt för den som vill bilda sig en uppfattning om forskningsfältet lexikografi såsom det presenteras i volymen.

4. Del 2: En brett anlagd historisk översikt

I bokens del 2 skissar Paul Bogaards den västerländska lexikografiska forskningens etablering och utveckling i kapitlet *A History of Research in Lexicography*. Som redaktör för *International Journal of Lexicography* under tio år fick han god insyn i och överblick över den pågående forskningen inom fältet. Hans resumé sträcker sig dock över längre tid än bara dessa tio år.

Bogaards belyser kort några av de fält som den lexikografiska forskningen omfattar: lexikografihistoria, ordbokskritik, ordbokstypologi, ordböckernas struktur, ordboksanvändning och inne-

hållet i ordböcker. Några av dessa områden återkommer och utvecklas i flera av volymens kapitel. Han återger de kriterier som vanligen ligger till grund för en typologisk indelning av ordböcker och påpekar att även distinktionen mellan akademiska och praktiska ordböcker borde uppmärksammas mer liksom den mellan deskriptiva och preskriptiva. När det gäller ordbokskritik slår han fast att de recensioner som publiceras i dagstidningar vanligen stannar vid ordbokens makrostruktur och sålunda främst är intressanta för presumtiva köpare. De vetenskapliga recensionerna är betydligt spretigare och präglade av den enskilda recensentens specialitet. Och trots att det snarast är kutym att nya ordböcker recenserar påpekar han att den kritik som framförs sällan lämnar avtryck i senare ordböcker.

Här framstår recensionernas tänkta mottagare som ett avgörande kriterium och distinktionerna akademisk vs praktisk och deskriptiv vs preskriptiv högst relevanta. Det är rimligt att en recension av ordboken som bruksföremål främst har karaktären av konsumentupplysning medan den sakkunniges kritiska granskning som vänder sig till ordboksredaktionen snarare beaktar andra aspekter. Området uppmärksammades tidigt i *LexicoNordica* och avsnittet aktualiserar en fråga som ställdes i temanumret om ordbokskritik: ”Finnes det en ordbokskritikk for redaksjonssjefer?” (Eek 2003). Ordbokskritik behandlas vidare i volymens kapitel 3.2.

Avslutningsvis berör Bogaards frågan om lexikografins teoretiska förankring. Han lyfter fram avsaknaden av en enda, heltäckande lexikografisk teori men ställer sig samtidigt tveksam till om en sådan teori behövs och undrar hur den i så fall skulle kunna utformas.

Bogaards del i volymen kompletterar Jacksons introduktion och tillsammans utgör de två första delarna en belysande och lättillgänglig introduktion till metalexikografen.

5. Del 3: Metoder och problem

Bokens tredje del samlar tre kapitel på temat ”Research Methods and Problems”.

Datoriseringens betydelse för ordboksarbetet är temat för Lars Trap-Jensens kapitel ”Researching Lexicographical Practice”. För varje enskilt led i arbetsprocessen belyser han datoriseringens inverkan: för planeringen och uppläggnings av ett ordboksprojekt, databasens design, den lexikografiska beskrivningen, valet av redigeringsystem, åtkomststruktur och presentation av lexikografiska data samt, slutligen, *var* och *hur* samt *i vilken form* användarna i framtiden kommer att hitta ordböcker och andra lexikografiska produkter.

Trap-Jensens kapitel spänner sålunda över vida fält. Han beskriver den radikala förändring som det innebär för lexikografen att inte ha siktet inställt på en konkret ordbok utan istället arbeta i databaser avsedda att ligga till grund för olika lexikografiska produkter, publicerade i olika plattformar och olika medier. Korpusarnas betydelse för ordboksarbetet betonas, liksom vikten av en medveten och kritisk inställning till relationen mellan korpuslingvistik och lexikografi. Korpusbeläggen måste analyseras, gallras, kompletteras, redigeras – kort sagt underkastas mänsklig hantering för att kunna tjäna lexikografen. Andra teman som lyfts är ordböckernas språksociala funktioner, språkpolitiska aspekter på lemmaselektionen och den lexikografiska beskrivningen samt användarmedverkan. Slutligen sneglar Trap-Jensen mot framtiden och den konkurrent som lexikografiska produkter av alla de slag har i Google. Frågan om hur lexikografen kommer att kunna hävda sig i den kampen lämnar han öppen.

I kapitlet ”Methods in Dictionary Criticism” återkommer diskussionen om ordbokskritik. Här påpekar Kaoru Akasu att det inte rått någon konsensus om vilka av ordbokens egenskaper som bedöms och enligt vilka kriterier. Att varken ordboksredaktörer

eller metalexikografer har tagit genren ordbokskritik på allvar menar han bl.a. framgår av att bara en av de omkring 330 artiklarna i HSK (1989–1991) har en rubrik som innehåller ”dictionary criticism”. Akasu presenterar den japanska metod för ”dictionary analysis”, som han varit med om att utveckla och använda. Vissa bestämda komponenter i ordboken analyseras av ett team av granskare med olika specialiteter enligt vissa fastställda bedömningskriterier, kvalitativa såväl som kvantitativa. Han menar att även om metoden behöver utvecklas ytterligare så har tillämpningen av den gagnat såväl japanska konsumenter som lexikografer. Bland annat underlättar den jämförelsen mellan olika ordböcker och gör det därmed lättare att välja den ordbok som bäst motsvarar den enskilde konsumentens krav.

När det gäller ordboksmarknaden i de nordiska länderna är läget naturligtvis ett helt annat än i Japan. Förutom de kollegiala granskningarna, som huvudsakligen publiceras i *LexicoNordica*, förekommer också recensioner av nya enspråkiga ordböcker i dagstidningarna. Dessa vänder sig då vanligen till allmänheten. Att tvåspråkiga ordböcker uppmärksammas sker däremot sällan. I små språksamhällen som de nordiska är ordbokskritik ett mycket smalt metalexikografiskt fält.

I den sista artikeln i kapitel 3, ”Researching Users and Uses of Dictionaries”, avhandlar Hilary Nesi studier i ordboksanvändning. Sedan 1980-talet har drygt 300 studier rapporterats, i princip alla med syftet att underlätta för ordboksanvändarna att hitta den information de söker. De olika metoder som används i studierna presenteras liksom de för- och nackdelar som är förknippade med de olika metoderna. Nesi konstaterar att studiernas vanligen beskedliga omfång och varierande utformning gör att det är svårt att dra generella slutsatser av de redovisade resultaten. Hon ställer också frågan om resultaten av användarundersökningarna har satt några märkbara spår i de stora ordboksförlagens produkter.

Den forskning Nesi diskuterar avser främst tryckta ordböcker

och digitala versioner av sådana. När det gäller nätordböcker konstaterar hon att användningen av dessa fortfarande till stor del är outforskad och att området är svårt att undersöka eftersom innehållet i nätordböcker inte är konstant. Avslutningsvis berör hon hur de tekniska framstegen har lett till att även de etablerade förlagens ordboksredaktioner krymper och att de lexikografiska produkter som lanseras på nätet i allt större utsträckning bygger på datateknikens och inte lexikografins utveckling.

I dessa tre kapitel presenteras intressanta perspektiv på ordboksarbetet, ordbokskritiken och användningen av ordböcker. Någon heltäckande bild av de stora utmaningar som lexikografen står inför kan de naturligtvis inte ge och vissa av problemen gäller det som varit snarare än den framtida utvecklingen. Men de lyfter på ett förtjänstfullt sätt fram ordböckernas roll i och samspel med det omgivande samhället. Den bild som tecknas av de stora ordboks-förlagen och -redaktionerna antyder dock en viss tröghet när det gäller att beakta kritikernas och användarnas synpunkter.

6. Del 4: Aktuell forskning

De tio kapitlen i bokens fjärde del har försetts med rubriken ”Current Research and Issues”. Först ut är Adam Kilgarriffs ”Using Corpora as Data Sources for Dictionaries”. Han konstaterar att fokus har flyttats från de två uråldriga lexikografiska metoderna plagiat och introspektion till det faktiska språkbruket såsom det dokumenteras i korpusar.

Kilgarriff är en auktoritet när det gäller korpusar, korpusredskap och konkordansverktyg avsedda för lexikografiskt arbete. Han beskriver hur korpusar och avancerade sökverktyg kan användas inte bara som källor utan också som redskap för lemmaselektion, för att urskilja eller belägga mönster i ordens kombinatoriska egenskaper och preferenser, som underlag för bruksmarkörer

samt för att hitta utgångspunkt för illustrativa och effektiva exempel på lemmats användning – kort sagt: det mesta av det lexikografiska arbetet. Till skillnad från Trap-Jensen tonar han ned lexikografens roll och menar att den alltmer kan begränsas till att kontrollera, bekräfta eller redigera de förslag som korpusredskapet genererar. Utgångspunkten för Kilgarriff är de produkter han själv har varit med och utvecklat, Sketch Engine och de ”word sketches” som den genererar.

Han diskuterar också parallellkorpusar och deras användbarhet inom den tvåspråkiga lexikografen. När det gäller de mindre språken är tillgången till sådana dock begränsad. Google Translate, som tillämpar statistisk maskinöversättning, framhålls som ett bra exempel på en lyckad kombination av språkteknologi och korpuslingvistik.

Ordboksanvändning återkommer som tema för kapitel 4.2, ”Researching the Use of Electronic Dictionaries” av Verónica Pastor och Amparo Alcina. Till skillnad från Nesi har Pastor och Alcina studerat elektroniska ordböcker. De konstaterar att de i sammanhanget två största problemen med dessa består å ena sidan i att de inte är särskilt användarvänliga och å andra sidan i att användarna inte är så bra på att använda ordböcker. En del tillkortakommanden kan tillskrivas diskrepansen mellan de avancerade sökmöjligheter som de elektroniska ordböckerna erbjuder och det faktum att många av dem bygger direkt på en tryckt ordbok som enbart är avsedd för sökning på lemman. Inom ramen för projektet ONTODIC har Pastor och Alcina studerat de söktekniker som används i elektroniska ordböcker, i korpusar och på nätet. Deras forskning har bl.a. resulterat i den klassificering av söktekniker som presenteras i kapitlet och som utgår ifrån de tre element som ingår i varje sökning: *frågan*, *resursen* och *resultatet*. Dessa element definieras, beskrivs och illustreras vidare med hjälp av drygt 20 skärmbilder av olika kvalitet. Resultatet av den jämförande analysen av drygt 30 elektroniska ordböcker redovisas i en bilaga.

Det tredje kapitlet i del 4 är skrivet av John Considine: "Researching Historical Lexicography and Etymology". Han är pessimistisk när det gäller framtiden främst för den etymologiska lexikografen. För den historiska lexikografen innebär däremot utvecklingen av korpusar med äldre texter av olika slag att källäget för historiska ordböcker nu har ändrats dramatiskt. Finansieringen framstår dock som ett reellt problem för båda dessa former av lexikografi.

Två kapitel i del 4, "Researching Pedagogical Lexicography" av Ami Chi och "Monolingual Learners' Dictionaries – Where Now?" av Shigeru Yamada, ägnas åt forskningen kring den "pedagogiska" lexikografen, dvs. enspråkiga ordböcker avsedda för användare med annat modersmål än det som behandlas i ordboken. I sitt kapitel tar Chi bland annat upp användaraspekten. Liksom Nesi påpekar hon glappet dels mellan lexikografernas intentioner och användarnas ordbokskompetens, dels mellan forskarnas resultat och förlagens intresse för dessa när det gäller ordböckernas utformning och innehåll. Här framträder tydligt hur olika villkoren är för lexikografen i stora språksamhällen med flera konkurrerande ordboksförlag och små språksamhällen som de nordiska. Yamada beskriver den utveckling som inlärorordböckerna har genomgått från 1940-talet och framåt. Beträffande de digitala ordböckerna betonar han att det obegränsade utrymme som lexikografen där har till sitt förfogande och möjligheten att bygga ut artiklarna med nya informationskategorier ställer stora krav på en ytterst tydlig mikrostruktur och på att resultatvisningen anpassas till användarens behov.

En artikel i volymen ägnas helt åt den tvåspråkiga lexikografen: "Issues in Compiling Bilingual Dictionaries" av Arleta Adamska-Salaciak. Hon inleder sitt kapitel med en översikt över den tvåspråkiga lexikografins centrala begrepp: källspråk respektive målspråk, vilket av dessa som är den tilltänkta målgruppens modersmål, dvs. om ordboken ska fungera som L1→L2 och vara

avsedd för produktion av något slag på det främmande språket, eller om den är en L2→L1 ordbok, avsedd för avkodning. Hon begränsar perspektivet till att bara behandla tryckta ordböcker och hänvisar till att elektroniska ordböcker behandlas på andra håll i volymen. Den utmaning som det digitala formatet innebär just för den tvåspråkiga lexikografen förbigår hon helt. De frågor som Adamska-Salaciak behandlar har förvisso varit grundläggande inom den tvåspråkiga lexikografen hittills men de hör till ett skede som snart är förbi.

I Danie J. Prinsloos kapitel, ”Issues in Compiling Dictionaries for African Languages”, behandlas några av de många, svåra problem som den lexikografiska beskrivningen av bantuspråken ställs inför. Bantuspråken är morfologiskt komplicerade vilket försvårar inte minst lemmatiseringen. Prinsloo redovisar de strategier som etablerats. De afrikanska ordböckerna är fortfarande huvudsakligen tryckta och effekten av den utveckling som korpusar och digital publicering innebär är ännu inte särskilt märkbar.

Ett annat lite smalt tema som beretts plats i volymen är teckenspråkslexikografi. Trots att denna gren av lexikografen inte är helt ny – i kapitlet omnämns en gloslista från slutet av 1700-talet – är den dåligt utvecklad. I kapitlet ”Issues in Sign Language Lexicography” ger Inge Zwitterlood, Jette Hedegaard Kristoffersen och Thomas Troelsgård en statusbeskrivning från fältet. Behovet av ordböcker är stort. Enspråkiga teckenordböcker finns inte; de är antingen från det talade språket till teckenspråk eller omvänt. Inte heller finns ordböcker mellan två teckenspråk. De flesta ordböcker som finns är små och omfattar avgränsade ämnesområden som exempelvis juridik och sjukvård.

Forskarna beskriver arbetet med att utveckla den ortografiska representationen av tecknens fonologiska struktur och en adekvat notation för andra typer av distinktiva drag. Även denna typ av lexikografi ändras fort när nya medier tas i anspråk. Fotografier och filmsekvenser har utnyttjats och nu även andra multimediala

presentationssätt. Ett problem som fortfarande är stort för denna lexikografi är bristen på korpusar med teckentexter. Kapitlet är rikt illustrerat vilket underlättar för den icke insatte läsaren att hänga med i diskussionen.

Medan de tidigare kapitlen i del 4 sakligt har rapporterat (forsknings)läget inom det aktuella området kan de två avslutande kapitlen ses som debattinlägg i en diskussion av övergripande teoretisk karaktär. I ”Identifying, Ordering and Defining Senses” reder Robert Lew ut relationen mellan betydelsebeskrivningen i ordböcker och den semantiska beskrivningen inom lingvistik. Medan man inom lingvistik beskriver språkets lexikala enheter går ordbokens betydelsebeskrivning ut på att urskilja och avgränsa olika betydelser och betydelsemoment för att inlemma dem i ordbokens artikelstruktur. Hur dessa strukturella element ordnas i relation till varandra är underkastat ordbokens syfte: att presentera det semantiska innehållet på ett enkelt och för lekmannen överskådligt sätt. Lew behandlar inte bara enspråkiga ordböcker utan även tvåspråkiga, där bl.a. indelningen i betydelsemoment kompliceras av att den involverar två lexikala system. Hans budskap kan sammanfattas som så, att alla sätt att i ordböcker beskriva och redovisa betydelse är lika riktiga och inget sätt är mera motiverat än ett annat ur en lingvistisk semantisk synvinkel. Den redovisning som väljs kan tillskrivas förlagets etablerade traditioner snarare än objektiva kvalitetskriterier.

Lew ger sig inte in i diskussionen om lexikografins teoretiska förankring. Det gör däremot Tadeus Piotrowski i fjärde delens sista kapitel: ”A Theory of Lexicography – Is There One?”. Kapitlet kan ses som ett debattinlägg i två olika diskussioner, dels den som anges i rubriken, dels diskussionen om lexikografins relation till informationsvetenskapen. När det gäller huruvida det finns en lexikografisk teori (eller flera) tar han sin utgångspunkt i den i vissa kretsar etablerade uppfattningen att hantverk inte är teoribaserat. Han drar en parallell mellan lexikografi och översättning: båda ak-

tiviteterna inbegriper hantverk vars resultat är rent språkliga produkter samtidigt som de också har en teoretisk dimension. Han menar vidare att den skepsis mot lexikografins teoribildning som ibland framförs snarast ligger i de humanistiska teoriernas allmänt svaga validitet. De är inte prediktiva utan generaliserar framtida skeenden utifrån redan pågående eller avslutade skeenden. Han konstaterar att lexikografen nu står inför ett paradigmskifte och att vi därför inte längre kan utgå ifrån att den i allt väsentligt kommer att fortsätta vara en direkt utveckling av det den hittills har varit. Resonemanget är spännande.

Piotrowski lyfter fram lexikografiteorins tre komponenter: en syntaktisk, dvs. ordbokens abstrakta informationsstruktur, en semantisk, dvs. det innehåll som ordnas i denna struktur samt en pragmatisk, dvs. ordbokens målgrupp. Att ge den pragmatiska aspekten företräde, såsom förespråkas inom funktionsteorin, är han tveksam till. De tre komponenterna ger istället lexikografen en unik ställning bland referensvetenskaperna. Hans inlägg är uppfriskande och debatten lär fortsätta.

De nedslag i fältet som belyses i de enskilda kapitlen i del 4 bidrar förvisso med en rad intressanta infallsvinklar på avancerade och spännande problem. Men det är Lews och Piotrowskis kapitel som lyfter hela denna del.

7. Del 5: Framtiden

Volymens femte del, "New Directions in Lexicography" innehåller tre kapitel. Pedro A. Fuertes-Olivera redogör för olika typer av online-ordböcker i kapitlet "e-lexicography: The Continuing Challenge of Applying New Technology to Dictionary-Making". Han definierar termen *e-lexikografi* som den vetenskapliga disciplin inom informationsvetenskapen som huvudsakligen är inriktad på utveckling, planering, sammanställning och publicering av elek-

troniska referensverktyg. I kapitlet lyfter han fram den intrikata relationen mellan teknologin, användarna, lexikografiska framsteg och ekonomiska aspekter. Han diskuterar spannet av ordböcker som finns på nätet, dvs. å ena sidan ordböcker som i alla avseenden utom publikationsformen är identiska med en tryckt förlaga och å andra sidan de individuellt anpassade referensverktyg som utvecklats med utgångspunkt i Århuskolans version av funktionsteorin. De största hindren för utvecklingen av effektiva digitala informationsredskap är inte i första hand tekniska utan snarare av institutionell, forskningsrelaterad samt intellektuell och social natur.

Framtidskapitel nummer två handlar om historisk lexikografi: "The Future of Historical Dictionaries, with Special Reference to the Online OED and Thesaurus". Där belyser Charlotte Brewer några av de paradoxala problem som redaktionerna för de mastodontordböcker som koncipierades på 1800-talet har att hantera i relation till nya, digitaliserade upplagor. Ett typiskt problem är hur exempelvis föränderliga moraliska och etiska aspekter ska hanteras. Hon illustrerar problemet bl.a. med ord för 'homosexualitet' där *unnatural* förekommer i definitionen. I sådana sammanhang aktualiseras en rad viktiga frågor såsom de vetenskapliga kraven på akribi, risken för språkhistorisk och/eller kulturhistorisk censur och allmän lexikografisk klåfingrighet. Relationen mellan en eller flera tryckta upplagor och en föränderlig online-upplaga framstår som ett i allra högsta grad intrikat problem. Här går tankarna osökt till SAOB, *Ordbok öfver svenska språket, utgiven av Svenska Akademien* (1893–), som inom några år kommer att vara komplett. En rad av de problem som Brewer beskriver kommer SAOB-redaktionen då att ställas inför. Hur genomgripande revidering och komplettering det kan bli frågan om för SAOB är ännu inte avgjort. Problemet med justerade lexikografiska principer och daterad moral- och människosyn i SAOB har diskuterats av Larsson (2009:35ff.).

Även det sista kapitlet i del 5, "The Future of Dictionaries,

Dictionaries of the Future”, är inriktat på det paradigmskifte för lexikografen som anslutningen till informationsvetenskapen skulle innebära. Sandro Nielsen ser lexikografiska produkter som bestående av många ytliga och tre underliggande strukturer: lexikografiska funktioner, data som stöder dessa funktioner samt en lexikografisk struktur som länkar data. Han berör delvis samma frågor som Trap-Jensen. Bland annat aktualiserar han relationen mellan en databas och de många lexikografiska produkter som kan genereras ur denna. Han för även in användaranpassningen och möjligheten att skraddarsy de lexikografiska produkterna till den specifika användaren och dennes vid varje tillfälle specifika behov.

Att de framtida utmaningar som lexikografen står inför är svåra och viktiga framgår tydligt i dessa tre kapitel. Det hävdvunnna lexikografiska hantverk som har utövats i århundraden måste i vissa avseenden ändras i grunden för att de möjligheter som de nya publikationsformerna och redskapen erbjuder skall kunna utnyttjas fullt ut. Det framstår klart att lexikografen inte kommer att kunna utvecklas vidare utan ett mycket nära samarbete med de språkteknologiska miljöerna. Men frågan är också hur långt språkteknologins lexikaliska resurser kan utvecklas utan de traditionella lexikografiska verktygen för den språkliga analysen.

8. Volymens avslutande delar

De avslutande delarna i boken innehåller översikter av olika slag. I del 6 rapporterar Reinhard Hartmann från sitt arbete med att kartlägga de aktörer som verkar inom det lexikografiska fältet och de sammanslutningar och fora som har etablerats på området. Del 7 innehåller Barbara Ann Kipfers förteckning över omkring 200 lexikografiska termer med förklaringar, del 8 en kommenterad, tematiskt indelad bibliografi över metalexikografisk litteratur, sammanställd av Howard Jackson, följd av det avslutande sakregistret.

9. Kommentarer och reflektioner

Till det yttre är *The Bloomsbury Companion to Lexicography* en tilltalande volym med hårda pärmar och sober layout. Däremot är priset knappast satt med tanke på studenter, verkets främsta målgrupp. Översikten i framför allt de två första delarna utgör goda introduktioner till fältet medan de övriga kapitlen ställer stora krav på den student som ger sig i kast med dem. Huruvida verket lever upp till ambitionen att utgöra ”the definitive guide” till lexicografen kan naturligtvis diskuteras: baksidestexten är sällan den bästa innehållsdeklarationen för vetenskaplig litteratur. Ändå är det den som de presumtiva köparna har att utgå ifrån när de står inför valet att köpa boken eller inte.

Uppläggningsen av boken kan bli förvirrande. Kapitlen är numrerade i förhållande till den del de ingår i; i del 4 sålunda från 4.1 till 4.10. Eftersom varje kapitel i sin tur delas in i numrerade avsnitt kan en hänvisning till exempelvis 4.2 dels avse kapitel 2 i del 4, dels avsnitt 4.2 i det aktuella kapitlet. Risken för onödigt bladdrande mellan kapitlen är dock liten eftersom hänvisningar mellan olika kapitel i volymen är sällsynta. De innehållsrutor som inleder varje kapitel ger en bra överblick över innehållet.

De forskare som engagerats för att medverka i volymen är etablerade som auktoriteter på sina respektive områden och de förstår att presentera sin forskning på ett intresseväckande sätt och samtidigt kritiskt. Det är sympatiskt att forskare från vitt skilda håll i världen inbjudits; möjligen spelar förlagets representation i fyra av de fem världsdelarna in här. Det är också tilltalande att teman som ligger något i utkanten av den centrala västerländska strömfåran bereds plats.

Innehållet i volymen har refererats ovan, utan större invändningar mot de enskilda kapitlen. Det starkaste intrycket av sträckläsningen av boken är dels redogörelsen för den diskussion som länge pågått om lexicografins teoretiska förankring och akademi-

ska status, dels beroendet av och relationen till korpuslingvistik och de korpusredskap som utvecklats i samarbetet mellan lexikografer och korpuslingvister. Det står klart att de framsteg som lexikografin gjort de senaste decennierna beror mer på datateknikens framsteg än på en teoribaserad metodutveckling. Också den komplicerade relationen mellan ordboksredaktionernas ambitioner och användarnas ordbokskompetens, såsom den skymtar i flera av kapitlen, är tänkvärd.

Presentationen av funktionsteorin såsom den renodlats inom Århuskolan och synen på lexikografin som en gren av informations- och referensvetenskapen och frikopplad från lingvistik är intressant och mer nyanserad och insiktsfull än den ibland har varit. Men de kritiska synpunkter som exempelvis Grønvik (2012) framför, avseende den långt drivna användaranpassningen, stöds av de rapporter om användarnas ordbokskompetens som ges i volymen. Också tillgången på så välstrukturerade och kompletta databaser som krävs för den grad av användaranpassning som funktionsteorin förespråkar har ifrågasatts (Kilgarriff 2012). En komplikation med funktionsteorin och dess terminologi – när den överförs till de nordiska språken – är, att lexikografin som disciplin är så nära förknippad med språklig verksamhet: analysen och beskrivningen av *ord* i *ordböcker*. Lika lite som alla lexikografiska produkter av tradition tar sin utgångspunkt i språkvetenskapen, lika långsökt är det att driva funktionsteorin för långt, exempelvis genom att förespråka totalt brott med språkvetenskapen även för den gren av lexikografin som i framställningen av ordböcker arbetar med analys och beskrivning av språkliga data. Relationen mellan ordböcker och andra typer av referensverktyg är därför i det närmaste dikotomisk och ordböckerna låter sig svårligen inordnas i en allmän informationsteknisk disciplin. Även om varken den syntaktiska eller den pragmatiska komponent som Piotrowski definierar kan ses som exklusiva för den

lexikografiska teorin, framstår det som svårt för språkvetare att anlägga i första hand ett informationstekniskt perspektiv på den semantiska komponenten: det ordförråd och dess beskrivning som ska inordnas i den syntaktiska strukturen.

Vissa områden berörs för ytligt i volymen. Särskilt beklaglig är den tvåspråkiga lexikografins undanskymda roll. I en tvåspråkig ordbok som är avsedd att publiceras på nätet ställs egentligen allting som hittills har gällt för den tryckta tvåspråkiga ordboken på huvudet. Båda språken kan göras sökbara för såväl L1-talare som L2-talare. Ett målspråk som också kan utgöra ingång i ordboken är inte längre bara ett målspråk vilket innebär helt nya krav både på ordboksarbetet och på metalexikografen (Hannesdóttir 2014). Den utveckling som datatekniken och publicering på nätet har inneburit för den enspråkiga lexikografen är bagatellartad i jämförelse med den revolution som den tvåspråkiga lexikografen nu står inför.

Volymens ambition är att visa vad som gjorts och görs inom lexikografen och att pekat ut en väg framåt. Den lyckas bäst med de två första momenten och mindre bra när det gäller att visa vägen framåt. Det hade varit intressant med större fokus på de utmaningar som lexikografen nu står inför och mindre på dagsläget.

Hur förhåller sig då innehållet i denna volym till tidigare publicerade antologier på området? Visst förekommer det en viss överlappning när samma forskare presenterar sin forskning i flera av volymerna, men naturligtvis också vissa kompletteringar av problem och perspektiv. Med tanke på det höga priset finns det all anledning för de utpekade målgrupperna, studenter och andra som vill bekanta sig med området lexikografi, att välja mellan de verk som nu finns på marknaden. Men även om de enskilda kapitlen sammantaget ger en utmärkt översikt över det lexikografiska fältet kommer *The Bloomsbury Companion to Lexicography* knappast att bli någon hörnsten i kurser i lexikografi.

Litteratur

- Eek, Øystein (2003): Finnes det en ordbokkritikk for redaksjons-sjefer? I: *LexicoNordica* 10, 45–51.
- Eide, Øyvind (2013): Solid utgivelse om elektronisk leksikografi. (Recension av) Sylviane Granger & Magali Paquot (red.): *Electronic Lexicography*. I: *LexicoNordica* 20, 205–221.
- Grønvik, Oddrun (2012): Digital leksikografi och funksjonsteoretiske briller. (Recension av) Pedro A. Fuertes-Olivera & Henning Bergenholtz (eds.): *e-Lexicography: The Internet, Digital Initiatives and Lexicography*. I: *LexicoNordica* 19, 239–254.
- Hannesdóttir, Anna Helga (2014): Lemman och ekvivalenter i nya roller – en reviderad ordbokstypologi. I: Ruth Vatvedt Fjeld & Marit Hovdenak (red.): *Nordiske studier i leksikografi 12. Rapport fra Konferanse om leksikografi i Norden. Oslo 13.–16. august 2013*. Oslo: Novus forlag, 193–211.
- Kilgarrieff, Adam (2012): (Review of) Pedro A. Fuertes-Olivera & Henning Bergenholtz (eds.). *e-Lexicography: The Internet, Digital Initiatives and Lexicography*. I: *Kernerman Dictionary News*. July 2012. Tel Aviv: K Dictionaries Ltd, 26–29.
- Larsson, Lennart (2009): Hur märker användaren av SAOB de 110 åren mellan A och TYNA? I: *Språk och stil* 19. *Tidskrift för svensk språkforskning*, 33–52.
- SAOB (1893–) = *Ordbok över svenska språket utgiven av Svenska Akademien* 1–. Lund.

Anna Helga Hannesdóttir
professor i nordiska språk
Göteborgs universitet
Institutionen för svenska språket
Box 200
SE-405 30 Göteborg
anna.hannesdottir@svenska.gu.se

Almqvist – en (meta)lexikografisk pionjär

Jon Helgason

Carl Jonas Love Almqvist: *Samlade verk. 17, Ordbok öfver svenska språket i dess närvarande skick*, redigerad och kommenterad av Sven-Göran Malmgren. Stockholm: Svenska Vitterhetssamfundet 2013, 562 s. Pris: 330 SEK.

1. Inledning

År 1841 stod Carl Jonas Love Almqvist i valet mellan att skriva ett geografilexikon över Sverige eller en svensk enspråkig ordbok. Det var P.M. Lindh, Almqvists förläggare, som till sist fattade beslutet att satsa på det sistnämnda. Resultatet blev en ordbok bestående av två häften som gavs ut på N.M. Lindhs förlag i Örebro 1842 respektive 1844, omfattande uppslagsorden *A–Brambär*. Mer färdig än så blev inte Almqvists ordbok. Detta ordboksfragment har nu utkommit i en textkritisk utgåva, redigerad och kommenterad av Sven-Göran Malmgren.

Almqvists insatser som språklärare och språkforskare har hamnat i skuggan av hans författarskap och ställning som en av den svenska 1800-talslitteraturens centralgestalter. Trots att hans gärning som språkforskare var omfattande och mångsidig så är den idag tämligen bortglömd. Efter flera år som lärare och informator anställdes Almqvist 1828 vid den då nyetablerade Nya elementarskolan i Stockholm, där han mellan åren 1829 och 1841 var rektor. I hans uppdrag ingick att författa läroböcker, bland annat i språk. 1829 utgav Almqvist dels *Allmän Språklära* (även i en andra upplaga 1835), dels *Svensk Rättstafningslära*, som kom att ges ut i 19 upplagor. 1832 gav han ut *Svensk Språklära*, som även den gavs ut i

flera upplagor. Där har särskilt den tredje upplagan från 1840 lyfts fram som viktig, eftersom den innehåller en omfattande översikt över de svenska dialekterna, inklusive ett antal ordlistor (Romberg 2000). Under 1830-talets andra hälft gav Almqvist därtill ut läroböcker i både grekiska och franska.

Almqvists ordbok får anses väl utforskad ur ett lexikografiskt perspektiv. Malmgren har själv i ett antal bidrag belyst den. Bland annat har han påvisat att Almqvists ordboksfragment med stor sannolikhet påverkat A.F. Dalin i dennes arbete med *Ordbok öfver svenska språket* (1850–1853) (Malmgren 1988). I andra bidrag har Malmgren diskuterat utgivningsprinciperna för ordboken, särskilt med utgångspunkt i svårigheterna kring tryckprocessen (Malmgren 2005). Almqvists ordbok är även föremål för en utförlig analys i Hannesdóttir (1999) där de olika informationskategorierna i Almqvists ordbok granskas och relateras till Karl Eduard Kindblads ordboksfragment respektive Dalins ordbok. Till forskningen om Almqvists ordbok bör också Johan Svedjedals litteratursociologiska avhandling (1987) räknas, som bland annat tar upp Almqvists förlagskorrespondens i samband med utgivningen av ordboken, samt Kurt Aspelins (1979) idéhistoriska contextualisering av ordboken och dess tillhörighet till en patriotisk och nationalistisk idétradition.

2. Utgåvans innehåll

Utgåvan består av tre delar. Första delen utgörs av Sven-Göran Malmgrens ”Inledning” (20 sidor). Där presenterar han kort Almqvists karriär fram till utgivningen av ordboken, hur Almqvists arbete med ordboken förlöpte och placerar in ordboken i en lexikografihistorisk kontext, för att sedan lyfta fram viktiga aspekter av Almqvists inledning. Malmgrens introduktion avslutas med kortare redogörelser för ordboksprojektets tillkomst, dess mottag-

ande i samtiden och tidigare forskning. Malmgrens introduktion är välskriven och läsvärd och ger inte minst en fascinerande inblick i Almqvists konstnärliga temperament och hur det uppenbarligen hamnar i konflikt med kraven på effektivitet vid skrivprocessen. Ett exempel är hur Almqvist trots, vad det tycks, akut tidsnöd i arbetet med ordboken ägnar hela spalter åt att redogöra för mycket subtila betydelsenyanser av ord som knappast kan sägas tillhöra det centrala ordförrådet. Ofta sker det med hjälp av redaktionella exempel. Malmgrens exempel i det här fallet är uppslagsordet *Bildhuggeri*, som i ordboken sträcker sig över nästan två spalter.

Den andra delen utgörs av editionen ”Ordbok öfver svenska språket i dess närvarande skick”. Editionen omfattar Almqvists ”Inledning” (28 sidor) och det i originaleditionen 650-sidiga ordboksfragmentet (479 sidor). Mer om editionen nedan.

Utgåvans tredje och avslutande del omfattar en kommentardel (50 sidor), bestående av en textkritisk kommentar, en redogörelse för ändringar gentemot originaleditionen samt speciella anmärkningar. Den innehåller även en textkritisk jämförelse mellan Almqvists inledning och två anonymt utgivna artiklar, som på goda grunder tillskrevs Almqvist redan av samtiden, publicerade med några dagars mellanrum i *Aftonbladet* i augusti 1842 på temat ”Om nyttan av en ordbok öfver svenska språket”. Denna del avslutas med en kortare ord- och sakförklaring till Almqvists inledning och till själva ordboksdelen, riktad till lekmannaläsare. Kommentardelen redovisar editionsarbetet på ett överskådligt vis.

Föreliggande utgåva ges ut inom ramen för det jättelika arbetet med utgivningen av Almqvists samlade verk, planerad att omfatta drygt 50 band, varav ett 20-tal är publicerade. Arbetet med Almqvists samlade verk påbörjades med Bertil Romberg som huvudredaktör (1993–2005), en post som numera innehas av Johan Svedjedal.

3. Almqvists ordbok

3.1. Almqvists inledning

Forskningen har särskilt framhållit Almqvists inledning som viktig från ett lexikografiskt och metalexikografiskt perspektiv (Malmgren 1988; Hannesdóttir 1999). Inledningen innehåller en noggrann genomgång av Almqvists tankar om och principer för ordboken. På så vis tillhandahåller inledningen en intressant historisk inblick i den tidens språkvetenskapliga och lexikografiska ställningstaganden. Almqvists diskussion av olika lexikografiska beskrivnings- och urvalsproblem är ofta skarpsinnig och i flera avseenden överraskande modern.

Nu är det i sig inget underligt. På flera ställen i Almqvists övriga språkvetenskapliga produktion framskymtar just påfallande progressiva ståndpunkter – flera av dem säkerligen kontroversiella i hans samtid. Det gäller inte minst frågan om de svenska dialekterna, i förlängningen (den lexikografiskt intressanta) relationen mellan skriftligt och muntligt språk:

Det språk, som man kallar *Svenska*, är en idé. Det finnes all-
enast i böcker och i grammatiskt bildade personers skrifter:
såsom verkligt lefvande språk (i tal) gifves det deremot icke,
om ej i någon persons mun, som tilljordt söker likna skrif-
språket, riks-idiomet, det allmänna. (Almqvist 1840:237,
kursivering i original)

Som det framgår av citatet betraktade Almqvist det talade språket som det verkligt ”levande” och menade vidare att impulserna till att föryngra ”riks-, bok-, och skrif-språket” (ibid., s. 239) kom via det språk som nyttjades av vanligt folk i tal. Språket utvecklades, enligt Almqvist, via de breda folklagren, underifrån. Almqvists organiska, romantiskt färgade språksyn avtecknar sig också i hans

ordbok, något som säkerligen bidrog till att ambitionsnivån på ordboksprojektet redan från början var orealistisk.

Inledning till ordboken uppvisar på samma sätt en överraskande fräschör. Almqvist uttalar sig tydligt för en deskriptiv lexikografi. En ordbok, skriver han, är inte ”en lagstiftning, utan en framställning af det varande” (s. 15, jfr äv. s. 20), en uppfattning och språksyn som för övrigt framgår redan av ordbokens titel. Han gör på upprepade ställen i inledningen en funktionsuppdelning mellan ordbok och grammatik. Likaså gör han en distinktion mellan en- och tvåspråkiga ordböcker och deras funktioner.

Almqvist anför i inledningen ett flertal skäl för nyttan av en ordbok av det slag han vill presentera. Några av dessa förtjänar särskilt att framhållas (se Hannesdóttir 1999 för en mer grundlig genomgång). För det första, menar Almqvist, bör varje ord, även de i allmänspråket kända, ges en definition i syfte att ”*gifva ordens betydelse sina rätta gränser*” (s. 9, kursivering i original). Utifrån denna grundprincip avhandlar Almqvist skarpsinnigt andra lexikografiska problem så som homonymi, polysemi och semantiskt relaterade ord. Almqvist bedrift är särskilt imponerande om man betänker att han var i färd med att ge ut den första svenska definitionsordboken. Just betydelsebeskrivningen var det drag som, av förklarliga skäl, var minst utvecklat inom det dåtida uteslutande tvåspråkiga lexikografiska området. För det andra vill Almqvist införliva dels facktermer, dels så kallade ”provincialismer” – dialektord – i ordboken. Anledningen till att Almqvist så helhjärtat lyfter fram dessa två aspekter får ytterst tillskrivas Almqvists språksyn. Han betraktade nämligen dessa båda språkliga områden som viktiga källor för berikande av ”riksidiomet”. I samband med det har Almqvist en intressant utläggning över skillnaden mellan allmänspråk och fackspråk. För det tredje framhåller Almqvist värdet av att ange uttal. I synnerhet framhåller han det historiskt förtjänstfulla av att utgöra ”en fixerad *anteckning öfver huru Svenskan i vår tid låter*” (s. 15, kursivering i original). Både ”ljud och tonvigt”

(s. 24) ska anges och Almqvist har en noggrann genomgång av hur uttalsangivelserna ska tolkas.

De längsta och mest utförliga avsnitten i inledningen berör behandlingen av verben. För varje verb vill Almqvist ange samtliga möjliga verbalsubstantiv (*-eri, -ande, -ende, -(n)ing, -nad, -an, -else*). Almqvists ambition avseende verben inskränker sig emellertid inte bara till verbens nominalavledningar. Denna ambition reflekteras också i själva ordboksdelen. Ordboken innehåller ingående information om verbens konstruktionssätt, fraseologi, tillhörande prepositioner, eventuella objektsväxlingar samt partikelverbens kombinatoriska egenskaper. Betydelsebeskrivningen av verben tar följaktligen påfallande stort utrymme, vilket gör att Almqvists ordbok med rätta kan betecknas som ”embryot till en svensk valensordbok” (Malmgren 1988:204).

3.2. Almqvists ordboksfragment

Almqvists ordboksfragment innehåller ca 6000 uppslagsord. Om den hade fullbordats skulle den ha innehållit närmare 70 000 ord, att jämföra med de 60 000 ord som finns i Dalins ordbok (Malmgren 1988). Som grund för lemmaurvalet för allmänspråket anger Almqvist Westes och Lindfors tvåspråkiga ordböcker (s. 18). I enlighet med Almqvist programmatiska inledning innehåller ordboken ett stort antal dialektala ord, i ordboken markerade med en asterisk framför uppslagsordet, och facktermer. Därtill innehåller den även geografiska ord och personnamn.

Ordboksartiklarna i Almqvists ordbok inleds med uppslagsord i fet stil, utdraget i marginalen i förhållande till artikeltexten. Vanligast är att varje uppslagsord representerar ett lemma. Verbavledningarna behandlas i regel som primära uppslagsord, så även lexikaliserade sammansättningar. Det gör att ordboken har en närmast rent glattalfabetisk makrostruktur. Efter uppslagsordet följer en uttalsangivelse markerad med ett kursivt ”u.”, följt av informati-

on om ordklass, deklination och genus för substantiv, deklination för adjektiv och konjugation för verb. Därefter kommer betydelsebeskrivningen. I förekommande fall följer sedan information om talesätt, särskilda ordförbindelser och konstruktionssätt m.m.

Uppslagsordet *Besked* får här utgöra exempel:

Besked, *u. beschéd*, s. n. 5. – 1) Befallning, utslag, afgörande; svar, upplysning. Konstrueras med *Gifva* och *Få*; samt *om*, *på* eller *öfver*. [---] 2) Gagn, ordentligt utförande af tjenst. Konstrueras mest reflexivt med *Göra* och *för*. [---] Icke reflexivt förekommer ordet med samma bemärkelse skämtvis, i talesättet ”*Göra besked för en skål*, vid pokalen, i dryckjom o. s. v.” – 3) Duglighet, drift, dugtighet; ord och aftal. Konstrueras mest impersonelt, med *Vara* och *med*. [---] Hit hör adverbsammansättningen *Med besked*, *Dugtigt*, *Mycket*. [---] – 4) Kännedom, Kunskap. Konstrueras med *Veta* och *om*, *på* eller *i*. [---] 5) Skriftligt intyg, bevis, betyg öfver personligt uppförande och beskaffenhet, hvilket *a*) af husbonde lemnas tjenstejon vid flyttning ur dess tjenst; *b*) af presterskapet gifves en ur församlingen afgående. Attest.

Denna artikel (här med de redaktionella exemplen utelämnade) upplyser på ett informativt vis om ordbokens rikedom avseende semantisk och syntaktisk information. En före detta redaktör för SAOB, som undertecknad, upplever också sättet att åtskilja ord-bokstextens informationsnivåer med kursiver, komman och semikolon som tydligt och överskådligt.

Almqvists ordbok är både språkhistoriskt och lexikografiskt intressant. Men, som också Malmgren lyfter fram i sin inledning (s. XVII), de enskilda artiklarna har ofta även ett sakligt innehåll av stort kulturhistoriskt värde. Almqvists betydelsebeskrivningar är därtill påfallande ofta roliga, spetsfundiga och intressanta. Be-

skrivningen av ord som *Artist* och *Banko* (äldre myntsort) utgör i sig en form av högst läsvärda minisexposéer.

4. Editionen av C. J. L. Almqvists *Ordbok öfver svenska språket*

Det föreligger dessvärre inget manuskript till Almqvists ordboksfragment. Avsaknaden av handskrifter i form av manuskript, arbetsversioner och utkast kännetecknar för övrigt större delen av Almqvists litterära och intellektuella kvarlätenskap. Det är förstås beklagligt att sådant material inte finns bevarat. Det hade kunnat ge en inblick i hur Almqvist rent praktiskt gick till väga vid arbetet med ordboken och vilket underlag han hade för sitt skrivande. Förlaga är alltså de redan utgivna två häftena av ordboken. Grundtexten till editionen är baserad på ett exemplar i Svenska Vitterhetssamfundets ägo, men ytterligare sju exemplar av ordboken har kollationerats, varav ett är Almqvists eget handexemplar (Böoksamlingen, Lunds universitet).

En målsättning med föreliggande utgåva (i det följande benämnd SV-utgåvan) är att så långt som möjligt rätta originaleditionens sättningsfel och andra former av slarv-, tryck- och konsekvensfel. Almqvist var fullt medveten om att ordboken var behäftad med ett stort antal sådana fel och skyller redan i inledningen till ordboken på att "Aflägsenheten från tryckningsorten förorsakat att ett antal tryckfel insmugit sig, särdeles af accenter" (s. 31). En del av felen försökte Almqvist själv åtgärda genom en rättelselista i slutet av det första häftet, men en motsvarande lista finns inte till häfte nr 2. Almqvist utlovade istället en rättelselista när första delen av ordboken var färdigställd (s. 516). Almqvists tryckta rättelser till det första häftet och Almqvists egna, relativt få, rättelser utförda i hans handexemplar är inarbetade i SV-utgåvan.

Malmgrens textkommentar och förteckning över emendingar och suppleringar ger Almqvist rätt i det att de flesta felen finns bland uttalsangivelserna i ordboken. Accenter för att markera uttal är felaktiga eller helt utelämnade på flera ställen i originalutgåvan. En mängd övriga suppleringar är även utförda i SV-utgåvan. Det gäller till exempel alfabetiseringsfel, sakfel i böjningsangivelser, felskrivna språkprov och de fall som Almqvist (eller sättaren) bryter mot de egenuppsatta lexikografiska principerna. Almqvist aviserar till exempel att sammansättningar markeras med förkortningen ”Ss.”. I de fall som en sådan markering har saknats i originaleditionen så har den förts in i SV-utgåvan osv.

Jag har utfört stickprov och jämfört SV-utgåvan med Almqvists handexemplar (första häftet) och har härvidlag också kunna jämföra mot Almqvists för hand införda ändringar. Det finns mycket lite att anmärka på. SV-utgåvans målsättning är såtillvida uppfylld, enligt mitt förmenande. Editionsarbetet tycks ha utförts med föredömlig noggrannhet. Då ska också framhåvas att Malmgrens arbete inte har inbegripit någon mekanisk rättning av felen i originaleditionen. För snart sagt varje enskild avvikelse har redaktören tvingats ta ställning till om eventuella inkonsekvenser är avsiktliga eller inte. En del dessa svårigheter har beskrivits av Malmgren själv (2005, 2013). Editionsarbetet, särskilt avseende accenterna, måste ha varit ett filigransarbete. En småsak emellertid: I kommentar delen anges att uttalet av den första förekomsten av uppslagsordet *Aj* (om trögdjurssläktet *Bradypus*) ändrats från originaleditionens *ai* till *ájj*. I handexemplaret är uttalet dock angivet som *ái*.

Malmgren anger att han använt uttalsangivelser i SAOB som stöd i arbetet med Almqvist och han varnar för att felaktiga eller tvivelaktiga angivelser i Almqvists ordbok stundom också okritiskt återges i SAOB. I det avseendet vill jag emellertid lugna både Malmgren och oroliga läsare och säga att problemet inte tycks vara särskilt omfattande. En sökning på det första häftet av Almqvists ordbok i OSA-versionen av SAOB visar att den använts 944 gånger i

SAOB.¹ I majoriteten av förekomsterna i SAOB används Almqvists ordbok som belägg på ord, i första hand sammansättningar – på vilka Almqvists ordbok är rik. Den används även som källangivelse för genus och för pluralformer. Däremot hänvisas det sällan till det första häftet för att ange uttal och det används aldrig, så vitt jag har kunnat konstatera, som intäkt för ett av SAOB rekommenderat eller existerande uttal. Ett av de få fall som jag har påträffat där SAOB övertagit en felaktig angivelse gäller uppslagsordet *Alf*:

ALF *al^tf* (a^tlfar (pl.) WESTE (1807); *alf* DALIN (1850); detta uttal, som beror på ordets egenskap af litterärt lån, afses tydligen af de flesta, som före vokal, t. ex. i pl. *alfer*, skrifva *f*, icke *fv*), äfv. *al^tv* (ållv ALMQVIST (1842); [---]).

I ovanstående citat anger SAOB för kännedom Almqvists uttalsangivelse och reproducerar därmed originaleditionens felaktiga grava accent. Baserat på Almqvists första häfte så tycks följaktligen problemet inte vara särskilt omfattande.

5. Sammanfattande omdöme

Fastän Almqvists ordbok troligen fick ett begränsad praktisk betydelse så rör det sig om ett pionjärarbete. Almqvist var den förste som gav ut en, låt vara ofullständig, svensk definitionsordbok. Almqvists ordbok utgjorde också i sig en klar brytning med den tradition som utvecklats inom den tvåspråkiga lexikografin. Det gäller särskilt betydelsebeskrivningen och konstruktionsupp-

1 Enligt Anton Lundqvist, kontrollant på SAOB, så har första häftet använts 896 gånger på A respektive B. 1 gång på S. Det andra häftet har använts 405 gånger på B. Almqvists inledning har använts som källa knappt 30 gånger spritt över hela SAOB. Notera att OSA-versionen omfattar uppslagsorden A–TYNA.

gifterna. Forskningen har även, som tidigare nämnts, visat att Almqvists lexikografiska principer med stor sannolikhet påverkade Dalin och hur dennes arbete kom att utformas (Malmgren 1988, 2005; Hannesdóttir 1999, 2001).

Almqvists ordbok har, som det har sagts, stora förtjänster. Hans ”Inledning” avhandlar en mängd metalexikografiska problem på ett intressant och för sin tid nydanande vis. Ordboksdelen är mer ojämn. Almqvist tar en hel del utrymme i anspråk på att förklara självklarheter, samtidigt som han blixtrar till i betydelsebeskrivningar och utredningar där han uppvisar sann briljans. I slutändan spelar det också roll att det är just Almqvist som håller i pennan. Det är bitvis svårt att bortse från kunskaperna om författaren Almqvist när man läser det ordboksredaktören Almqvist har skrivit. Ordboken kompletterar således bilden av ett för den svenska litteraturhistorien centralt författarskap.

Det är följaktligen en kulturgärning att Almqvists ordbok nu görs både mer tillgänglig och mer tillförlitlig. För språkhistoriska och lexikografihistoriska studier är det förstas utmärkt och skäl nog i sig. När det gäller just Almqvists ordboksfragment, så kan man emellertid förmoda att SV-utgåvan i kraft av Almqvists författarnimbus kan få en något större läsekrets än den strikt inomvetenskapligt lexikografiska. Det faktum att den förväntade läsekretsen för den föreliggande SV-utgåvan troligen är något större än blott det lexikografiska vetenskapssamhället gör att jag gärna hade sett en mer utförlig, kontextualiserande inledning. Detta emellertid sagt utan kännedom om de redaktionella avvägningar som eventuellt satt ramarna för inledningens omfång.

SV-utgåvan finns också tillgänglig på Litteraturbanken (se litteraturförteckningen) som faksimil. Den lexikografiska användbarheten hade varit än större om den digitala versionen gjorts fullt sökbar. Det är trots det föredömligt att Almqvists ordbok på så vis görs allmänt tillgänglig.

Litteratur

- Almqvist, Carl Jonas Love (1840): *Svensk språklära. Tredje upplagan, öfversedd och tillökad med samlingar öfver tio svenska landskapsdialekter*. Stockholm: M. Wirsell.
- Aspelin, Kurt (1979): *Studier i C.J.L. Almqvists författarskap åren kring 1840, del 1, "Det europeiska missnöjet". Samhällsanalys och historiespekulation*. Stockholm: Norstedt.
- Hannedóttir, Anna Helga (1999): *Lexikografihistorisk spegel. Den enspråkiga svenska lexikografins utveckling ur den tvåspråkiga*. (Meijerbergs arkiv för svensk ordforskning 23.) Göteborg: Göteborgs universitet.
- Hannedóttir, Anna Helga (2001): Funktioner som försvann: det kontrastiva perspektivet i äldre svensk lexikografi. I: Martin Gellerstam et al. (red.): *Nordiska studier i lexikografi* 5. (Meijerbergs arkiv för svensk ordforskning 27.) Göteborg: Göteborgs universitet, 3–17.
- Litteraturbanken: <www.litteraturbanken.se> (maj 2014).
- Malmgren, Sven-Göran (1988): Almqvist, Dalin och den svenska definitionsordbokens födelse. I: Gertrud Pettersson (red.): *Studier i svensk språkhistoria*. Lund: Lunds universitet, 195–213.
- Malmgren, Sven-Göran (2005): Författaren Carl Jonas Love Almqvists Ordbok öfver svenska språket i dess närvarande skick (1842–1844) och dess återutgivning. I: Ruth Vatvedt Fjeld & Dagfinn Worren (red.): *Nordiske studiar i leksikografi* 7. Oslo, 279–286.
- Malmgren, Sven-Göran (2013): Almqvists Ordbok öfver svenska språket i dess närvarande skick i Samlade verk. I: Gunnar Bergh, Rhonwen Bowen & Mats Mobärg (eds.): *Language, football and all that jazz: A Festschrift for Sölve Ohlander*. Göteborg: Acta Universitatis Gothoburgensis, 265–272.

- Romberg, Bertil (2000): C.J.L. Almqvist betraktar språket. I: Arne Jönsson & Anders Piltz (red.): *Språkets speglingar. Festskrift till Birger Bergh*. Lund: Lunds universitet, 56–64.
- Svedjedal, Johan (1987): *Almqvist – berättaren på bokmarknaden. Berättartekniska och litteratursociologiska studier i C.J.L. Almqvists prosafiktion kring 1840*. Uppsala: Uppsala universitet.

Jon Helgason
forskare, fil. dr
Språk- och litteraturcentrum
Lunds universitet
Box 201
SE-221 00 Lund
jon.helgason@litt.lu.se

Norsk Ordbok på nettet

Henrik Hovmark

Netudgave af Norsk Ordbok (eNO), afsnittet *i-tåvрут* (<no2014.uio.no>). Testet og anmeldt marts-april 2014.

1. Norsk Ordbok: trykt bog – og meget mere

Projektet Norsk Ordbok 2014 (NO2014) er tæt på målet. 12. og afsluttende bind af den trykte udgave, Norsk Ordbok (NO), udkommer i 2014, ganske vist ikke på selve 200-årsdagen for grundloven d. 17. maj, men dog inden årets udløb. Dermed afsluttes et usædvanligt leksikografisk projektførløb hvor en af de store nationale dokumentationsordbøger er blevet færdigredigeret på rekordtid, takket være en betydelig tilførsel af resurser i 2002 der resulterede i omlægning til fulddigital produktion og nye arbejdsgange og redaktionsregler. Projektet startede i 1930, bind 1-4 udkom 1966-2002 og bind 5-12 2005-2014.

Men projektet rækker videre end den trykte bog. NO2014 har i hele projektperioden haft fokus på at udnytte internettet til at give brugeren adgang til det kildemateriale der ligger bag den trykte bogs redigerede fremstilling (jf. fx Grønvik & Ore 2013, Ore & Wetås 2014): dels netværket af baser i Metaordboka (digitaliserede seddelsamlinger, kortsamling, billedsamling, xml-opmærkede ældre ordbøger m.m.), dels Nynorskkorpusset (p.t. ca. 100 mio. ord). Det er desuden blevet klart at ordbøger ikke længere kan nøjes med at være bøger, men (også) må tilgængeliggøres digitalt, først og fremmest på internettet og dets afledte former (apps). Den 15. marts 2012 lancerede man således første testversion af NO på nettet (en app findes ikke endnu). Netudgaven omfatter de dele af NO

som er blevet redigeret efter omlægningen til fulddigital produktion, dvs. p.t. afsnittet *i-tåvrrut*.

Arbejdet med netudgaven af NO har fået stigende prioritet undervejs, men det er klart at fokus hidtil primært har ligget på at få den trykte bog færdig til tiden. Når jeg i det følgende anmelder eNO, sker det derfor vel vidende at eNO stadig er i en udviklingsfase. På den anden side er eNO heller ikke længere en testversion – der er tale om en fuldt funktionsdygtig applikation, og det er i øvrigt også her man ledes hen hvis man søger på *Norsk Ordbok 2014* i Google (bortset fra at underlinket *Søk i Norsk Ordbok* på Google-oversigten fører brugeren hen til søgesiden på den gamle hjemmeside).

Anmeldelsen vil koncentrere sig om eNO som digital ordbog og bygger på en række repræsentative opslag. Inden jeg gik i gang, havde jeg kun hørt mundtlige præsentationer af eNO, men ikke brugt værktøjet. Jeg havde altså mulighed for at få en fornemmelse af situationen for en ny bruger. Undersøgelsen foregik marts-april 2014 – og når disse ord trykkes, kan meget have ændret sig. Følgende tekniske platform er brugt ved undersøgelsen: pc med Windows 7 Enterprise, Service Pack 1, og Mozilla Firefox 28.0. Med hensyn til sekundærlitteratur om e-leksikografi vil jeg generelt henvise til den anoterede bibliografi i Jackson (2013:415), til eLEX-konferencerapporterne og til Lew (2012) med henvisninger.

2. Opbygning, søgefaciliteter og design

Skærbilledet i eNO er inddelt i to hovedområder: 1) et rødt topfelt og 2) et centralt beigefarvet aktivitetsfelt (jf. figur 1). Al søgning og resultatvisning finder sted i aktivitetsfeltet, mens topfeltet er statisk – her vises kun projektets titel, *Norsk Ordbok 2014*, og nogle få knapper. Hertil kommer en tynd, sort topbjælke med klikbart link til projektets hjemsted: *UiO – Universitetet i Oslo*, og et bund-

felt (jf. figur 4) med 1) klikkbare logoer for alle partnere i projektet, 2) logo for NO2014 med kontaktinformasjon og 3) klikkbare logoer til NO2014's sider på Twitter, Facebook og Google+.

Søk i ordboka

niste Søk Bibliografi Avansert søk

FORRE TREFFLISTE ARTIKKEL BESTE

ALFABETISK LISTE

- ▲ nissemerke
- nissen
- nissespil
- nisseløys
- nissig
- 1 nist
- 2 nist
- 1 nista
- 2 nista
- niste
- nistebeite
- nistebidne
- nistebooks
- nistebrømm
- nistebrød
- nistebrøddelv
- nistedåse
- nistefer
- nistehey
- ▼

niste substantiv, hokjønn

1. mat ein tek med for å eta på ein tur, på arbeid el i (Heimfesting: Vanl, Kirf)

smarja niste
ta, ha med niste
laga til niste
leggja på niste (Heimfestingar: NLand, ØvHall, YSander, Froland KRf)

ha med mat heimetå for lenger tid

ORDNA: det er ingen naud um nista, når ein heve pening i pungen (Litterære kjelder: A.Ordspr 108, R.Th.-Christ.GV 58)

◆ BILETTEG

ha niste under trøya vera pukkestrygja (Heimfesting: Voss KRf)

◆ form (11a), farring (8,3d) (Heimfestingar: VTel A2, Set A2 KRf; Litterær kjelde: GOS 1,31)

2. OVERFØRD TYDING, utrusting, fide (8,2)

Årsboka i ranselen vil vere snau åndelig niste for dagens unge (Litterær kjelde: SS 1976,341)

**den beste niste på livsens ferd* (Litterær kjelde: V.SkrU 111,130)

ORDNA: kunst og næme er god niste på vegen (Litterær kjelde: A.Ordspr 80)

3. stor mengd av noko

han hadde til sluttet ei heil niste med papir i pengeboki si (Litterær kjelde: Ba.VR 11,464)

Figur 1: Søgeresultat ved søgning på *niste*.

Centralt i aktivitetsfeltet lidt til venstre for midten finder man søgefeltet (jf. figur 2). Umiddelbart til højre ses søgeknappen (*Søk*). Søgeresultatene vises neden under søgefeltet. Hvis søgeresultatet fylder mere end den afsatte plads i resultatfeltet, aktiveres en rullemenu i højre side (jf. figur 3). Når der er blevet søgt på et ord, vises der en alfabetisk lemmaliste til venstre for resultatfeltet (listen inkluderer også flerordsforbindelser, der i basen er lagret som selvstændige artikler). Den alfabetiske liste aktiveres også automatisk som en dropdown direkte under søgefeltet når man har tastet to bogstaver (fx *ni*, *po* osv.).

The screenshot shows the Lexiconordica search interface. At the top, there is a search bar with the text 'niste' and a 'Søk' button. To the right of the search bar are links for 'Bibliografi' and 'Avansert søk'. Below the search bar are tabs for 'FØRRE', 'TREFFLISTE', 'ARTIKKEL', and 'NESTE'. On the left side, there is an 'ALFABETISK LISTE' (Alphabetical List) with a scrollable list of words including 'nissemerke', 'nissen', 'nissespel', 'nissetrys', 'nissig', '1 nist', '2 nist', '1 nista', '2 nista', 'niste', 'nistebeite', 'nistebedne', 'nisteboks', 'nistebomme', 'nistebrød', 'nistebrødviv', 'nistedåse', 'nistefar', and 'nistehey'. The main search results area is titled 'niste substantiv, hokjønn' and contains a list of entries. The first entry is '1. mat ein tek med for å eta på ein tur, på arbeid el i (Heimfesting: Vanl, Kirf)' and includes sub-entries for 'smearja niste', 'ta, ha med niste', and 'laga til niste'. The second entry is 'luggja på niste (Heimfestingar: NLand, ØvHåll, YSandavær, Froland Kirf) ha med mat heimetrå for lenger tid' and includes a citation: 'ORDTAK: det er ingen naud um nista, når ein heve pening i pungen (Litterære kjelder: A.Ordgr 108, R.Th.-Christ.GV 58)'. Below this are two more entries: '♦ BILETEG' with the example 'ha niste under traya vera pukkeingja (Heimfesting: Voss Kirf)' and '♦ FORM (L1a), fering (L3d) (Heimfestingar: VTel A2, Set A2 Kirf, Litterære kjelder: GOS L31)'. At the bottom, there is a link to '2. nysporen tyttar utteltar ferja # 2'.

Figur 2: Søgefelt, resultatfelt og alfabetisk liste (søgning på *niste*).

Øverst i søgeresultatet ses til venstre det søgte lemma og til højre seks knapper. Ved tryk på de fire første knapper kommer eventuelle yderligere oplysninger frem om lemmaet: *Eldre kjelder*, *Alternative skriftformer*, *Uttale*, *vokalkvalitet* og *målføreformer* og *Etymologi*. Ved tryk på den femte knap vises et kort over Norge, inddelt i kommuner ca. 1950 (jf. Berg-Olsen & Wetås 2012:117) (jf. figur 3). Kortet vises også ved klik på udbredelsesangivelser (betydninger og målføreformer). Ved klik zoomes ind på den pågældende lokalitet som fremhæves med rødt. Der kan vises enkeltlokalitet – eller alle lokaliteter hvor en form eller en betydning er blevet stedfæstet (*sumstad*), i figur 3 målføreformerne *nistå*, *-u*. Hvis man er i tvivl om hvor en lokalitet befinder sig, kan man let zoome ud (og ind igen) ved tryk på +/- i øverste venstre hjørne af kortet. Den sjette og sidste knap folder eksempelstof m.v. ved alle betydninger ind (og ud igen). Eksempelstoffet kan også foldes ind og ud ved klik på en lille knap ved hver enkelt betydning.

The screenshot shows the Norwegian Lexicon interface. At the top, the title is "Søk i ordboka (i-tåvrut)". Below it is a search bar containing the word "niste" and a "Søk" button. To the right of the search bar are links for "Bibliografi" and "Avansert søk". Below the search bar are four tabs: "FØRRE", "TREFFLISTE", "ARTIKKEL", and "NESTE". The "TREFFLISTE" tab is selected. Below the tabs is a breadcrumb trail: "niste > nisteboks > nist > nist". On the left side, there is a vertical list titled "ALFABETISK LISTE" containing various words starting with "niste". The main content area shows the selected entry: "**I nist** substantiv, hankjønn". Below this, there is a definition: "vond attåtsmak (Halmfesting; Stryn R Kirø)". To the right of the definition are several icons for actions like print, copy, and share.



Det Norske Samlaget



NTNU

Det skapende universitet



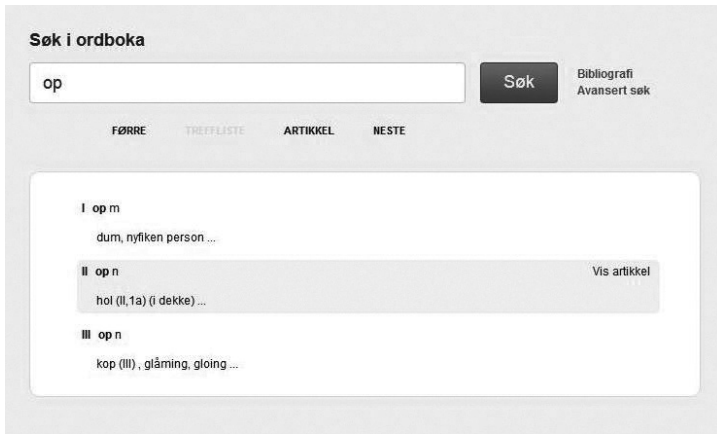
KULTURDEPARTEMENTET

Norsk Ord

E-post: no-pos

Figur 4: Søgehistorik ved oprindelig søgning på *niste* og med yderligere ord hentet ind fra den alfabetiske liste, senest *I nist*.

Umiddelbart under søgefeltet findes fire klikbare overskrifter: FØRRE, TREFFLISTE, ARTIKKEL og NESTE (jf. figur 5). En søgning viser som default en fuld artikel, men ved tryk på TREFFLISTE fremkommer en oversigtsvisning hvor man kun ser opslagsordet og et udsnit af første definition. Hvis en søgning matcher flere resultater, dvs. hvis der er homografer, vises fundene samlet i kortform i resultatvinduet. Hvis man herefter klikker på ARTIKKEL eller direkte på et af fundene (der fremhæves med lillablå farve ved mouseover, jf. figur 5), kan man klikke frem og tilbage mellem fundene med FØRRE og NESTE. Jeg skal vende tilbage til denne funktion (afsnit 3.4).



Figur 5: Oversigtsvisning med mouseover på *II op* og mulighed for at klikke frem og tilbage mellem homograferne.

Til højre for den blå søgeknapp findes henvisningsknapper til *Bibliografi* og *Avansert søk*. Yderligere til højre findes fire klikbare overskrifter, indledt af en pil: *Ordbok*, *Korpus*, *Metaordboka* og *Andre kjelder* (jf. figur 3). Herved kan man skifte mellem selve ordbogen (*Ordbok*) og andre søgebaser, delvis inden for samme webinterface som ordbogen: De to sidste knapper fremkalder en kort beskrivelse af *Metaordboka* og *Andre kjelder* i resultatfeltet, men selve adgangen opnås uden for eNO's interface ved at klikke på eksterne links. Søgning i *Korpus* er til gengæld integreret i samme søgeflade som ordbogsdelen og giver mulighed for både simpel og *Avansert søk* (fx højre-/venstre-sortering, brug af regulære udtryk m.v.).

Øverst til højre i det røde topfelt findes fire klikbare overskrifter: *FAQ*, *Om oss*, *Rettleiing* og *Andre ordbøker*. Ved tryk på overskrifterne aktiveres det i øvrigt selvforklarende indhold i resultatfeltet. Henvisningerne er til lignende projekter i Norden og Europa, fx *Den Danske Ordbog* (DDO) og *Oxford English Dictionary* (OED). Titlen *Norsk Ordbok 2014* er også klikbar – herved

indlæses siden forfra og tidligere søgninger i historikken slettes, hvilket undertiden kan være en nyttig mulighed.

Som det fremgår, er der mange knapper til rådighed i eNO, men de virker godt grupperede, og siden som helhed virker ikke uoverskuelig. Aktiviteter bliver dog komplicerede i det øjeblik man inddrager eksterne kilder – bl.a. fordi søgesiden vil være nulstillet og søgehistorikken slettet når man vender tilbage til ordbogssiden. Dette er ikke optimalt (jf. videre afsnit 3.6). Den afdæmpede diskretion kan også være en svaghed: Knapperne til højre i det røde topfelt (FAQ osv.) er der næppe mange der vil lægge mærke til – de er små og står uden for og langt væk fra aktivitetsfeltet. Dette er især uheldigt for de to hjælpeknapper *FAQ* og *Rettledning* som burde have en mere prominent plads eller rykkes tættere på brugerens aktiviteter i aktivitetsfeltet.

Jeg skal vende tilbage til andre komplikationer ved søgninger og visninger. Et enkelt problem er dog relevant at nævne her, nemlig spørgsmålet om visning af lange og omfattende søgeresultater. Dette er et klassisk problem som det er vanskeligt at finde en ultimativ løsning på, selv i en digital ordbog med rige muligheder for at folde ud og op og i. I betragtning af hvor stort materialet bag NO og eNO er, og hvor mange og lange søgeresultater en bruger kan komme ud for, er det en stor luksus ikke at udnytte hele skærmens højde ved visning af resultater – og generelt at lade så stor en del af skærmen være uudnyttet. En løsning kunne være at anbringe søgefelt og resultatfelt ved siden af hinanden så skærmens højde kan udnyttes bedre ved resultatvisning – eller i det mindste gøre de store, statiske top- og bundfelter mindre. Til gengæld er opbygningen med et dominerende søgefelt i centrum lidt til venstre ideel når applikationen skal fungere på en større skærm (men ikke nødvendigvis i en app). I DDO står søgefelt og resultatfelt ved siden af hinanden – men søgefeltet, som er noget mindre end i eNO, er ved denne manøvre endt øverst i højre side, med lemmalisten nedenunder. Opbygningen på eNO er klart bedre, med et promi-

nent søgefelt inden for det område hvor brugeren naturligt kigger hen. Men der kan med fordel eksperimenteres med placeringen af resultatfeltet og med udnyttelsen af skærmfladen som helhed, også uden at det går ud over æstetikken (jf. fx OED).

eNO fremstår generelt som en meget overskuelig digital ord-bog, med et afdæmpet funktionelt design. Der er god plads mellem de enkelte aktivitetsområder, og de afrundede hjørner på knapper og bokse giver et harmonisk indtryk. Der bruges ganske mange farver på siden: de norske farver, blå, rød og hvid, samt beige, grå, sort og lyslilla. Men de fremstår alle i afstemte, behagelige nuancer, og de udnyttes diskret og elegant til at inddele siden i forskellige områder og til at adskille og fremhæve forskellige elementer.

3. eNO og den digitale platform

Omsætningen fra trykt bog til webpublikation rummer velkendte muligheder for at frigøre sig fra den trykte bogs fysiske og omfangsmæssige begrænsninger, og hertil kommer særlige muligheder i den digitale platform. eNO udnytter i varierende grad alle disse muligheder.

3.1. Udfoldning af forkortelser m.v.

Da Norsk Ordbok blev til Norsk Ordbok 2014, blev der foretaget en række forbedringer af layoutet i den trykte bog, fx tydeligere betydningsmarkering og sammensætningsartikler på ny linje. NO er dog stadig en meget komprimeret og ikke særlig læservenlig ordbog. Af særligt berygtede unoder kan nævnes at der ikke bruges ny linje i artikler, og at forkortelser virkelig er meget korte: Det lakoniske *A* står fx for selveste Ivar Aasen.

Disse unoder er oplagte at rydde af vejen i en netudgave, og det har man også gjort som noget af det første i eNO. Resultatet er en

stor gevinst for brugeren. Ny linje anvendes ved alle betydninger og underbetydninger (inklusive fx *ordtak*, ordsprog og talemåder) samt ved hvert enkelt sprogeksempel og ved henvisningsrækker (jf. figur 1). Lige så væsentligt er det at det dybe betydningshierarki (op til fem niveauer) også fremstår tydeligere: Der er indsat et indryk ved nyt niveau i hierarkiet, og den helt ulæselige dobbeltbjælke er erstattet af et stjernelignende symbol. Endelig er forkortelser skrevet ud, enten i form af det fulde ord på skærmen, fx *målf* til *målføreformer*, eller i form af mouseover-informationer om kilder. Det sidste er velmotiveret i og med at kildeangivelserne kan være meget lange og ville optage alt for meget plads på skærmen. Til gengæld kan man i øjeblikket ikke få adgang til en samlet kilde-liste: Man kan ikke klikke sig videre fra kildeforkortelsen selvom den er fremhævet som et link, og knappen *Bibliografi* til højre for søgefeltet ender også blindt.

Det skal også nævnes at henvisninger fremstår tydeligere for brugeren i eNO. I NO kan det være næsten umuligt at få øje på henvisningerne, de markeres hverken med font eller markør og ses kun indirekte hvis der optræder et betydningsnummer bagefter i parentes (jf. fx *overgå* (1A) under *IV slå A5* vs. *innsjå* under *IV slå A6c*). I eNO er henvisninger fremhævet med blå. De har dog heller ingen markør når de er del af en betydningsbeskrivelse, og en fælles henvisningsmarkør, fx en pil, burde måske overvejes.

3.2. Forskellige visninger

Det dybe betydningshierarki kan være kompliceret at navigere i, især i store artikler. Selve hierarkiet er en arv fra den trykte ordbog, og en eventuel omorganisering må afvente en ny udgave af ordbogen. Men der er alligevel gjort tiltag som prøver at udnytte den digitale platform til at gøre fx store artikler mere overskuelige, nemlig ved at indbygge mulighed for forskellige, kortere eller mere omfangsrige visninger.

Det første tiltag er enkelt og genkendes fra andre digitale ud-

gaver af trykte ordbøger. Som udgangspunkt vises alle artikler i fuldt udfoldet form, men ved klik på det lille kryds kan man folde eksempel materialet m.v. væk i betydningsbeskrivelserne (jf. afsnit 2). Basis for faciliteten er imidlertid den trykte udgaves lineære fremstilling, og her kan definitioner være vanskelige at forstå uden den efterfølgende tekst (jf. fx Lorentzen & Trap-Jensen 2012:98, Nilstun 2012:134). Ved *niste* står BILETLEG fx alene tilbage som underbetydning. Trods disse uhensigtsmæssigheder er faciliteten nyttig og især en hjælp i de store artikler der som nævnt er ekstra svære at overskue på grund af den manglende udnyttelse af skærmens højde ved resultatvisning (*opp* har fx 16 hovedbetydninger og adskillige underbetydninger).

Noget som til gengæld er foldet sammen i udgangspunktet, er oplysningerne om målføreformer osv., dvs. det der befinder sig i den firkantede parentes i den trykte bog, mellem opslagsord og betydningsdel. Oplysningerne kan hentes frem ved at trykke på ikonknapperne øverst til højre i resultatfeltet. Men får brugeren øje på informationen? Risikoen for at brugeren overser en nyttig information (jf. Lew 2012:353), er her vejet op imod hensynet til at oplysningerne er forholdsvis specialiserede, og at en særligt interesseret bruger nok vil finde dem alligevel. Dispositionen synes fornuftig nok på nuværende tidspunkt – som netop vist er problemerne med det lille resultatfelt store nok, og nogle af informationerne fremstår også ret uformidlede når de hentes direkte fra den trykte bogs fremstilling (indholdet under *Alternative skriftformer* ved *niste* er fx *e skr* og *nyste* (Fedrah 1883) – som ikke forklares nærmere, heller ikke med mouseover). Noget andet er at ikonknapper nok er enkle, men også lidet oplysende for dem der ikke kender ikonerne i forvejen. Det eneste ikon som kan siges at være selvforklarende, er kortikonet, de øvrige er afhængige af mouseover-funktionen. På længere sigt kunne man dog godt forestille sig en mere iøjnefaldende visning af data om fx målføreformer, kombineret med kort og lyd.

3.3. Vejledningstekst

Det er glædeligt at eNO udnytter mouseover, fx ved kilder og ved ikonknapper. Der findes også en mere klassisk vejledningstekst til siden som helhed, men den befinder sig som nævnt ret langt væk i øverste højre hjørne i det røde topfelt. Teksten er meget pædagogisk og forbillidligt klart stillet op, men den indeholder primært oplysninger om artiklernes indhold (fx oplysninger om ordklasse og bøjning), hvorimod der er meget få oplysninger om selve sidens funktionalitet, fx søgemuligheder, ordlisten eller søgehistorikken. Ikonknapperne forklares, men den interessante kortvisning af geografisk lokalisering får desværre ingen ord med på vejen. Derudover står vejledningsteksten som én lang fortløbende tekst – der kan ikke laves emnesøgning e.l. Teksten gør bestemt ingen skade, men den er svær at få øje på, og i videreudviklingen af hjælpefunktioner og -tekst bør der også være vejledning til søgefaciliteterne, ligesom FAQ må aktiveres. Det vil også være interessant at overveje hvordan vejledningen kan anbringes tættere på brugerens handlinger på siden. eNO viser allerede vejen med mouseover-funktionerne som kunne udvides, måske med hjælpefunktion aktiveret med højreklik og med relevant vejledningstekst i et popup-vindue dér hvor brugeren befinder sig på siden?

3.4. Nye faciliteter

Allerede på nuværende tidspunkt introducerer eNO forskellige faciliteter som er muliggjort af den digitale platform, fx ordlisten i venstre side. Ordlisten genereres faktisk allerede som en dropdown-liste når man taster ind i søgefeltet. Dette kaldes bl.a. *type-ahead search* (Lew 2012:351f.). eNO udnytter brugerens indtastning til at sende et vejledende signal (*cue*) til brugeren mens indtastningen foregår. Den e-leksikografiske verden er meget optaget af cue-tanken netop nu, især fordi noget tyder på at såkaldt

almindelige brugere ikke gider bruge tid på at tilpasse en søgeside – de ønsker kun at skrive i ét søgefelt og har uforkammet høje forventninger om et meget hurtigt svar (jf. fx Trap-Jensen 2010b). Hvis en digital ordbog gerne vil have en høj succesrate og holde fast i sine utålmodige brugere, kan det derfor være en god idé at udtænke og inkorporere forskellige mekanismer der registrerer og analyserer brugerens adfærd og sandsynlige ønsker i baggrunden mens brugeren søger. På basis heraf kan man prøve at hjælpe brugeren på rette vej ved at prioritere søgeresultatet på bestemte måder (jf. fx Google-søgninger) eller ved at generere synlige vejledende signaler (cues). Ordlisten under søgefeltet er i princippet en sådan enkel cue. Den fungerer glimrende i eNO, ikke mindst fordi den også viser flerordsforbindelser (et eksempel på det både nyttige og rigtige i at disse forbindelser er oprettet som selvstændige lemmata i NO2014's base – med mulighed for interne links).

Det er oplagt at arbejde videre med denne type cue i eNO. For det første er placeringen af flerordsforbindelser ikke optimal. Den alfabetiske liste er i sandhed alfabetisk, og forbindelsen *til råds* dukker fx kun op i listen ved opslag på *til*, ikke på *råd* hvor den måske ville være mere relevant. Yderligere er basen ikke konsekvent med hensyn til henvisninger: *til råds* er oprettet som henvisning i grundmanuskriptet og optræder som følge heraf to på hinanden følgende gange i den alfabetiske liste i lighed med fx *til og med*, *til overs* m.fl., i modsætning til en række andre forbindelser med *til*, fx *til orlogs*, *til reie* m.fl. der kun optræder én gang. For det andet kan arbejdet med at identificere hvilket opslagsord brugeren tænker på, videreudvikles (jf. Lew 2012:345ff.). Brugere indtaster ikke nødvendigvis opslagsformen, men en variantform eller bøjet form – og fra Google er man desuden vant til at indtaste flere ord. Der knytter sig også særlige komplikationer til variantformer og opslagsform i eNO: Opslagsformerne følger retskrivningsnormen fra 1938, og Metaordboka har brug for én grundform for at kunne fungere (jf. Berg-Olsen & Wetås 2012). Men netop derfor er det

oplagt at tænke i hvordan de nye teknologier kan blive til hjælp her, for man kan ikke forvente at det langt større antal brugere NO vil få med eNO, skal eller vil kunne gennemskue de særlige forhold omkring opslagsformerne. Man er ikke nødvendigvis dårligt stillet på NO2014 i et sådant arbejde, for som dialektordbog har man netop registreret og lagret en række sprogbugsdokumenterede variantformer (målføreformerne) som kunne danne udgangspunkt for en mere intelligent fortolker af brugerens indtastning. Under alle omstændigheder bør man overveje hvordan variantformerne anføres: Af de fem variantformer under *niste*, er det i øjeblikket kun én, *nyste*, der står i ordlisten som henvisningsartikel.

En anden velkendt cue som også introduceres i eNO, er oversigtsvisning af homografer (jf. figur 5). Jeg synes også godt om dette tiltag – men måske kunne man i visningen i højere grad markere at søgeresultatet kun er foreløbigt, og at den egentlige information findes ved at klikke sig videre. I øjeblikket ses det først ved mouse-over over en af valgmulighederne, hvorefter *Vis artikel* dukker op i højre side. Det er blevet fremført at brugere let kan lade sig irritere af insisterende tekst der afkræver svar og kan synes at stå i vejen for at man får vist et ordentligt søgeresultat (fx ”Mente du xx, yy, zz?”). Jeg er for så vidt enig, men i eNO får man heldigvis mere end bare ordklassen præsenteret, nemlig meningsgivende betydningsdefinitioner eller dele heraf, en cue-model som brugere i forvejen er vant til fra Google (jf. Lew 2012:357). Men en tydeligere markering i form af et ikon og/eller en klikbar tekst vil alligevel være fordel.

Endelig introducerer eNO søgehistorikken som en facilitet. Dog kun i begrænset udgave: Ved homografer kan man benytte knapperne FØRRE og NESTE, og når der vælges ord fra ordlisten, registreres og anføres ordene under søgefeltet (jf. figur 4). Klik på homografer ved oversigtsvisning opsamles også i søgehistorikken. Der knytter sig imidlertid en række begrænsninger til faciliteten, fx at søgehistorikken forsvinder når man laver en ny søgning via

søgefeltet (jf. afsnit 2). Man kan ganske vist komme tilbage til tidligere søgninger og resultater ved at klikke på browserens Tilbage-knap, men søgehistorikken i eNO er endnu ikke nået helt i mål. Et eksempel på en mere integreret og fuldstændig søgehistorik-facilitet ses fx i OED.

3.5. Udnyttelse af andre medier: visning af kort

eNO beskriver også nynorsk dialekt, dvs. talesprog, og det ville være oplagt at koble udtaler, dvs. lyd, til fx målføreformer. Det kan desværre ikke lade sig gøre endnu, men man må håbe at tale og lyd hurtigt kan få en plads i eNO. Det vil være en afgørende forbedring og et sted hvor den teknologiske udvikling kan gøre en markant forskel.

Til gengæld giver eNO allerede nu mulighed for visning af lokalisering (*heimfesting*) af målføreformer og betydninger på kortet til højre for resultatfeltet (jf. figur 3). Brugeren gøres ikke opmærksom på kortets eksistens som udgangspunkt – det vises først når der klikkes på enten en lokalisering (en kommune) eller på KART efter lokalisering. Kortet i eNO fungerer enkelt og effektivt og giver nyttig information til både læg og lærd. Der er virkelig tale om merværdi, både indholds- og formidlingsmæssigt. Måske burde muligheden for kortvisning fremhæves på eNO, for lokalisering og kort er i dén grad blevet trendy overalt: Den brede almenhed har længe nydt godt af gps-systemer og stadigt mere detaljerede udgaver GoogleMaps o.l. – og med smarttelefonernes indtog bliver vi konstant opfordret til at ”dele vores placering” (hvorved man kan registrere og udnytte vores spor). Mere specialiserede brugere er også vilde med kort og lokalisering, det gælder fx inden for dialektologi, sociolingvistik og kulturgeografi.

Men faciliteten (og de bagvedliggende baser) kan udnyttes yderligere. Næste skridt kunne være at udvikle mere generelle kort over lydformer og variationer – man kan her tænke på Jysk Ord-

bog der i mange år har udnyttet kort til at vise udbredelser. Den slags kort vil uden tvivl have interesse for andre end fagfolk. Man kunne også tænke sig at brugeren kunne få vist former, ord og betydninger som er knyttet til en kommune eller et større område. Dermed ville man nærme sig en mere multimodal resurse hvor indgangen for brugeren ikke kun er et eller flere ord i søgefeltet, men fx også en lokalitet (på et kort).

3.6. Integrering med andre resurser

Som nævnt er NO2014 ikke bare en ordbog – ordbogsprojektet trækker på et netværk af databaser, organiseret af Eining for Digital Dokumentation (EDD) ved Universitetet i Oslo. NO2014 har længe givet offentligheden adgang til flere af disse baser (fx de digitaliserede sedler), også på nettet, men offentliggørelsen af netudgaven sigter tydeligvis – med rette – mod at udnytte den digitale platform til at integrere disse bagvedliggende, linkede resurser på en bedre og mere brugervenlig måde, jf. tankerne om Linked Open Data. Der er dog et stykke vej endnu. Søgning i korpus foregår som nævnt i samme ramme som ordbogen, mens Metaordboka og Andre kjelder (fx Sedelarkivet) åbner i et separat link. Hvis en bruger udnytter de øvrige resurser i basen, vil det normalt være for at sammenligne data eller for at få mere information om søgeresultatet i ordbogen. Det vil derfor være en fordel hvis man kan se oplysningerne sidestillet (fx i separate vinduer) sådan som det er tilfældet i integreret form i eNO's ordbogsdel når kortet vises. I øjeblikket vises imidlertid en frisk side når der klikkes på de tre eksterne links, mens søgeresultatet forsvinder. Det kan kaldes frem igen ved at klikke på browserens Tilbage-knap, men jeg ville foretrække en visning hvor søgeresultatet bliver stående, og hvor man kan arbejde parallelt med data i fx ordbog, korpus og seddelarkiv, fx i separate vinduer – eller hvor de er eller kan linkes løbende og vises parallelt sådan som det gøres med kortvisningen i selve ord-

bogen. At lave en sådan total integrering af ressourcerne er et ikke helt ukompliceret fremtidsønske (jf. Trap-Jensen 2010a), for slet ikke at tale om muligheden for at eksportere eller lagre forskellige søgeresultater i de forskellige ressourcer, men det er alligevel et vigtigt mål for projektet: Det vil tilgodese specialiserede brugere, og muligheden for at integrere ordbog med kilderessurser fremhæves til stadighed som en styrke af projektet selv (senest fx Grønvik & Ore 2013).

4. Opsamling og diskussion

Der er grund til at rose NO2014 for at lancere eNO allerede inden den trykte bog er færdig. Hermed kan langt flere brugere få glæde af ordbogens righoldige oplysninger og mange års ihærdig redaktørindsats. Arbejdet med netudgaven er nået langt, og det er især glædeligt at man ikke kun omsætter det vanskeligt læselige bogtryk til en mere tilgængelig form, men også begynder at inkorporere faciliteter som er muliggjort af den digitale platform og den digitale kontakt med brugeren. eNO er stadig i en udviklingsfase, og denne anmeldelse kan forhåbentlig bruges i det videre arbejde.

Generelt synes jeg at eNO tenderer mod at være lidt for diskret i sine signaler og markører (måske spiller æstetiske hensyn en rolle?). eNO kunne godt være mere eksplicit og udnytte den ekstra plads som netpublikationen giver mulighed for. I OED står der fx *Quotations: Show all* | *Hide all* – i eNO ses kun et kryds/plus-tegn. I DDO markeres de forskellige oplysningstyper også i ord, fx SYNONYM og EKSEMPLER, men også *Vis overblik*, UDTALE og OPRINDELSE. Oplysninger om fx udtale og oprindelse skjuler sig i eNO bag ikoner som man nok lærer at kende efterhånden, men en ny, travl og utålmodig bruger vil let kunne overse dem. Især er det som sagt ærgerligt at kortfunktionen ikke fremhæves mere tydeligt når funktionen allerede er så godt integreret i applikationen og

virkelig er en ekstragevinst. Jeg tror ikke man skal være bange for at tale direkte til brugeren og fx fortælle at ”Her kan du finde hjælp og vejledning” eller ”Her kan du søge i ordbogens kilder” e.l.

Jeg har undervejs også givet eksempler på forskellige fejl, mangler og uhensigtsmæssigheder i eNO. Men der bør skelnes mellem typerne. Der er 1) fejl som ikke har noget at gøre med eNO som applikation, men som skyldes indholdet i basen, og dem ser jeg umiddelbart bort fra, bortset fra de tilfælde hvor de genererer inkonsekvenser, jf. fx angivelsen af flerordsforbindelser i den alfabetiske liste (jf. afsnit 3.4). Dernæst er der 2) uhensigtsmæssigheder som skyldes at netudgavens visning af oplysninger er bundet til kategoriseringen af oplysninger i databasen – som på en række punkter atter har rødder tilbage til den lineære fremstilling i den trykte bog (jf. afsnit 3.2). Der er også 3) fejl i mere banal forstand: Småting der ikke fungerer, eller klikbare elementer der ikke fører til noget. Det gælder fx den manglende bibliografi og den manglende mulighed for *Avansert søk*. Hertil kan fx tilføjes at man ved klik på det henvisningsfremhævede ord *forn* i *niste 1* ikke får den relevante meddelelse om at ordet befinder sig uden for det alfabetiske afsnit som p.t. er med i eNO, men en generel meddelelse om at søgestrengen ikke har givet noget resultat, og at man kan forsøge sig med trunkering e.l. Der er ikke uacceptabelt mange fejl af denne type i eNO på nuværende tidspunkt, men dog alligevel så mange at man bemærker dem. Endelig er der 4) mangler som skyldes at nyudviklede faciliteter endnu ikke fungerer optimalt. Det gælder fx søgehistorikken.

Til punktet om fejl vil jeg knytte to kommentarer. For det første viser gennemgangen tydeligt at det at udgive til internettet og andre digitale platforme er en selvstændig redaktionel opgave. At data ligger i en velformet base, giver fantastiske muligheder, herunder fx visning på nettet e.l., men man får ikke leveret et fikt og færdigt værktøj blot ved at trykke på en knap. Arbejdet er altså ikke gjort med at rette simple fejl – der skal afsættes resurser til at

omsætte data til den særlige kommunikationsplatform og formidlingstype som nettet e.l. udgør, og man må også være forberedt på at skulle lave justeringer og konsekvensrettelser i basen. Men i betragtning af hvor vanskeligt det er at læse og bruge den trykte NO, er eNO nået langt. For det andet kan man overveje strategien i forhold til antallet og typen af fejl. Simple fejl er selvfølgelig aldrig god reklame, og funktionaliteter der ikke rigtig fungerer endnu, kan skræmme interesserede brugere væk. Måske skulle man vente med at vise knapperne *Avansert søk*, *FAQ* osv.? Man kunne også overveje at vente med at introducere søgehistorikken som funktion til den er mere fuldstændig, og benytte lejligheden til at lave et friskt fremstød for hjemmesiden. På den anden side er det fuldt forståeligt at eNO har lyst til at introducere søgehistorikken i en miniudgave, og det er under alle omstændigheder rigtigt set af eNO at give sig i kast med netplatformens særlige muligheder. Registrering af spor er i lighed med funktioner der arbejder i baggrunden, i eksplosiv vækst i disse år, jf. fx brugen af cookies. En digital ordbogsbruger efterlader sig også spor, som leksikografien kan udnytte til brugerundersøgelser, cue-funktioner osv.

Internettet og andre digitale platforme er kommet for at blive. Her opsøger brugerne løbende information, og her er de også aktive og kommunikerer i både arbejde og fritid. Her bør de derfor også kunne finde solide sproglige hjælpemidler af høj kvalitet både indholds- og formidlingsmæssigt – og kunne regne med at disse hjælpemidler løbende bliver opdateret og holdes levende. Digitalisering er tidsbesparende og giver nye muligheder, men afføder også nye resursekrævende arbejdsopgaver. Jeg vil derfor slutte denne anmeldelse af med et kraftigt ønske om at der kan skaffes midler ikke kun til at konvertere og implementere A-I i eNO, alt andet er meningsløst, men også til at vedligeholde og videreudvikle værktøjet. Talrige erfaringer fra ordbogsprojekter i hele Norden peger helt entydigt på at det er en ualmindelig god investering (for fonde eller offentlige myndigheder) at give penge til både udvik-

ling og løbende vedligeholdelse. Hvis man forsømmer almindelig vedligeholdelse, vil tingene hurtigt forfalde eller forældes, den akkumulerede viden på stedet vil forsvinde, og man må i gang med en meget kostbar hovedreparation eller en total udskiftning.

Litteratur

Ordbøger

DDO = *Den Danske Ordbog* på nettet. <ordnet.dk/ddo> (april 2014).

eNO = Netudgave af *Norsk Ordbok* (NO), afsnittet *i-tåvrrut*. <no2014.uio.no> (marts-april 2014).

Jysk Ordbog = *Jysk Ordbog*. <www.jyskordbog.dk> (april 2014).

NO = *Norsk Ordbok* (1966/2002-2014). Hovedredaktører: Lars S. Vikør m.fl., projektdirektører: Åse Wetås m.fl. Oslo: Det Norske Samlaget.

OED = *Oxford English Dictionary online*. <www.oed.com> (april 2014).

Anden litteratur

Berg-Olsen, Sturla & Åse Wetås (2012): *Norsk Ordbok på nett: Utfordringer og løysingar*. I: Birgit Eaker m.fl. (red.): *Nordiska studier i lexikografi 11. Rapport frå Konferensen om lexikografi i Norden, Lund 24-27 maj 2011*. Lund: Nordiska föreningen för lexikografi, 115-127.

Grønvik, Oddrun & Christian-Emil Ore (2013): What should the electronic dictionary do for you – and how? I: Iztok Kosem et al. (eds.): *Electronic lexicography in the 21st century: thinking outside the paper. Proceedings of the eLex 2013 conference, 17-19 October 2013, Tallinn, Estonia*. Tallinn: Eesti Keele Instituut, 243-260.

- Jackson, Howard (ed.) (2013): *The Bloomsbury Companion to Lexicography*. London & New York: Bloomsbury Publishing.
- Lew, Robert (2012): How can we make electronic dictionaries more effective? I: Sylviane Granger & Magali Paquot (eds.): *Electronic Lexicography*. Oxford: Oxford University Press, 343-361.
- Lorentzen, Henrik & Lars Trap-Jensen (2012): Nødvendigt, tilstrækkeligt, typisk? – nogle tanker om definitionspraksis. I: *LexicoNordica* 19, 79-101.
- Nilstun, Karina (2012): Fra dype betydningshierarkier til flattere struktur – et innblikk i revisjonen af *Norsk Riksmålsordbok*. I: *LexicoNordica* 19, 123-140.
- Ore, Christian-Emil S. & Åse Wetås (2014): Den store ordboka som digital ressurs: Norsk Ordbok i den digitale tidsalderen. I: *LexicoNordica* 21 (dette bind).
- Trap-Jensen, Lars (2010a): Access to Multiple Lexical Resources at a Stroke: Integrating Dictionary, Corpus and Wordnet Data. I: Sylviane Granger & Magali Paquot (eds.): *eLexicography in the 21st Century: New Challenges, New Applications. Proceedings of eLex 2009, Louvain-la-Neuve, 22-24 October 2009*. Louvain: Presses universitaires de Louvain, 295-302.
- Trap-Jensen, Lars (2010b): One, Two, Many: Customization and User Profiles in Internet Dictionaries. I: Anne Dykstra & Tanneke Schoonheim (eds.): *Proceedings of the XIV Euralex International Congress*. Leeuwarden, 1133-1143.

Henrik Hovmark
 lektor, ph.d.
 Ømålsordbogen
 Nordisk Forskningsinstitut
 Afdeling for Dialektforskning
 Københavns Universitet
 Njalsgade 136
 DK-2300 København S
 hovmark@hum.ku.dk

Pionjärprojekt om svenska översättare

Hans Landqvist

Svenskt översättarlexikon, Södertörns högskolebibliotek, 2009–,
<www.oversattarlexikon.se>.

1. Inledning

På senare år har forskare inom det översättningsvetenskapliga fältet intresserat sig allt mer för översättning av skönlitteratur ur ett diakront perspektiv. Två exempel är *The Oxford History of Literary Translation in English* (2006–) och *Suomennoskirjallisuuden historia* (2007). Det förstnämnda projektet har hittills resulterat i fyra omfattande volymer och den femte och sista ska föra projektet fram till år 2000 (Oxford University Press). Det senare projektet, redovisat i två digra band, tar sin utgångspunkt i medeltida översättningar av religiösa texter till finska, men när den kronologiska redogörelsen når fram till 1800-talets början ändras fokus till skönlitterära översättningar. Inom projektet undersöks såväl högprestige- som lågprestigelitteratur och därtill litteratur för barn (SKS Kirjat). För närvarande finns det inte någon svensk motsvarighet till de båda projekten; svensk översättningshistoria som en helhet väntar fortfarande på att bli utforskad och beskriven (Kleberg 2013:355).

Däremot föreligger olika studier av översättare med svenska som målspråk. Här kan nämnas Stina Hanssons undersökning av svensk latinöversättning under 1600-talet och Lars Wollins studier av översättning till svenska från medeltiden och framåt (Hansson 1982; se t.ex. Wollin 1998, 2007, 2013). Även Gunnel Furulands undersökning av skönlitterära häftesserier 1833–1851 ger viktiga upp-

lysningar om ett antal översättare bakom verk i serierna (Furuland 2007). Sedan 2009 finns det därtill en fritt tillgänglig och gratis internetkälla för den som söker upplysningar om skönlitterära översättare till svenska.

Databasen *Svenskt översättarlexikon* (i fortsättningen SÖL) är ”ett biobibliografiskt uppslagsverk med artiklar om svenska och finlandssvenska översättare från äldsta tid till idag” (Kleberg 2009:173). Enligt Kleberg (2013:357) kan SÖL vara ett ”internationellt pionjärprojekt” utan tidigare motsvarigheter. Initiativet till SÖL kom från Södertörns högskola och dess bibliotek, och därifrån bedrivs arbetet (se vidare Lembke 2006, Kleberg 2009, Kleberg 2013).

I det här sammanhanget har *lexikon* alltså betydelsen ’encyklopedi’, inte betydelsen ’ordbok’, vanligen preciserad till ’tvåspråkig ordbok’ (Svensén 2004:4). Härigenom skiljer sig SÖL från de lexikon som vanligen uppmärksammas i *LexicoNordica*. Trots detta kan SÖL vara av intresse även för de av tidskriftens läsare som i första hand intresserar sig för lexikologi och lexikografi, särskilt metalexikografi. Också i fråga om encyklopedier finns det intressanta frågor om insamling, urval och strukturering av ingående data.

I det följande ger jag först en kortfattad presentation av SÖL. Därefter diskuterar jag tre utvalda aspekter på databasen, nämligen den övergripande strukturen, artikelstrukturen i lexikonet samt (det nuvarande) urvalet av översättare i det. Till sist reflekterar jag över betydelsen av att en resurs som SÖL finns tillgänglig och utvecklas.

2. Presentation

SÖL är ett webbaserat lexikon, där artiklar tillkommer efterhand. Min granskning bygger på den version som är tillgänglig i början

av april 2014, där den senast publicerade artikeln gäller japansk litteratur på svenska (2014-04-02). Som utgångspunkt för min presentation fungerar figur 1, som visar SÖLs toppmeny med de sju rubrikerna Artiklar A–Ö, Språk, Kronologi, Om lexikonet, Medarbetare, Kontakt och Nya artiklar. Här återfinns också en länk till SÖLs Facebooksida och en sökfunktion i lexikonet.



Figur 1: SÖLs toppmeny.

Under Artiklar A–Ö i toppmenyn kan användaren dels söka bland samtliga publicerade artiklar i alfabetisk ordning, dels välja en av tre möjligheter: Översättare, Temaartiklar och Översättarpriser. Temaartiklarna gäller bland annat bokserier och tidskrifter, t.ex. den litterära tidskriften *All världens berättare* (1945–1956), och särskilda typer av översättning, t.ex. *Schlageröversättning*. (För hänvisningar inom SÖL använder jag i fortsättningen artikeltitlar i kursiv stil.) Dessutom behandlas också andra teman än de båda ovanstående, t.ex. *Översättarnas materiella villkor i Sverige*.

Under Språk i toppmenyn ges tre typer av uppgifter. Den första anger Originalspråk, vilket innebär att ”språk och översättare [listas] utifrån originalverkens språk. Eftersom verk [även] kan vara andrahandsöversättningar visar dessa listor inte nödvändigtvis vilket språk översättaren [i fråga] har översatt ifrån”. Antalet förtecknade originalspråk uppgår till 52. Detta innebär dock inte att översättarna i lexikonet tillsammans har översatt från 52 olika originalspråk till svenska. För det första förekommer *Katalanska* två gånger och norska tre gånger, under rubrikerna *Norska*, *Norska (Bokmål)* och *Norska (Nynorsk)* (se Fjeld Halvorsen 1997:47 om

bokmål och nynorsk som former av samma språk respektive olika språk). För det andra har alla förtecknade översättare faktiskt inte utfört alla sina översättningar från originalspråket i fråga, utan översättningar kan även ha gjorts via ett annat språk, dvs. andrahandsöversättning, indirekt översättning, kedjeöversättning alternativt reläöversättning (*Indirekt översättning*).

Ytterligare en svårighet när det gäller att fastställa grunden för en viss översättning är att det inte alltid är möjligt att klarlägga vilket språk som en översättare har utgått från. Ett illustrativt exempel erbjuder Anders Fredrik Dalin, 1806–1873, antagligen mer känd som språkman och lexikograf än som översättare (se Hannesdóttir 1998:462–467 om Dalin som lexikograf och om hans *Ordbok öfver svenska språket*). Under 1830- och 1840-talen var Dalin emellertid en flitig översättare, och bland hans översättningar återfinns verk med danska, engelska, franska, nederländska och tyska som originalspråk. Mestadels översatte Dalin verk av engelskspråkiga författare, men i artikeln om honom konstateras följande:

Det är svårt att fastställa vilket källspråk Dalin översatte från, men det är inte otroligt att flera av romanerna översattes från tyskan, som var det språk bildade personer behärskade bäst vid denna tid.

Den andra typen av upplysningar under Språk gäller vilket språk som fungerar som källspråk för en viss översättning till svenska. Som framgår ovan kan översättningar till svenska antingen vara gjorda från originalspråket eller via ett annat språk. Ett exempel på andrahandsöversättningar är Hugo Hultenbergs översättning av Mahatma Gandhis självbiografi. Den svenska översättningen är gjord från den engelska, inte från Gandhis originaltext på gujarati (*Hugo Hultenberg, 1870–1947*).

Den tredje typen av upplysningar under Språk gäller vilka an-

dra språk än svenska som ett visst verk är översatt till. Här förtecknas dock endast två översättare, två språk och två verk: Elmer Diktonius med en översättning från svenska till finska och Ellen Wester med en från svenska till franska (*Elmer Diktonius, 1896–1961, Ellen Wester, 1873–1930*). Upplysningar om översättningar från svenska till andra språk ges dock även i vissa artiklar om översättare. Ett exempel är översättningar till ungerska av en rad svenska 1900-talsförfattare, utförda av János Csatlós, 1920–1993.

Under Kronologi i toppmenyn kan användaren söka översättare verksamma under en vald tidsperiod, just nu från 1437 till 2014. Det är utgivningsåren för översättarnas verk som ligger till grund för kronologin, men den aktuella översättarens levnadsperiod ger en begränsning. Exempelvis förtecknas den nyss nämnda Ellen Wester, 1873–1930, inom det tidsspann som hennes födelse- och dödsår ger. Moderniserade versioner av några av Ellen Westers översättningar har dock utgivits fram till början av 2000-talet, och dessa förtecknas i Westers bibliografi.

Information om tankarna bakom lexikonet och arbetet med det ges under toppmenyns Om lexikonet. Här är det särskilt intressant med den information som ges om artiklar som är under arbete och de artiklar som redaktörerna hoppas ska kunna tillkomma. Bland artiklar under arbete finns sådana som ska ägnas författarna – och översättarna – Artur Lundqvist och Bertil Malmberg, medan några önskade artiklar gäller författarna och översättarna Karin Boye, Vilhelm Ekelund och August Strindberg.

Artikelförfattare i SÖL förtecknas under Medarbetare. Länkad information ger kortfattade uppgifter om varje författare samt uppgift om vilken artikel eller vilka artiklar som han eller hon har skrivit. Under Kontakt i toppmenyn ges möjlighet att skicka e-post till redaktionen. De senast publicerade artiklarna i SÖL återfinns under Nya artiklar, ordnade efter publiceringsdatum.

3. Diskussion

Som framgår i avsnitt 1 ovan väljer jag här att diskutera den övergripande strukturen för SÖL, artikelstrukturen i lexikonet och urvalet av översättare som (hittills) behandlas i det.

3.1. Övergripande struktur

I skrivande stund omfattar SÖL, enligt toppmenyns Alla artiklar A–Ö, totalt 193 artiklar. Målet är att lexikonet ska innehålla åtminstone 500 artiklar (Kleberg 2013:357). Utifrån Kleberg (2013:356) går arbetet för närvarande raskt framåt: i oktober 2013 ingick ungefär 150 artiklar i SÖL och i april 2014 närmar sig antalet alltså 200.

Utifrån informationen om de tre olika typerna av artiklar fördelar sig de 193 artiklarna i den här aktuella versionen av SÖL på 159 för Översättare, 16 för Temaartiklar och 18 för Översättarpriser. Kategorin Översättare är alltså den klart dominerande, vilket är förväntat med tanke på ambitionen att SÖL ska bidra till att ”fylla en stor lucka i kunskapen om den litterära översättningen, om översättarna och deras plats i svensk kulturhistoria” (Kleberg 2009:173). Men också de tematiska artiklarna och artiklarna om översättarpriser är intressanta. De förra ger möjligheter att behandla ett visst tema med såväl djup som bredd, t.ex. *Svenska Koranöversättningar*, och de senare lyfter fram uppmärksammade översättare, bl.a. mottagare av *Elsa Thulin-priset*.

Man kan kanske likna SÖLs webbplats med startsidan, toppmenyns sju möjligheter och de tre typerna av artiklar vid en ordboks *megastruktur* (jfr Svensén 2004:447–457). Mitt intryck är att webbplatsen är logiskt uppbyggd, vilket bidrar till att göra den lätt att använda. Också länkarna i artiklar, vilka kan ses som en form av *hänvisningar*, bidrar till att det är lätt att använda SÖL (jfr Svensén 2004:458–469 om *hänvisningsstruktur* i ordböcker). Exempelvis

finns det i artikeln *Sigurd Agrell, 1881–1937* länkar till artiklar om två andra översättare av ryska originalverk, nämligen Ellen Rydelius och Walborg Hedberg. De båda länkarna underlättar för läsare som intresserar sig för just översättningar av ryska originalverk till svenska under 1800-talets andra halva och 1900-talets första (se vidare Kleberg 2013:357 om korsreferenser i SÖL).

3.2. Artikelstruktur

Grundstrukturen för artiklarna i SÖL utgår från två huvuddelar: artikeltext och bibliografi. Artikeltexten ska ge biografiska upplysningar om den aktuella översättaren och en karakteristik av översättaren. Därtill kommer normalt en bild på översättaren (Kleberg 2009:178, 173). Möjligen kan man jämföra dessa grundlinjer med principerna för *mikrostrukturen* i en viss ordbok, i betydelsen 'artikelstruktur' (Svensén 2004:409f.). "Normalartikeln" i SÖL beräknas omfatta mellan 300 och 2.000 ord, och därtill bibliografien (Kleberg 2009:173).

När det gäller den bibliografiska informationen är ambitionen "fullständighet och exakthet beträffande den enskilda översättarens verk" och därtill referenser till böcker eller artiklar om översättaren – ifall sådana finns. Som verk räknas i första hand tryckta böcker men även andra bidrag kan vara aktuella: "framförda pjäsöversättningar, viktigare antologibidrag och tidskriftspublikationer och dessutom – i förekommande fall – opublicerade manuskript som finns i tillgängliga arkiv" (Kleberg 2009:179). Däremot redovisas normalt inte en översättares samtliga publikationer i tidningar, om nu sådana publikationer skulle vara aktuella.

Det är mer eller mindre svårt för artikelförfattare i SÖL att skaffa underlag till artiklar och bibliografier. Artiklar om vissa översättare ingår i *Svenskt biografiskt lexikon* (SBL), t.ex. Homeros-översättaren Erland Lagerlöf, medan andra är välkända också genom andra verksamheter än översättning (Kleberg 2009:180,

Wollin 2013:30). Det senare gäller bl.a. författarna Anna Maria Lenngren och Gustaf Fröding. Men för att producera en artikel krävs ofta grundforskning, både för att få fram uppgifter om översättaren som person och om översättarens verk. Detta gäller särskilt översättare som enbart ägnade sig åt att översätta och dessutom kanske inte signerade sina egna översättningar. Ett exempel på den senare situationen erbjuder ”Ellen Sofia Wester (1873–1930), som för att inte bli nonchalerad för sitt köns skull signerade sina översättningar ’E. Weer’” (Kleberg 2009:180).

Mitt intryck är att artikelstrukturen är funktionell och att den underlättar användningen av SÖL. Artikeltexterna är överlag väl strukturerade, och de innehåller ofta illustrativa exempel ur den aktuella översättarens produktion, sakkunnigt kommenterade. Exempelvis får användare av SÖL veta att den som läser Fredrik Niklas Bergs översättning av *Les trois mousquetaires* (1844) – *De tre muskötörerne. Historisk roman af Alexandre Dumas* (1846) kan ”märka en viss åtstramning av Dumas pratiga stil, särskilt gällande berättarens kommentarer av den egna berättelsen och läsningen av den”.

De bibliografiska upplysningarna i SÖLs artiklar är i många fall imponerande, och de har sannolikt krävt omfattande arbetsinsatser. Till exempel meddelas det att Ellen Wester översatte runt 80 böcker och därtill många noveller. I den tillhörande bibliografin upptas både ”Översättningar i bokform” och ”Bidrag” om tillsammans hela 103 poster (*Ellen Wester, 1873–1930*).

3.3. Urvalet av översättare

För produktionen av ordböcker är frågor om dataurval centrala. Det handlar dels om vilka lemman som ska ingå i en viss ordbok, *yttre selektion*, dels om vilka uppgifter som ska lämnas om de ingående lemmarna, *inre selektion* (Svensén 2004:79). I fråga om urvalet av översättare i SÖL, dvs. en form av yttre selektion, kon-

staterar Lars Kleberg, redaktör för lexikonet, att det krävs olika överväganden, men ”[i] princip kan alla betydelsefulla översättare till svenska språket – hittills endast avlidna personer – sägas förtjäna en plats i vårt lexikon” (Kleberg 2009:179). Översättarna som behandlas i SÖL, antingen de är kända eller okända, ska emellertid ha bidragit ”till den svenska översättningslitteraturen i någon eller några av dess former (skönlitteratur, essäistik, filosofi, populärvetenskap, barnböcker)” (Kleberg 2009:179). Urvalet får sägas vara brett, och det är väl enbart översättare av specialiserade facktexter som kan sägas bli uteslutna från SÖL, möjligen med undantag för översättare av facktexter inom filosofins område. Det breda urvalet kan illustreras med att SÖL, förutom artiklar om översättare med olika inriktningar, innehåller tematiska artiklar som ägnas såväl svensk bibelöversättning från äldsta tider till nutid som schlageröversättning från 1900-talet och framåt (*Svensk bibelöversättning, Schlageröversättning*).

Därtill har redaktionen för SÖL följt principen ovan om att enbart behandla avlidna författare (se dock nedan om Maibrit Westrin, f. 1924). Begränsningen går att förstå och att försvara, även om den innebär att exempelvis Thomas Warburton (f. 1918) för närvarande inte behandlas (Kleberg 2013:358). Warburton är kanske mest känd för att vara den förste som översatte James Joyces *Ulysseus* till svenska (*Odysseus*, 1946, reviderad 1993) (Warburton 2003:27–32, 135). I just det här fallet kompenseras dock ”bristen” i SÖL delvis genom att Warburton i sina memoarer ”[k]larsynt, elegant och roande beskriver [...] översättandets vedermödor och glädjeämnen” (Warburton 2003, baksidestexten).

Ambitionen är att SÖL ska behandla översättare av högprestigelitteratur, översättare av massproducerad lågprestigelitteratur och ”översättare som överskridit genregränserna” (Kleberg 2009:181). Den önskade spännvidden i fråga om översättare av högprestige- och lågprestigelitteratur kan illustreras med artiklarna om Erland Lagerlöf och Karl-Rune Östlund. Den förre översat-

te och utgav bland annat *Homeros' Odyssee* (1908) och *Homeros' Iliad I–II* (1912), medan den senare presenteras så här:

Trots betydande insatser som redaktör och översättare var Karl-Rune Östlund en obemärkt figur i det svenska litterära livet under efterkrigstiden. Obemärkt eftersom han under så gott som hela sitt verksamma liv som redaktionschef vid B. Wahlströms förlag ansvarade för den kioskboksutgivning som av kritiker och kultursidor betraktades som en skamfläck i det svenska läsutbudet. Obemärkt också för att han som översättare framförallt arbetade med de författare som betraktades som de värsta av de värsta. Östlund översatte 25 böcker av Mickey Spillane, 26 av Peter Cheyney, över 40 av James Hadley Chase. Det förvånar knappast att den som arbetade med böcker som aldrig recenserades och avfärdades som ”smutslitteratur” inte blev namnkunnig.

Det blir inte helt klart för mig vad Kleberg (2009:181) avser med ”översättare som överskridit genrer”. Är det fråga om översättare som inte enbart översätter skönlitteratur med stort litterärt värde utan även andra sorters skrivna texter: skönlitteratur av mindre god kvalitet, barnlitteratur, populärvetenskap osv.? Är det i stället fråga om översättare som översätter t.ex. sångtexter eller serietidningstexter, där det översatta verbala språket ska samspela med icke-verbala modaliteter? Däremot är det helt klart att SÖL innehåller artiklar som ger stöd för båda de här tolkningarna av Klebergs formulering om genreöverskridande översättare.

SÖL uppmärksammar rentav en och annan översättare som aldrig har haft fysisk gestalt. Mellan 1933 och 1967 utkom 35 romaner i översättning av Håkan Bergstedt, men någon person med detta namn har inte utfört översättningarna i fråga. Namnet användes nämligen ”som förlagspseudonym på Albert Bonniers förlag under ca trettio år i mitten av 1900-talet”. Pseudonymen

användes av två skäl; förlaget bearbetade en översättares verk så genomgripande att översättaren inte ville förknippas med den bearbetade översättningen eller den verkliga översättaren av ett verk ville inte uppträda under eget namn (*Håkan Bergstedt, 1933–1967*).

Ambitionen är att SÖL ska behandla översättare till svenska som har varit verksamma under alla tidsperioder (jfr Kleberg 2009:179), men hittills har främst översättare verksamma under ”äldre perioder och 1900-talet” uppmärksammas (Kleberg 2009:181). En sökning under toppmenyns Kronologi ger följande uppgifter om antalet översättare under de perioder som anges i tabell 1.

Tidsperiod	Antal översättare
1400–1699	3
1700–1799	15
1800–1899	51
1900–1999	122
2000–2014	22
Totalt	213

Tabell 1: Antal översättare under de angivna tidsperioderna utifrån en sökning under Kronologi i SÖL.

Allra först måste det sägas att det inte går att sätta likhetstecken mellan antalet översättare i tabell 1 och antalet översättare som behandlas i SÖL, eftersom en och samme översättare kan uppträda under två perioder. Detta gäller t.ex. Johan Gothenius (1700–1799, 1800–1899) liksom Mathilda Langlet (1800–1899, 1900–1999) (*Johan Gothenius, 1721–1809, Mathilda Langlet, 1832–1904*). Än mera tydligt är detta när det gäller de 22 översättarna för perioden 2000–2014: samtliga av översättarna har även varit verksamma under delar av 1900-talet.

Också med dessa reservationer är det klart att perioden 1900–1999 dominerar klart med 122 av 213 översättare i tabell 1 och att

perioden 1800–1899 kommer därefter med 51 belägg. Översättarna i SÖL verksamma före 1800-talet är få, och de som är verksamma före 1700-talet är blott två. Dessa är Jöns Budde, karakteriserad som ”den svenska senmedeltidens kanske flitigaste latinöversättare, utan tvivel också en av de skickligaste”, och Nicolaus Balk, som översatte fyra andliga skrifter och en fabelsamling från tyska. Därtill återfinns ”en ansevärd mängd språkprov” från hans översättningar i *Svenska Akademiens ordbok* (SAOB) (*Jöns Budde, 1437–1491, Nicolaus Balk, 1540–1611*).

De jämförelsevis få behandlade översättarna som är verksamma före 1800-talet är lite av en överraskning. Med tanke på exempelvis Hanssons studie av 1600-talets tryckta översättningslitteratur (1982) borde möjligheterna att behandla också ”äldre” översättare vara större av vad SÖL för tillfället visar. Förhoppningsvis kommer lexikonet därför att utvidgas med fler artiklar om översättare verksamma före 1800-talet.

Översättaryrket har ofta betraktats som ett yrke dominerat av kvinnor, särskilt under 1800-talet (jfr Furuland 2007:441f., 448). De hittills publicerade artiklarna i SÖL ger dock en lite annan bild. Utifrån tabell 1 ingår totalt 195 översättare i SÖL under perioden 1800–2014; 131 är män och 64 är kvinnor. Av de 51 översättarna verksamma 1800–1899 är 38 män, medan 78 av de 122 verksamma översättarna 1900–1999 är män. För perioden 2000–2014 är det 15 av 22 översättare som är män. Två av de kvinnliga översättarna i SÖL ägnas dock inte några egna artiklar. I stället porträtteras Lisbeth Renner och Maibrit Westrin tillsammans med sina respektive äkta män, som de samarbetade med i sitt arbete, i var sin temartikel (*Lisbeth och Louis Renner, P A Westrin*). Också med reservation för att en och samma översättare kan registreras under två tidsperioder är det, i alla fall hittills, en dominans för manliga översättare som ägnas egna artiklar i SÖL. Det kan finnas en eller flera anledningar till detta.

Mitt intryck är att ”högprestigelitteraturöversättare” ägnas fler artiklar än ”lågprestigelitteraturöversättare” i SÖL. Möjligen är det därtill vanligare att män översätter högprestigelitteratur än att kvinnor gör det, i alla fall fram till 1900-talet. Om det faktiskt är så, kan dessa omständigheter leda till att fler manliga än kvinnliga översättare behandlas i SÖL, i alla fall hittills. En annan möjlig förklaring är att det kan vara svårare för artikelförfattare att finna uppgifter om kvinnliga översättare och deras arbete:

Föga är känt om översättaren Carolina Wancke. På så sätt är hon typisk för en stor del av den svenska översättarkåren, i synnerhet den kvinnliga delen, under romanens genombrottsperiod i Sverige. (*Carolina Wancke, 1808–1879*)

Slutligen kan man ställa frågan hur väl ambitionen att behandla både svenska, dvs. sverigesvenska, och finlandssvenska översättare i SÖL har uppfyllts så här långt (jfr Kleberg 2009:173). På den här punkten kan det finnas en del att göra, eftersom en sökning på ”finlandssvensk” och ”finlandssvenska” ger endast 13 träffar. Dessa fördelar sig på 8 artiklar inom kategorin Översättare, 2 artiklar inom kategorin Temaartiklar samt 3 träffar inom kategorin Verk. De totalt 13 träffarna redovisas i tabell 2.

Översättare	Temaartiklar	Verk
Jöns Budde, 1437–1491	<i>Moderna poeter</i>	<i>Finland berättar: trettio noveller</i>
Otto Adolf Meurman, 1817–1850	<i>Om svenskt översättarlexikon</i>	<i>Två människobarn; Hiltu och Ragnar: en berättelse om två människobarn</i>
Sigrid Elmblad, 1860–1926		<i>Svart blir aldrig vitt</i>
Rafael Lindqvist, 1867–1952		
Elmer Diktonius, 1896–1961		
Cilla Johnson, 1911–2002		
Ralf Parland, 1914–1995		
Bo Carpelan, 1926–2011		

Tabell 2: Träffar på ”finlandssvensk” och ”finlandssvenska” i SÖL:
översättare, temaartiklar och verk.

En närmare granskning av de artiklar som förtecknas i tabell 2 gör att frågan om innebörden hos ”finlandssvensk” och ”finlandssvenska” i SÖL blir aktuell. Handlar det enbart om att en översättare eller ett verk ska ha en geografisk förbindelse med det område som sedan den sjätte december 1917 är det självständiga Finland eller kan också andra kriterier vara aktuella?

Vanligen anses det att en finlandssvensk identitet började utvecklas bland svensktalande i det dåvarande ryska storfurstendömet Finland mot slutet av 1800-talet och början av 1900-talet (Fougstedt & Westerholm). Utifrån detta perspektiv är det en anakronism att beteckna Jöns Budde och Otto Adolf Meurman som finlandssvenskar, medan man – med tanke på deras födelseår

– kan diskutera klassificeringen beträffande Sigrid Elmblad och Rafael Lindqvist.

Av översättarna i tabell 2 är det även värt att fundera över Cilla Johnson, som var född, uppvuxen och utbildad i Finland men som flyttade till Sverige där hon var verksam som översättare från danska, engelska och norska. Här är det intressant att jämföra med artikeln om Harry Järv. Han var född och uppvuxen i Finland och började sina universitetsstudier där. Men Harry Järv flyttade till Sverige, slutförde sina studier vid Uppsala universitet och översatte från originalspråken engelska, finska, franska, latin, tjeckiska och tyska. Men det blir inte någon träff på Harry Järv vid sökningen på ”finlandssvensk” och ”finlandssvenska” i SÖL (*Harry Järv, 1921–2009*). Däremot är bedömningen av de tre återstående översättarna i tabell 2, dvs. Elmer Diktonius, Ralf Parland och Bo Carpelan, som finlandssvenskar okontroversiell med tanke på deras levnadstid och deras geografiska hemvist.

Inte heller är klassificeringen av de båda temaartiklarna i tabell 2 kontroversiell. I artikeln *Moderna poeter* om bokserien med detta namn, som utgavs av Norstedts förlag mellan 1985 och 1989, nämns en översättning av den finlandssvenska författarinnan Mirjam Tuominen, och i artikeln *Om Svenskt översättarlexikon* meddelas att både svenska (= sverigesvenska) och finlandssvenska översättare ska behandlas i SÖL (även Kleberg 2009:173).

När det gäller de tre verken i tabell 2 innehåller *Finland berättar: trettio noveller* (1943) ett urval av noveller ursprungligen skrivna på finska respektive svenska av finländska författare, vilket gör att klassificeringen med ”finlandssvenska” kan motiveras. De återstående båda verken i tabellen är översättningar till svenska med delupplagor avsedda för den finländska marknaden.

Den svenska översättningen av Frans Eemil Sillanpääs *Hiltu ja Ragnar*, i ”bemyndigad översättning av Hjalmar Dahl”, gavs ut i Stockholm 1923 under titeln *Två människobarn* och i Helsing-

fors samma år under titeln *Hiltu och Ragnar: en berättelse om två människobarn*. På motsvarande sätt publicerades Ralf Parlands översättning av den ungerske författaren Illés Kaczérs roman *Ikongo nem hal meg* (översättningen gjord från den tyska översättningen *Roman eines Negers*) i både Sverige och Finland 1941. Utgivningen skedde under olika titlar: *Svart blir aldrig vitt* respektive *Svart erotik*. De båda översättarna Hjalmar Dahl och Ralf Parland är att betrakta som finlandssvenskar, men liksom i fråga om artikeln som behandlar Harry Järv blir det inte någon träff i artikeln *Hjalmar Dahl, 1891–1960* vid en sökning med ”finlandssvensk” och ”finlandssvenska”. De båda översättningarna ovan utkom i både Sverige och Finland. Frågan är om det är översättarna och/eller utgivningen under olika titlar som ligger till grund för klassificeringen ”finlandssvensk” i SÖL.

Sammanfattningsvis kan man fråga sig hur sökorden ”finlandssvensk” och ”finlandssvenska” definieras i SÖL och vilken information som en användare som intresserar sig för just finlandssvenska översättare får tillgång till genom att använda de båda sökorden.

4. Avslutning

Enligt min bedömning kan SÖL, i sin nuvarande och framtida form, vara till nytta och glädje för alla som intresserar sig för översättare till svenska och litterära verk som är översatta till svenska. Användare kan söka efter specifik information eller rentav läsa artiklar för nöjes skull. Genom SÖL börjar de litterära översättarna till svenska få en plats i offentligheten något som de länge har saknat. Lars Wollin (2013:31) beskriver översättarnas betydelse som att deras ”samlade insats i språk- och skriftkulturen är av en helt annan omfattning och dignitet än man ofta föreställer sig” (jfr Kleberg 2013:357). Vidare finns det inom översättningsvetenska-

pen ett allt större intresse för översättare som individer, verksamma i ett visst samhälle under en viss tid (Kleberg 2013:355f.), och SÖL kan bidra till fortsatt forskning. Lexikonet kan även fungera som förebild för andra motsvarande projekt. Sådana förbereds i Tyskland och Danmark, och fler kan tillkomma (Kleberg 2013:357).

Jag instämmer i Lars Klebergs förhoppning om framtida motsvarigheter till SÖL. Därtill skulle jag gärna se att lexikonet utvidgas så att det uppmärksammar även översättare med en tydlig inriktning på facklitteratur. Betydelsen av deras arbete är minst lika stor som betydelsen av de skönlitterära översättarnas, men facköversättarna får sägas vara ännu mer okända än sina skönlitterära kollegor (jfr Wollin 2013:28f.).

Litteratur

- Fjeld Halvorsen, Eyvind (1997): Norsk språk. I: Allan Karker, Birgitta Lindgren & Ståle Løvland (red.): *Nordens språk utgitt av Nordisk språksekretariat*. Oslo: Novus förlag, 47–72.
- Fougstedt & Westerholm = Gunnar Fougstedt & Johan Westerholm. Finlandssvenskar. I: *Uppslagsverket Finland – webbutgåva*. Helsingfors: Schildts, 2009–2012, Svenska Folkskolans Vänner (SFV), 2012–. <<http://www.uppslagsverket.fi>> (april 2014).
- Furuland, Gunnel (2007): *Romanen som vardagsvara. Förläggare, författare och skönlitterära häftesserier i Sverige 1833–1851 från Lars Johan Hierta till Albert Bonnier*. (Skrifter utgivna av Avdelningen för litteratursociologi vid Litteraturvetenskapliga institutionen i Uppsala 53.) Stockholm: LaGun.
- Hannedóttir, Anna Helga (1998): *Lexikografihistorisk spegel. Den enspråkiga svenska lexikografins utveckling ur den tvåspråkiga*. (Meijerbergs arkiv för svensk ordforskning 21.) Göteborg: Meijerbergs institut för svensk etymologisk forskning.

- Hansson, Stina (1982): ”Afsatt på svensko”. 1600-talets tryckta översättningslitteratur. (Skrifter utgivna av Litteraturvetenskapliga institutionen vid Göteborgs universitet 5.) Göteborg: Göteborgs universitet.
- Kleberg, Lars (2009): Att göra översättningar synliga. Om översättningshistoria och databasen Svenskt översättarlexikon. I: Erland Jansson (red.): *En bok om böcker och bibliotek tillägnad Louise Brunes*. Huddinge: Södertörns högskola, 173–186. <<http://sh.diva-portal.org/smash/get/diva2:213960/FULLTEXT01.pdf>> (april 2014).
- Kleberg, Lars (2013): Litterärt översättarseminarium och Svenskt översättarlexikon. Erfarenheter från Södertörns högskola. I: *Nordisk tidskrift för vetenskap, konst och industri utgiven av Letterstedtska föreningen* 2013:4, 349–358.
- Lembke, Jonas (utg.) (2006): *Svenskt översättarlexikon. Rapport om ett nytt projekt*. Huddinge: Södertörns högskola. <http://www.oversattarlexikon.se/filer/oversattarlexikon_rapport_2006.pdf> (april 2014).
- Oxford University Press = Oxford History of Literary Translation in English <<http://ukcatalogue.oup.com/category/academic/series/literature/hlte.do>> (april 2014).
- SKS Kirjat = Suomenoskirjallisuuden historia 1–2. <<https://kirjat.finlit.fi/index.php?showitem=1831>> (april 2014).
- Svensén, Bo (2004): *Handbok i lexikografi. Ordböcker och ordboksarbete i teori och praktik*. 2 omarb. o. utök. uppl. Stockholm: Norstedts Akademiska Förlag.
- SÖL = *Svenskt översättarlexikon*. Södertörns högskolebibliotek <<http://www.oversattarlexikon.se>> (april 2014).
- Warburton, Thomas (2003): *Efter 30 000 sidor. Från en översättares bord*. Helsingfors & Stockholm: Söderström & Co & Bokförlaget Atlantis.

- Wollin, Lars (1998): Från Heliga Birgitta till Barbara Cartland. Kring den svenska översättningens språkhistoria. I: Lars Kleberg (red.): *Med andra ord. Texter om litterär översättning*. Stockholm: Natur & Kultur, 62–91.
- Wollin, Lars (2007): Öfversättaren i det nittonde seklet. I: Yvonne Lindqvist (red.): *Gränslösa texter. Perspektiv på översättning*. (Ord och stil. Språkvårdssamfundets skrifter 38.) Uppsala: Hallgren & Fallgren, 161–194.
- Wollin, Lars (2013): Den svenska översättningskulturen – hur gammal är den? I: *Från ett språk till ett annat. Om översättning och tolkning*. (Språkrådets skrifter 15.) Stockholm: Norstedts, 22–31.

Hans Landqvist
docent, universitetslektor
Institutionen för svenska språket
Göteborgs universitet
Box 200
SE-405 30 Göteborg
hans.landqvist@svenska.gu.se

Skånsk dialektordbok i behändigt format

Caroline Sandström

Carl-Erik Lundbladh: *Skånska dialektord*. Skrifter utgivna av Dialekt- och ortnamnsarkivet i Lund 12. Lund: Institutet för språk och folkminnen, Dialekt- och ortnamnsarkivet i Lund 2012. 379 sidor. Pris: 200 SEK.

1. Inledning

Carl-Erik Lundbladh vid Dialekt- och ortnamnsarkivet i Lund har sammanställt och publicerat en ordbok där äldre skånska dialektord förklaras. Boken, som bär titeln *Skånska dialektord*, har ett tilltalande och behändigt format i normal romanstorlek 145 x 220 mm. Den är försedd med hårda pärmar i en dämpad, grå färgskala och pryds på framsidan av en stork – en fågel som är starkt förknippad med Skåne.

Ordboken är indelad i tre avdelningar som är markerade med grå färg vilket gör det lätt att orientera sig i verket. Avdelningarna *normaliserade former*, *dialektformer* och *betydelser* ger läsaren olika ingångar till det skånska dialektmaterialet. De mest utförliga artiklarna finns i avdelningen med normaliserade former, medan artiklarna i de två senare avdelningarna är kortare, och i många fall utformade som hänvisningsartiklar.

Ordboken är försedd med omtexter som förklarar uppbyggnaden och ger vägledning i hur ordboken kan användas. Under rubriken *Äldre och yngre skånska, talspråk med variation* beskrivs de skånska dialekterna och placeras in i tid och rum genom jämförelser med andra nordiska varieteter. I en avslutande text belyses bakgrunden till den syd- och västsvenska dialektdatabas som ord-

boken bygger på. Ett av ordbokens syften är att leda användarna vidare till det mer omfattande materialet om skånska dialekter som finns tillgängligt i den digitala databasen via nätet.

Texterna är föredömligt läsarvänligt skrivna. Lundbladh beskriver på ett enkelt och klart sätt principerna för hur dialektologer med språkhistorisk förankring brukar ordna sitt material. Av introduktionen framgår att det ofta är komplicerade språkhistoriska fakta som förklarar varför normaliserade dialektala uppslagsformer kan vara svåra att förstå och hitta fram till för en vanlig dialektalande person. Att på ett så elegant och tillgängligt sätt presentera de språkhistoriska förklaringarna är ett konststycke.

2. Urval

Lundbladh inleder med en introduktion där han redogör för de avgränsningar han gjort. Ordboken förklarar ord i äldre skånskt talspråk. Innehållet består huvudsakligen av ”ord som funnits i äldre skånska dialekter och har eller hade ett dialektalt uttal” (s. 6). Att den är avgränsad till äldre skånska dialektord utesluter dock inte att orden också kan ha förekommit i andra svenska dialekter.

I de inledande texterna ger Lundbladh en kort beskrivning av de skånska talspråksvarieteterna. Han diskuterar skånska dialekter och lokala ord i förhållande till både äldre språkskeden och dagens sydsvenska.

Skånska dialektord ger smakprov på det rika dialektmaterial som finns i Dialekt- och ortnamnsarkivet i Lund. Det är fråga om ord och uttryck som är förknippade med det äldre folkliga språkbruket och livet på landsbygden förr. Som framgått av studier som Lilja (1996) och Skott (2008) intresserade sig traditionsarkiven i första hand för landsbygden och den gamla bondekulturen; materialet i dialektarkiven är därför sällan representativt för hela befolkningen när det gäller kön, klass, ålder eller yrke.

De människor som levde då, tänkte och kände på många sätt precis som vi gör i dag. Lundbladhs beskriver livsvillkoren för dem som förr levde på landsbygden i Skåne (s. 13):

Man förälskade sig, gladde sig, oroade sig och förargade sig, man var elak och snäll, man gladde sig åt djur och natur, man trampade upp en stig i snön så barnen kunde komma till skolan och hade ett ordförråd för att beskriva både det yttre och det inre.

Urvalet av ord som tas upp i lemmalistan och de uppgifter som ges verkar lämpligt utförliga och ger ett enhetligt, balanserat och genomtänkt intryck. Man lägger särskilt märke till att både simplexord som *anka* 'ångra', *arp* 'mjäll, skorv', *fly* 'räcka, överräcka', *led* 'elak; ful', *lid* 'backe' och sammansatta ord som *ladufogde* 'ladugårdsförman', *lidköp* 'köpskål', *middagsbete* 'middagsrast' och *störvärdig* 'högdragen, kaxig' presenteras. Partikelverb behandlas så att partikeln anges i lemmat, t.ex. *skiftas till* 'turas om', *stubba på* 'bli försvagad' och på liknande sätt får vi information om reflexiva verb, t.ex. *bavsas sig* 'ångra sig', *begås sig* 'klara sig'. Alla ordklasser är väl representerade. Glädjande stort utrymme får adjektiv och adverb (*ackurat* 'precis, just', *akava* 'plötsligt', *görveten* 'nyfiken', *hickret* 'rädd, ängslig', *närmer* 'närmare; höger', *stacket* 'kort', *ständig* 'istadig', *vindad* 'narig'), som kan uppvisa en kreativ mångfald i dialekterna. Också pronomen (*hansa(s)* 'hans', *henna(s)* 'hennes', *hine* 'denne', *hocken* 'vilken') är väl företrädda bland de ord som tas upp. De har ofta nedslitna former och uppvisar stor variation i uttal vilket gör dem till en viktig grupp att behandla i en dialektordbok – i synnerhet eftersom pronomen liksom t.ex. förstärkannde adverb genom sina särpräglade former, som ofta har en lokalt begränsad användning, är uttrycksfulla och därför kan bära upp en dialektal identitet; jfr t.ex. Ivars (2003).

Dialekt har på svenska definierats som en "språkvariant som

talas av invånarna inom ett avgränsat geografiskt område” (Edlund 1990) eller som ”varje geografiskt avgränsad underart eller varietet av ett språk” (Dahlstedt 1982). Enligt dessa definitioner har dialekter således anknytning till geografiskt avgränsade områden.

För en användare som inte är förtrogen med Skåne och de skånska byarna, städerna och dialektsocknarna, samt indelningen av dem i dialektområden, hade det varit till stor hjälp om ordboken hade försetts med en schematisk karta över de olika områden där det skånska dialektmaterialet har samlats in. De orter orden kommer ifrån nämns över huvud taget sällan, trots att det av otexterna framgår att dialektskillnader förekommer och att variationen i Skåne är stor. För en utomstående kan det verka förbryllande att ordboken ger sken av att de ord som tas upp i den är allmängiltiga över hela området. För många ord anges visserligen varianter och sidoformer, men den variation som framträder i ordbokens artiklar kopplas inte konkret till systematiska dialektala skillnader mellan olika delar av Skåne.

3. Avdelningar

Genom att dela in ordboken i urskiljbara avdelningar har Lundblad valt att ge läsarna tre vägar in i de skånska dialektordens värld. Den mesta informationen om dialektorden ges som sagt i den första avdelningen med normaliserade former på omkring 130 sidor. Avdelningarna med dialektformer och betydelser har kortare artiklar och omfattar vardera drygt 100 sidor. Den tydliga tredelade strukturen hjälper användaren att lotsa sig fram till det dialektord eller den betydelse som han eller hon letar efter.

I varje avdelning finns grovt uppskattat omkring 2 000 uppslagsord. Men eftersom samma ord återkommer med olika ingångar i de tre avdelningarna kan ordboken möjligen kritiseras för att den med sina 379 sidor vid en första anblick ser ut att vara mer

omfångsrik och förklara ett större antal ord än den i själva verket gör. Det saknas dock explicita uppgifter om antalet uppslagsord i förklarande texter och på omslaget. I bokens introduktion nämns bara i anspråkslösa ordalag att den förklarar ”några ord i det äldre skånska talspråkliga ordförrådet” (s. 5). Men att omfånget inte är tydligt angivet uppvägs av att användaren via den överskådliga utformningen lätt kan hitta fram till den mångfald av ord och betydelser som ordboken erbjuder. Artiklarnas konsekventa upplägg och klara struktur visar att ordboken sammanställts av en erfaren lexikograf.

De mest ingående uppgifterna ges alltså i den första avdelningen (s. 22–151) där man kan söka orden i deras standardsvenska, normaliserade form. Den inleds med en modellartikel som visar hur artikeluppbyggnaden ser ut. De normaliserade uppslagsformerna belyser ordbildning och ger tydligare än dialektformerna ledtrådar till ordens ursprung. I artiklarna i denna avdelning ges utförliga uppgifter om uttal, ordklass, betydelse och exempel på hur ordet används. De valda exemplen är ofta stående uttryck, ramsor eller talesätt med ett målande innehåll. Det framgår av introduktionen att de bygger på autentiska språkprov ur dialektsamlingarna i Lund (s. 11). Artiklarna *gevunt* och *skumra* nedan belyser uppbyggnaden i avdelningen med normaliserade former.

GEVUNT, subst.

jevónt herravälde: *Den nye tocken ('tuppen') tog gevunten [jevónten] på den gamle.*

SKUMRA, verb

skåmra skymma: *Faselet vad dan sätter av, klockan är ju inte mer än fyra på dan och det börjar skumra [skåmra].*

Till variantformer för de normaliserade uppslagsorden hänvisas med kapitäl, såsom i exempelordet *haffeltag* nedan.

HAFFELTAG (Jfr **AFFELTAG**), subst.

haffeltag krafttag: *Han tog ett jäkla haffeltag.*

I många fall ges förutom variantformer ytterligare uppgifter om variation i böjning och uttal. Detta gäller i synnerhet verb som *hjälpa*, *lydas* 'lyssna', *löpa* 'springa', *vålla*, *växa* och *ärja* 'köra med årder'; jfr utformningen av artikeln över adjektivet *leret* nedan, som kan ha flera olika uttalsformer.

LERET, adj.

lared, **laired**, **lered** lerig:
Byxorna har blivit lereta
[*laireda*].

Den naturligaste vägen in till de skånska dialektorden för en dialektkunnig skåning är den avdelning som är ordnad utifrån dialektformer (s. 152–256). Artikeluppbyggnaden utgår från uppslagsformer som ligger nära det dialektala uttalet, i en del fall ges även uttalsvarianter. Också avdelningen med dialektformer inleds med en modellartikel. I artiklarna finns uppgifter om betydelse, där efter följer en hänvisning till uppslagsord i avdelningen för normaliserade former och slutligen anges ordklass. Att det även här ges en kortfattad betydelse underlättar användningen, eftersom det innebär att enbart de användare som vill få mer uppgifter om ordet behöver följa hänvisningen och gå till själva artikeln i första avdelningen. Uppbyggnaden av artiklarna i avdelningen med dialektformer belyses med utdrag av artiklarna *jöben* och *sjäckenerad*.

JÖBEN, JÖVEN

handfull, näve; vad som
kan omslutas av två händ-
er, Se **GÖPEN**, subst.

SJÄCKENERAD

chikanerad, utskämd. Se
CHIKANERAD, adj.

Den sista avdelningen som utgår från betydelse (s. 257–370) är avsedd att vara en praktisk hjälp att finna dialektord för en viss betydelse, dvs. man får veta hur en viss betydelse kan uttryckas i skånsk dialekt. Liksom de två föregående avdelningarna inleds också den med en illustrerande modellartikel. Därpå följer en kort text som förklarar principerna för artiklarnas struktur i avdelningen. Då uppslagsordet till sin betydelse är mer eller mindre synonymt med dialektordet står det ensamt versalt i artikelns inledning; se t.ex. verbet *lindra(s)* nedan. Av artikeln framträder att det i skånska kan finnas synonymer som *godna* och *lena* för verbet *lindra(s)*, och dessutom anges här uttalsformer för dessa.

LINDRA(S)

– *gona*. Se **GODNA**, verb
– *lena, lina, läna*. Se **LE-**
NA, verb.

En del av uppslagsorden kan ha en bestämmande del eller sidosform med kapitåler som tillägg. Exempel på uppslagsord med en bestämmande del är bl.a. *akterdel* (med tillägget *utrymme längst akterut i eka*), *maskstungen* (kompletterad med synonymen *maskäten*) och *urinsamling* (med tillägget *vid gödselstad*).

En betydelsegrupp som uppvisar många synonymer i ord-

boken är *andfådd* och *andtruten* som på skånska kan uttryckas med adjektiv som *ånnafårkaffad*, *ånnafårtryden*, *ånnafårtan* eller *stackaonad* i dialektform. I artiklarna *andfådd* och *andtruten* ges därtill hänvisningar till normaliserade former som *andaforkaffad*, *andafortruten*, *andafortagen* och *stackandad*. Också pronomen kan man lätt nå fram till via denna avdelning, t.ex. *vilken* som pekar vidare mot de dialektala pronomenvarianterna: *hocken*; *hytjen*, *sicken*.

Maträtter och arbetsredskap är en given grupp ord som förtjänar stort utrymme i en ordbok som förklarar dialektord. I avdelningen med betydelser finns bl.a. flera uppslagsord med *bakugn-*, *mjöl-*, *mjölk-*, *svagdricka-* och *säd-* som förled.

Det brukar i dialekter även finnas ett stort antal synonyma ord i fråga om djur och växter. Både flora och fauna är väl företrädde i *Skånska dialektord*, och ord ur dessa betydelsegrupper kommer verkligen till sin rätt i det upplägg som ordboken har. Genom att bläddra i de olika avdelningarna hittar man lätt fram till ord för växter och djur. När man inte fram via de normaliserade uppslagsorden kan man hitta det ord man söker via de dialektala formerna eller också via betydelserna. Utifrån betydelsedelen kan man t.ex. finna skånska dialektala ord för *höna*, *tupp*, *blåklint*, *gullviva*, *blåbär*, *lingon* eller *vinbär*.

För växten kaveldun finns i skånska många ord med förleden *dun-*. Dem kan man komma fram till via de normaliserade uppslagsorden *dunemannar*, *dunemark*, *dunemärg*. En användare som enbart känner till de dialektala uttalen kan kanske finna dem via de bekanta dialektformerna *dönemanna*, *dönemar*, *dönemark*. Olika varianter för denna växt kan dessutom hittas smidigt genom att ta *kaveldun* i betydelsedelen som utgångspunkt.

4. Avslutande ord

I introduktionen och förordet anger Lundbladh ramarna för sitt arbete med ordboken. Det framgår att arbetet gjorts medvetet, engagerat, med stor kunskap och insikt både i dialekterna i Skåne och i lexikografi, men att det också funnits sakkunniga personer som har kunnat stödja hans arbete. I introduktionen svarar Lundbladh på frågor som varför och för vem ordboken skrivits och han redogör också för materialet som ligger som grund för den. Tonen och innehållet i texten visar att han behärskar och kan förhålla sig till både den traditionella dialektforskningen med stadig förankring i språkhistoria och till dagens mer sociolingvistiskt inriktade dialektforskning. Materialet som ordboken bygger på är ett traditionellt insamlat arkivmaterial, där man strävat efter att samla in och återge det traditionella språkbruket. Men som ovan framgått presenteras materialet medvetet och genomtänkt på ett sätt som gör det lättillgängligt för moderna ordboksanvändare. En del vill kanske studera äldre uttal, böjning, konstruktioner och betydelser eller kanske bara påminnas om ord från förr i tiden. Lundbladh sammanfattar olika skäl till att intressera sig för äldre dialektord (s. 6):

Man kan vilja ta reda på vad ett uttryck i en äldre bouppteckning betyder eller vad som menas med ett talesätt man hörde en gång. Man kan vilja få bekräftat ett ord som man vakt erinrar sig, att det faktiskt har använts. Man kan vilja veta vad den som uttryckte sig så särpräglat lokalt faktiskt sa och menade.

Ordboken har en klar struktur och förklarar dialektorden på ett lättfattligt och inkluderande sätt. *Skånska dialektord* har alla möjligheter att nå ut till en bred publik, speciellt till dem som söker sina språkliga rötter i den skånska jorden.

Litteratur

- Dahlstedt, Karl-Hampus (1982): Talspråkets metamorfoser. I: Mats Thelander & Sture Allén (red.): *Talspråksforskning i Norden. Mål – material – metoder*. Lund: Studentlitteratur, 10–32.
- Edlund, Lars-Erik (1990): Uppslagsordet *dialekt*. I: *Nationalencyklopedin*. Stockholm: Bra Böcker.
- Ivars, Ann-Marie (2003): Lokalt och regionalt i svenskan i Finland. Tendenser i språkutvecklingen i norr och söder. I: Gunnstein Akselberg, Anne Marit Bødal & Helge Sandøy (red.): *Nordisk dialektologi*. Oslo: Novus forlag, 51–81.
- Lilja, Agneta (1996): *Föreställningen om den ideala uppteckningen. En studie av idé och praktik vid traditionssamlade arkiv – ett exempel från Uppsala 1914–1945*. Ser. B: 22. Uppsala: Dialekt- och folkminnesarkivet.
- Skott, Fredrik (2008): *Folkets minnen. Traditionsinsamling i idé och praktik 1919–1964*. Institutet för språk och folkminnen i samarbete med Göteborgs universitet. Göteborg: Institutet för språk och folkminnen.

Caroline Sandström
huvudredaktör, fil.dr
Ordbok över Finlands svenska folkmål
Institutet för de inhemska språken
Berggatan 24
FI-00100 Helsingfors
caroline.sandstrom@sprakinstitutet.fi

Nye kvalitetsverktøy for brukere av kinesisk i Skandinavia

Rune Svarverud

Lars Ragvald, Wai-Ling Ragvald, Susanna Björverud (redaktører): *Norstedts kinesisk-svenska ordbok*. Stockholm: Norstedts 2012. 1064 sider, 58 000 ord og fraser, 1. utgave. Pris: 460 SEK.

Lars Ragvald, Wai-Ling Ragvald, Susanna Björverud (redaktører): *Norstedts kinesisk-svenska ordbok*. Stockholm: Norstedts 2013. App versjon 1.0, krever iOS 5.0 eller senere. Kompatibel med iPhone, iPad og iPod touch. 58 000 ord og fraser. Pris: 299 SEK.

1. Historikk og funksjonalitet

Norstedts kinesisk-svenska ordbok (NKSO) er den første omfattende kinesisk-svenske ordbok som er laget, og bygger på det langt mest omfangsrike og ambisiøse tospråklige kinesiske ordboksprosjekt i Skandinavia.¹ Grunnlagsarbeidet for den foreliggende ordboken og appen startet ved Lunds universitet tidlig på 1990-tallet under ledelse av Lars Ragvald og med Magnus Nordenhake som medarbeider. Arbeidet med å samle oppslagsord og legge alt inn i en database gikk raskt fremover gjennom 1990-tallet, noe som førte til utgivelsen av en mindre kinesisk-svensk papiordbok med drøye 8000 oppslagsord i 2000 (Ragvald 2000). Men etter at fi-

1 I Norge ble et større ordboksprosjekt startet ved Universitetet i Oslo i 1988, noe som førte frem til utgivelse av papiordboken *Kinesisk-norsk ordbok* (Rekkedal et al. 1997) og senere den reviderte utgaven *Stor kinesisk-norsk ordbok* (Rekkedal et al. 2005), begge med ca. 80 000 oppslagsord. Dette norske prosjektet har ikke ført frem til en kommersialisert elektronisk ordbok på nett eller som app. Andre mindre omfattende tospråklige ordbøker mellom kinesisk og norsk er også utgitt. I Danmark har det ikke vært gjennomført noe større akademisk ordboksprosjekt til kinesisk.

nansieringen av prosjektet stanset opp, gikk arbeidet med databasen fremover i kun sakte tempo frem til 2007. Fra 2007 og frem til ferdigstilling i 2012 skjøt arbeidet på ny fart gjennom en utvidet stab bestående av, foruten Lars Ragvald selv, Wai-Ling Ragvald og Susanna Björverud. Redaksjonsarbeidet har i hovedsak blitt utført av Lars og Wai-Ling Ragvald, mens Susanna Björverud har hatt et hovedansvar for databasen og de tekniske løsningene som ligger til grunn for både papirordboken og appen.

Databasen består per i dag av ca. 90 000 ord og fraser. Av dette er det gjort et utvalg på drøyt 58 000 ord og fraser som har funnet veien inn i app og papirordbok. Papirordboken er på 1064 sider, appen foreløpig kun tilgjengelig på Apples plattform for iPhone, iPad og iPod touch, ikke for plattformene Android eller Windows. Hovedfokus i valg av ord og fraser har, ifølge redaktørene selv, ligget på det som kan kalles det grunnleggende ordforrådet i moderne kinesisk, komplettert med et stort antall fagtermer samt ord og fraser fra de siste tiårs utvikling av det kinesiske språket. I denne første versjon/utgave av ordboken (app versjon 1.0, papirordbokens første utgave) er kun de forenklete tegnvarianter (*jiǎntǐzì* 简体字) tatt med. Man kan altså ikke slå opp eller søke etter de uforenklete/tradisjonelle tegnvarianter (*fántǐzì* 繁體字) i denne første versjonen/utgaven.

Ordboken er bygget opp som en kinesisk-svensk ordbok, slik at de primære oppslagsordene er de kinesiske. Disse oppslagsordene er så oversatt eller forklart ved hjelp av ord eller fraser på svensk. I appen er det mulig også å søke etter svenske ord, men man vil da kun få treff i de svenske oversettelser eller forklaringer av kinesiske ord og fraser. Hvert oppslagsord på kinesisk kan bestå av enkelttegn eller av kinesiske ord sammensatt av to eller flere kinesiske skrifttegn. Det kinesiske oppslagsordet med kinesiske skrifttegn er etterfulgt av tegnets eller tegnenes uttale angitt ved hjelp av standard pinyin-transkripsjon (*Hanyu pinyin*) inkludert angivelse av tonetegn (tonem). Hvert oppslagsord kan ha én eller

flere betydninger. For oppslagsord med flere betydninger er disse adskilt med tallmarkeringer i tillegg til at betydninger tilhørende ulike ordklasser er adskilt med en prikk (•). Hver betydning er angitt med ordklasse og/eller fagområde, stilverdi eller annen informasjon. Noen av oppslagsordene er etterfulgt av eksempel-setninger (kinesiske tegn, transkripsjon og oversettelse), men disse er av plasshensyn redusert til et mindre antall.

Kinesisk-svensk ordbok	
Q 汉语, pinyin, pin2yin2, svenska	
呵	ā interj
啊	ā interj ja sā
阿	ā pref
阿飞	āfēi subst huligan
阿哥	āgē subst storebror
阿訇	āhōng subst imam
阿拉伯	Ālābó attr adj Arab-
阿拉伯文	Ālābówén subst Arabiska
阿拉伯语	Ālābóyǔ subst Arabiska
阿姨	āyí subst moster

Figur 1: Appens søkeside for både kinesiske tegn, pinyin-uttale og svensk oversettelse.

Ordklasser er til dels problematiske kategorier på kinesisk, men denne ordboken har allikevel valgt å innlemme informasjon om ordklasser knyttet til svensk oversettelse/forklaring av kinesiske ord som en hjelp for brukeren.² I tilfeller hvor det er vanskelig å

2 Kinesiske leksem kan som oftest ikke knyttes til en fast ordklasse, men vil opptre i en setning i en funksjon knyttet til ordklassekategorier. Ord-

angi oppslagsordets og/eller betydningens ordklasse, angis disse som ukategoriserte (okat). Av ordklasser som angis, finner vi: adjektiv, adverb, attributivt adjektiv, fast frase, interjeksjon, klassifikasjonsmerke/målord, konjunksjon, konstruksjon, lokativ, morfologisk del, onomatopoetikon, grammatisk partikkel, prefiks, preposisjon, pronomen, tallord, slektsnavn, substantiv, substantiv/verb, suffiks, verb-objekt-konstruksjon og verb.

Som kjent skrives kinesisk med egne skrifttegn og ikke med latinske eller andre bokstavsystemer. Både i appen og i papirordboken lar det seg gjøre å søke både på kinesiske skrifttegn og på transkripsjon. Gitt de ulike mediene, er søkefunksjonene til en viss grad forskjellige.



Figur 2: Appens radikalindeks for oppslag etter radikal og fonetikum.

klasseangivelser er derfor i denne ordboken, som i de fleste tospråklige kinesiske ordbøker, knyttet til de svenske oversettelser av eller forklaringer av kinesiske ord snarere enn til det kinesiske leksemet selv.

Søk etter kinesiske tegn i appen: Man kan søke etter kinesiske tegn på tre ulike måter. Det enkleste er å kopiere et oppslagsord fra en elektronisk tekst og lime dette direkte inn i søkefeltet i appen. Hvis det ikke lar seg gjøre, kan man søke i ordboken ved hjelp av håndskriftidentifisering. Dette krever at man har installert kinesisk som input-system med håndskriftidentifisering på sin iPhone, iPad eller iPod touch, og man må være godt kjent med prinsippene for strekrekkefølge i kinesiske tegn for at man skal få treff. Hvis ingen av disse to metodene lar seg gjennomføre, kan man søke ved hjelp av radikal og fonetikum. Alle tegn i ordboken er kategorisert etter 229 ulike radikaler. Man identifiserer derfor først det aktuelle tegnets radikal og teller så antall streker i den resterende delen av tegnet (fonetikum). Dette er en omstendelig prosess for å identifisere tegn i ordboken, men er et alternativ hvis de to foregående metoder ikke fører frem.

Søk etter uttale i appen: Kinesiske oppslagsord kan søkes opp etter uttale ifølge pinyin-systemet. Man kan søke på enkelttegn eller tegnsammensetninger, de siste ved å skrive pinyin for tegnene i oppslagsordet sammen uten mellomrom (for de fleste ords vedkommende – se diskusjon under). For mer presise søk kan man også søke ved å angi tone/tonem med tallangivelse (1, 2, 3 eller 4) etter hver stavelse. Det lar seg gjøre å søke etter *ü* gjennom å skrive *y*.

Søk etter kinesiske tegn og uttale i papirordboken: Papirordboken er organisert slik de fleste tospråklige kinesiske ordbøker er bygget opp. Oppslagsordene er oppført under første tegn i ordet, og alle disse enkelttegn med deres påfølgende ordsammensetninger er sortert alfabetisk etter pinyin, inkludert tonem, fra *ā* til *zù*. Hvis man ikke kjenner uttale av første tegn i ordet, lar første tegn i ordet seg søke opp etter radikal (de samme 229) og fonetikum etter samme prinsipper som i appen.

Søk etter svensk oversettelse i appen: Man kan søke etter ord eller fraser på svensk i appen og vil få treff på kinesiske oppslagsord som inneholder det søkte ord eller den søkte frasen. Man skal dog

være klar over at denne funksjonen ikke vil være å betrakte som en fullverdig svensk-kinesisk ordbok, men kun som en mulighet til å søke etter svenske oversettelser og forklaringer til kinesiske oppslagsord. Man kan selvfølgelig ikke på samme måte søke etter svenske oversettelser til kinesiske oppslagsord i papirordboken.

Appen følges av en rekke tilleggsfunksjoner og appendikser, som et meget godt kart over Kina, en oversikt over det tradisjonelle tidsregningssystemet (*Tiāngān Dizhī*), de 24 solperiodene, de 28 månehusene, de kinesiske dynastier og tidsepoker i kinesisk historie, samt de 8 trigrammene fra boken *Yijing* (*Forvandlingens bok*). Papirordboken har alle disse tilleggene, bortsett fra kartet over Kina.

Appen inneholder også et forord (Förord), som forklarer ordbokens historikk og institusjonelle bakgrunn, og en introduksjon til bruk av ordboken (Anvisningar). Denne siste er svært nyttig for brukere som ikke er godt kjent med bruk av app-ordbøker. Videre følger en oversikt over forkortelser brukt i ordboken (Förkortningar i ordboken), hvor også de grammatiske kategoriene er forklart. Et vedlegg om ordbokens oppbygning (Ordbokens oppbyggnad) forklarer brukeren hvordan ordboks-appen er bygget opp. Blant tilleggene i appen finner vi også en kort beskrivelse (Om ordboken) av hvem forfatterne er, hvem som har utgitt ordboken, når den er kommet ut og hvilken versjon det er. Alle disse tilleggene inneholder nyttige og brukervennlige forklaringer til brukeren, selv om organiseringen av informasjon og inndelingen i disse 5 dokumenter kunne vært gjort noe mer intuitiv.

Papirordboken inneholder i hovedsak de samme innledninger og forklaringer, men da knyttet til organiseringen av ordboken på papir. Det forekommer noe enklere å sette seg inn i dette når man får det presentert på større boksider på papir og ikke i en rekke mindre dokumenter på en liten skjerm.

2. Ordtilfang

Som alle språk, er også det kinesiske i stadig endring. Språkets grunnleggende oppbygging og syntaks endrer seg med lav takt, men mange nye ord kommer og går raskt. Ofte skapes nye ord i språket i nærkontakt med andre språk, hvor engelsk også i Øst-Asia er et språk med stor påvirkningskraft. Japansk er et annet språk som både har hatt og fortsatt har stor påvirkningskraft på kinesisk. I redaksjonsarbeidet med en ordbok er det ofte et vanskelig spørsmål når man skal regne med at et ord har stabilisert seg nok i språkets leksikon til at det fortjener plass i ordboken. Samtidig vil ikke en ordbok som NKSO ha noen normgivende funksjon, men først og fremst være et hjelpemiddel til brukere og studenter av språket. En slik funksjon berger for at redaktørene bør ha en relativt fleksibel politikk med hensyn til hvilke ord som kan antas å ha oppnådd en så frekvent og allmenn bruk at man som bruker av ordboken vil kunne ønske å slå disse opp. Ordboken dekker godt det grunnleggende ordforrådet i moderne kinesisk. Men siden NKSO hevder nettopp å ha tatt inn begreper og nyord fra de siste tiårenes sosiale endringer i Kina, kan det være interessant å undersøke i hvor stor grad gjengse ord i dagligtalen i Kina er å gjenfinne i ordboken.

En viktig trend i Kina de siste tiårene er endringer i familiemønstre og migrasjon av unge arbeidere fra landsbygda og inn til byene. Dette har også ført til at det er blitt dannet en rekke nye og i dag ofte brukte ord i det kinesiske språket. Nye relasjoner i familien, mellom venner og mellom menn og kvinner har ført til en rekke nyord, som for eksempel *xiǎosān* 小三 'tredjepart i et forhold, ofte brukt om kvinner som har et forhold til en gift mann', *zháinán/zháinǚ* 宅男/宅女 'unge menn og kvinner som holder seg mest hjemme, gjør seg liten flid med eget utseende og har liten sosial omgang med andre', *shèngnǚ* 剩女 'godt voksen kvinne som har forblitt ugift', *yǐzú* 蚁族 'nylig uteksaminerte, unge kinesere

som har lav inntekt og bor som maur sammen i trange og enkle boliger i de store byene’, *bāshíhòu/jiǔshíhòu* 80后/90后 ‘generasjonen av unge kinesere født på 1980- og 1990-tallet, som har helt andre livserfaringer og -forventninger enn kinesere som har opplevd for eksempel kulturrevolusjonen på 1960- og 1970-tallet’, *fènqīng* 愤青 ‘misfornøyd ungdom’, *kěnlǎozú* 啃老族 ‘unge kinesere i arbeidsfør alder som gjør lite eller intet for å finne seg arbeid og i stedet lever av sine foreldres inntekt’. En rekke nye begreper for homofile/lesbiske relasjoner er også blitt trendy, for eksempel at det politisk befengte *tóngzhì* 同志 ‘kamerat’ er blitt tatt i bruk for mannlige seksualpartnere. Arbeidsmigrasjon i Kina har også ført med seg en rekke nye ord og termer for denne trenden, som for eksempel *nóngmíngōng* 农民工 ‘bonde som har tatt midlertidig arbeid, som oftest i en større by’ og *dǎgōngmèi* 打工妹 ‘ung kvinne som tar midlertidig arbeid, ofte i kles- og skoindustrien’. De siste tiårenes kapitalistiske økonomiske politikk i Kina har skapt store forskjeller mellom fattig og rik, og især er det en generasjon av barn av rike forretningsfolk eller embetsmenn som har vokst opp under svært gode betingelser. Disse refereres ofte til som *fù’èrdài* 富二代 ‘rik annengenerasjon’ og *guān’èrdài* 官二代 ‘embetsmenns annengenerasjon’. Dessverre har ingen av disse ord og betydninger funnet veien til NKSO.

同义词 *tóngyìcí* *subst* språkv. **synonym**
 同音词 *tóngyīncí* *subst* språkv. **homofon**
 同志 *tóngzhì* *subst* **kamrat** numera framför allt om
 partimedlemmar
 同舟共济 *tóngzhōu gòngjì* *fras* **samarbeta mot ett
 gemensamt mål**

Figur 3: Fire artikler i papirutgaven av NKSO.

Helt nye former for kommunikasjon over nettet og i nye sosiale medier i Kina har også ført til at nye sett med ord og begreper er tatt i bruk i kinesisk, men disse er tilsynelatende dessverre heller ikke å finne i NKSO. Blant disse har *qīn* 亲 (kjære) blitt vanlig som tiltaleform på Kinas mikroblogg, Weibo 微博. Ordet *wēibó* i betydningen 'mikroblogg' er heller ikke å finne i ordboken. Blogging er også en vanlig sosial plattform for den yngre generasjon i Kina, kalt *bókè* 博客 for 'blogg' og *bókèwén* 博客文 for 'bloggtekst'. Som kjent er det viktig å ha 'fans' i sosiale medier, disse kalles *fěnsī* 粉丝, som er et fonetisk lån fra engelsk. Disse betydningene er heller ikke å finne forklart i Norstedts ordbok. Når man lager 'show' av noe, enten det er i nye eller mer tradisjonelle medier, brukes ordet *xiù* 秀 ofte som substantiv (også fonetisk lån). Denne betydningen er ikke med i ordboken, selv om verb-objekt-konstruksjonen *zuòxiù* 作秀 'lage show', er kommet med. Hvis man går for langt i sine uttalelser i Kinas nye internett-medier, kan man få et ublidt møte med politimakten. Dette kalles 'å bli invitert til å drikke te' – *qǐng hēchá* 请喝茶, et uttrykk som dessverre heller ikke har kommet med i ordboken. Begrepet *hūyou* 忽悠, som opprinnelig betyr 'å vingle/vakle', er nå blitt et mye brukt ord for å 'lure noen' eller 'føre noen bak lyset', både i dagligtale og på Weibo og andre sosiale medier. Ingen betydninger av dette ordet er å finne i Norstedts ordbok.

Som i alle språk, finnes det et sjikt av tildels «uanstendige» og seksuelt ladete ord og begreper som sjelden finner veien inn i ordbøker. Norstedts ordbok er ikke å klandre for å ha utelatt de fleste av disse, selv om man som student av språket nok ofte gjerne skulle kunnet slå dem opp. Men mer vanlig brukte og mindre «uanstendige» ord, som for eksempel *niú* 牛 for 'arrogant', 'oppblåst' og *tāmāde* 他妈的 i betydningen 'faen', kunne godt vært tatt med. *Gěilì* 给力, som et vanlig ord i dag i betydningen 'gi alt', 'legge seg i selen', burde også vært tatt med. Et annet ord som har blitt helt sentralt i den kinesiske debatten de siste årene, er *wùmái* 雾霾 for 'smog', som man ikke finner i ordboken. Når man søker på 'smog' i

Norstedts, finner man *yānwù* 烟雾, men det er altså ikke det ordet som er blitt standard for den smogen som nå plager befolkningen i alle større kinesiske byer.

Egennavn har NKSO i hovedsak unngått, men de har tatt med navn på de fleste større land i verden, som Frankrike, Tyskland, Sveits, Italia, Belgia, Holland, Portugal, England, Sverige, Finland, Danmark, USA, Canada, Japan, Korea (dog ikke de mest brukte navnene på Sør-Korea, Hánguó 韩国 eller Nán-Hán 南韓), Vietnam, Russland (dog ikke det ofte brukte navnet i dagligtale *É'guó* 俄国). Merkelig nok er også ganske mange land som man kunne ha forventet å finne i denne ordboken, utelatt, for eksempel Spania, Polen, Norge, Island, Irland, Egypt, Brasil og Argentina.

Det skrives intet i introduksjonen til ordboken om forholdet til det klassiske kinesiske språket, som jo må kunne sies fortsatt å være levende som del av moderne uttrykksformer. Et søk på noen av de mest vanlige klassiske betydninger av grunnleggende kinesiske tegn viser at Norstedts ordbok har valgt å inkludere de viktigste klassiske betydninger av enkelttegn i kinesisk. Det faktum at ordboken kun angir de forenklete tegn og ikke de tradisjonelle formene, skaper dog noen utfordringer i det å skille betydninger der forenklingen av tegn har slått sammen to ulike klassiske tegn med ulik betydning.

Ordtilfanget i NKSO dekker godt det grunnleggende ordforrådet i det moderne kinesiske språket, men som bruker av boken skulle man gjerne sett at redaktørene hadde gjort et noe grundigere arbeid med å innarbeide et utvalg av de mest brukte nye ord og termer som er i flittig bruk i det kinesiske språket i dag. De fleste brukerne av ordboken, og især brukerne av appen på iPhone, iPad eller iPod touch, vil støte på mange av disse ordene i sin omgang med språket, og da ville det vært en fordel om en oppdatert ordbok, og især på en fleksibel plattform som en app, i større grad hadde tatt denne delen av det samtidige kinesiske leksikon i betraktning.

3. Grammatisk informasjon, eksempelsetninger og transkripsjon

Den grammatiske informasjonen som gis i Norstedts ordbok, er rettet mot brukerne av ordboken snarere enn utarbeidet som et strengt analytisk verktøy i seg selv. En slik pragmatisk innretning av informasjonen er til brukernes fordel og gir nyttig tilleggsinformasjon om hvert oppslagsord. Dessverre har brukerveiledningen som gis til ordboken helt rett når det skrives at antallet eksempelsetninger er redusert til et mindre antall. I en papirordbok kan man lett forstå et slikt valg pga. omkostningene ved å produsere en svært omfangsrik papirordbok, men den samme argumentasjonen virker noe mer malplassert for brukere av appen. Kanskje kan brukerne håpe på at antallet eksempelsetninger som finner veien inn i appen blir større ved neste versjon, især hvis det er slik at disse allerede er å finne i databasen som ligger til grunn for alt dette materialet.

Transkripsjonene følger standard pinyin. Ved oppslagsord med to stavelser (to tegn) skrives pinyin som hovedregel sammen uten mellomrom. Tostavellesord som er angitt som verb-objekt-konstruksjoner (v/o), skrives dog ofte med mellomrom mellom de to stavelsene, men ikke alltid. Det virker ikke som om redaktørene har gjennomført en streng linje på dette feltet. Ved ord eller uttrykk med fire stavelser (fire tegn) skrives som regel de to første sammen uten mellomrom, så et mellomrom, og så de to siste stavelser sammen uten mellomrom, men også her finner brukeren variasjoner som ikke uten videre lar seg forklare. Nå er ikke dette med sammenskrivingsregler for pinyin-stavelser helt avgjørende for å vite hvordan man skal uttale et gitt oppslagsord. Men som vi skal se under, har denne mangel på gjennomført system visse konsekvenser for hvilke treff man finner ved søk på ord etter uttale i appen.

4. Søkefunksjoner

Søkning etter kinesisk enkelttegn og tegnsammensetninger etter håndskriftmetoden i appen fungerer svært godt, gitt at man behersker de grunnleggende regler og prinsipper for skriving av kinesiske tegn. Søk etter tegn gjennom radikalindeks og antall streker både i appen og i papirordboken fungerer også slik man er vant til i slike ordbøker. Som tidligere nevnt, kan man også søke etter ord og ordsammensetninger i den svenske forklaringen til de kinesiske oppslagsord i appen. Dette fungerer også utmerket så lenge man er klar over at denne funksjonen ikke gir resultater som i en fullverdig svensk-kinesisk ordbok. Det er dog en ulempe i noen tilfeller at man søker pinyin uten toner og svensk oversettelse i samme søkefelt, som for eksempel ved søk etter ord som *han*, *ni*, *la*, *ting* osv. Ved slike søk får man treff både på kinesiske tegn som uttales *han*, *ni*, *la*, *ting* og på svenske oversettelser som har disse ordene i seg.



Figur 4: Søk etter pinyin *yuyan* (uten tonem) gir treff på fire ulike kinesiske leksem.

Når det gjelder søk etter kinesiske oppslagsord ved hjelp av pinyin-transkripsjon i appen, møter brukeren på noen flere utfordringer. I hovedsak fungerer systemet utmerket, og funksjonen at man kan søke både med og uten toner i appen øker anvendeligheten av denne sett i forhold til papirordboken. Men ved bruk av store forbokstaver i egennavn har redaktørene skapt en utfordring for pinyin-søkefunksjonen. Redaktørene har i hovedsak valgt å angi egennavn med stor forbokstav i pinyin, så både navn på land og på språk er som en hovedregel skrevet med stor forbokstav i pinyin. Dette er dog ikke systematisk gjennomført, ved at tysk språk for eksempel er skrevet *Déyǔ* med stor forbokstav, det samme er fransk, russisk, og japansk, mens engelsk språk er skrevet *yīngyǔ* og England *yīngguó*, begge med liten forbokstav. Frankrike er skrevet med liten forbokstav i pinyin, mens Tyskland er skrevet med stor forbokstav. Dette ville kanskje ikke hatt så store konsekvenser hvis det ikke var for det faktum at søkefunksjonen er sensitiv for forskjellen på store og små bokstaver i pinyin. La oss ta Frankrike og fransk språk som eksempler. Hvis man søker på *Faguo* med stor forbokstav (uten toner), får man ingen treff, men når man søker på *faguo* med liten forbokstav, finner man det kinesiske oppslagsordet for Frankrike. Hvis man derimot søker på *fayu* med liten forbokstav, får man faktisk treff på fransk språk selv om dette er skrevet med stor forbokstav i oppslagsordet, mens om man søker på *Fayu* med stor forbokstav, som jo er slik det er skrevet i oppslagsordet, får man ingen treff på fransk språk. Skriver man inn tonene, får man treff på begge. Brukeren treffer på lignende utfordringer med store og små bokstaver på flere av disse egennavnene. Her har altså redaktørene en funksjon som klart bør forbedres til neste versjon av appen.

De samme søkeutfordringene møter vi på når vi skal søke etter ord som redaktørene har kategorisert som verb-objekt-konstruksjoner (v/o). Som nevnt over, varierer praksis med hensyn til om pinyin er skrevet sammen eller delt mellom verb og objekt. Hvis

redaktørene har valgt å dele pinyin for et v/o-ord, får man ingen treff i pinyin-søket om man skriver pinyin for sammensetningen uten mellomrom. Man får på samme måte heller ikke treff om man søker etter delt pinyin for et v/o-ord som redaktørene har valgt å skrive samlet. La oss ta et eksempel for å illustrere dette. Hvis man ønsker å søke opp et ord som består av de to stavelsene *da* og *gong*, kan man først søke ved hjelp av pinyin på sammensetningen *dagong* uten mellomrom, og man får treff på ordet 大功, som er oversatt med 'förtjänstfull insats' (pluss at man får treff på et par fire-tegns-uttrykk som ikke er relevante her). Søker man på *da gong* med delt pinyin, får man treff på v/o-ordet 打工, som er oversatt med 'utföra manuellt korttidsarbeite ...'. Dette siste ordet kunne nok like gjerne vært skrevet uten mellomrom i pinyin, og det vil under alle omstendigheter være vanskelig for brukeren å vite hva som er rett søkemåte – især fordi brukeren kanskje ikke alltid vil reflektere over at det ord man søker, er et v/o-ord. Hvis man som bruker ikke har valgt samme skriveform, med eller uten mellomrom, som forfatterne, får man altså ikke rett treff. Så kan man jo gjerne prøve seg med toner i tillegg for om mulig å øke sjansene for mer presis treff. Søker man etter v/o-ordet med delt pinyin og toner, *dǎ gōng*, får man rett treff, men søker man med toner og ikke delt pinyin, *dǎgōng*, får man ikke treff på v/o-ordet, men på et ord hvor v/o-ordet inngår i en sammensetning, *dǎgōngzú* 打工族, 'gästarbetare'. Dette siste ordet *dǎgōngzú* 打工族 får man ikke treff på hvis man søker *dagong* uten toner, men kun når man søker *dagongzu* uten toner. Her mangler forfatterne og utviklerne av appen et gjennomarbeidet system for søk etter v/o-ord. I denne pinyin-søkefunksjonen ligger det altså et klart forbedringspotensial til neste versjon av appen.

5. Oppsummering

NKSO, både som app og som papirordbok, er en svært god ressurs både for svenske og andre nordiske studenter som arbeider med å lære seg kinesisk, og for kinesere som ønsker å tilegne seg et vokabular på svensk. Materialet er omfattende og godt dekkende for det grunnleggende ordforrådet i moderne kinesisk, med fylldige forklaringer og mange nyttige eksempelsetninger knyttet til forklaringene. Skandinaviske kinesiskstudenter er vant til å benytte seg av den multifunksjonelle kinesisk-engelske app-ordboken Pleco. Med Norstedts app har vi fått et meget godt skandinavisk tillegg til vår gamle venn Pleco. Brukeren av ordboken kan også ha god nytte av de grammatiske kategorier som forklaringene og oversettelsene er inndelt i. Når det gjelder ordtilfanget, kunne denne anmelderen ønsket seg innføring av flere nyord i språket til nytte for den som orienterer seg også i de nye medier i Kina og blant yngre talere av kinesisk. Alle søkefunksjonene i papirordboken er gode, og det samme er søkefunksjonene for kinesiske tegn og de svenske forklaringer i appen. Søkefunksjonen for pinyin er svekket av en litt lite gjennomført – og kanskje lite gjennomtenkt – ordde-lingpolitikk for transkripsjon av uttale for oppslagsordene. Men når alt dette er sagt, er det ikke tvil om at det ligger et solid arbeid bak begge disse funksjonalitetene, app og papirordbok, av NKSO, og især appen representerer en helt ny tilførsel til de hjelpemidler vi har for å arbeide med og oversette kinesisk språk i Norden.

Litteratur

- Pleco: <<https://www.pleco.com/products.html>> (mars 2014).
 Ragvald, Lars (2000): *Kinesisk-svensk ordbok*. Lund: Institutionen för östasiatiska språk.

Rekkedal, Harald et al. (1997): *Norsk-kinesisk ordbok*. Oslo: Kunnskapsforlaget.

Rekkedal, Harald et al. (2005): *Stor norsk-kinesisk ordbok*. Oslo: Kunnskapsforlaget.

Rune Svarverud
professor i kinesisk
Institutt for kulturstudier og orientalske språk
Universitetet i Oslo
Postboks 1010 Blindern
NO-0315 Oslo
rune.svarverud@ikos.uio.no

Etymologi på ein ny måte

Lars S. Vikør

Yann de Caprona: *Norsk etymologisk ordbok. Tematisk ordnet*. Oslo: Kagge forlag 2013. 1920 sider. Pris: 599 NOK.

1. Innleiing

Norsk språk er ikkje rikt forsynt med etymologiske ordbøker. Klassikaren er Falk & Torp (1903–06) (med ei tyskspråkleg utgåve frå 1911). Meir omfattande er Torp (1919), der oppslagsorda er på nynorsk og i all hovudsak basert på ordtilfanget i Aasen (1873) og Ross (1895) sine ordbøker over dei norske dialektane. Så gjekk det åtti år før Bjorvand & Lindeman (2000) kom ut (revidert og utvida utgåve i 2007). Det er sjølvsagt den mest oppdaterte reint fagleg, men den omfattar berre eit utval arveord, som rett nok får ganske omfattande beskriving, men det er mange lakunar der frå eit brukarsynspunkt.

Derfor var det ikkje så lite av ein sensasjon da det i fjor (2013) vart lagt fram eit kjempearbeid på dette feltet, attpåtil laga av ein amatør, ein person som har utdannings- og yrkesbakgrunnen sin på heilt andre felt, og har produsert dette verket mest på hobbybasis. Yann de Caprona (heretter forkorta YdC) har ein internasjonal, i alle fall inter-europeisk, bakgrunn med svensk mor og fransk-amerikansk far, oppvekst i Italia og utdanning i samfunnsfag frå Paris. Han har budd i Oslo i snart tretti år og arbeid i Utanriksdepartementet, Raudekrossen og Riksrevisjonen. Han snakkar fleire vesteuropeiske språk (norsk, svensk, engelsk, fransk, italiensk) og har studert latin og nygresk. I arbeidet med dette verket har han hatt kontakt med mange sentrale fagpersonar innanfor ulike språk.

Det står på permen at ordboka inneheld 12 000 oppslagsord, og det er utvilsamt den mest voluminøse som er gitt ut. Ordtilfanget er det allment norske, basert på bokmålet, men dei fleste orda er sjølvstg felle for bokmål og nynorsk. Det mest originale ved boka er likevel ikkje noko av dette, men det at ho er tematisk ordna. Med andre ord har vi her noko så spesielt som ei onomasiologisk etymologisk ordbok, ein kombinert tesaurus og etymologisk ordbok. Eg har ikkje sett noko anna eksempel på dette; det normale er at etymologiske ordbøker er alfabetisk ordna (semasiologisk). Eg skal komme tilbake til ei vurdering av dette prinsippet. Det er eit relevant spørsmål om dette oppsettet er ein fordel ut over det å vere originalt og eigna til å skilje ut denne ordboka frå alle andre.

2. Disposisjon

Hovuddisposisjonen for boka er slik: Først kjem det ei innleiing med opplysning om oppbygginga av boka, kjeldene, den indo-europeiske språkfamilien, oversikt over regelrette lydkorrespondansar innanfor denne familien og over lydteikn som er brukt i boka. Så kjem hovuddelen, som altså er tematisk ordna. Det er 57 kapittel fordelt på sju hovudbolkar, med titlane *Mennesker*, *Naturen*, *Dyrelivet*, *Hverdagen*, *Samfunnet*, *Noen yrker og fag*, og til slutt *Språket*. Da er vi komne til side 1593, og så får vi over førti sider med bibliografiar (éin ordna etter språk og tema, éin ordna alfabetisk). Til slutt (s. 1639–1920) kjem ei alfabetisk liste over alle oppslagsorda, med tilvising til kapittel og sidetal.

Innanfor kvart kapittel er lemmalista ordna alfabetisk, og der kjem forklaringane til dei enkelte orda, som vi snart skal gi nokre eksempel på. Mange stader kastar YdC òg blikket utover, både i språket og i verda, og gir til dømes ei oversikt over kva eit ord eller eit begrep heiter på ulike språk. Alle kapitla blir innleidd av ein oversiktsartikkel om fellestrekk innanfor den ordgruppa kapitlet dekker. I tillegg er det ca. tretti tabellar over ord av spesielle in-

doeuropeiske røter, med utgangspunkt i eit norsk ord. Ein tabell har til dømes tittelen «Ord beslektet med *li*» (s. 221); den omfattar ord av den indoeuropeiske rota **klei-* ‘bøye, lene’, der samanhen- gen med bl.a. *klima* og *klinikk* blir vist gjennom ein trestruktur med latin, gresk og germansk som knutepunkt. Ein annan tabell heiter «Ord beslektet med *mye*» (s. 1576), av den indoeuropeiske rota **meg-* ‘stor, storartet’, og inkluderer bl.a. *magister* og *major* via latinske mellomledd.

Stoffmengda i denne boka gjer det til ein risikosport å prøve å gi ein heilskapleg omtale. Eg må berre «dukke ned» her og der og vise korleis boka er oppbygd gjennom eksemplifisering. Først går eg inn på makrostrukturen ved å sjå nærare på to kapittel: *Idrett* og *Helse*. Deretter plukkar eg ut nokre tilfeldige ord som skal repre- sentere litt ulike ordtypar, og ser på enkeltartiklane om dei.

Kapitlet *Idrett* er plassert i hovudbolken *Hverdagen*, som elles inneheld emne som mat og drikke, klede og hårstell, husinnreiing, arkitektur, kunst, litteratur, musikk og dans, leikar og spel, reiser, medium og data. Innleiinga til *Idrett* gir ei kort historisk oversikt over dei aktivitetsområda det omfattar, og gir deretter ein liten språkleg profil av feltet. Hovudspråket er engelsk, sidan England og USA har vore dei viktigaste inspirasjonskjeldene til mange av dei idrettsgreinene vi praktiserer eller har høyrte om (*bandy*, *baseball*, *basket*, *cricket*, *curling*, *tennis*, *volleyball* og fleire). Færre innslag frå andre nærskyldte språk (som fransk og italiensk) blir nemnt, sjølvstundt dei (få) norske *ski*-orda vi har eksportert og det austasia- tiske innslaget med bl.a. *judo*, *karate*, *kung fu*, *sumo* og *taekwondo*. Denne innleiinga fungerer som «appetittvekkar», og blir avslutta med ei kort liste over ord «med en særlig spennende etymologi og historie» (*aikido*, *ake*, *akrobat*, *birdie*, *bogey* osv.). Deretter kjem sjølve etymologilista på ca. 25 sider med eit par hundre oppslags- ord – ein del rett nok berre tilvisingar til andre artiklar på same ord under andre emnekapittel (til dømes *bane* under *Landreiser*, *bue* under *Arkitektur* og *urbanisme*, *dommer* under *dom* i *Lov* og

forbrytelser). Endeleg inneheld kapitlet ein figur over «ord beslektet med *ake*», der bl.a. *agere*, *amt* og *pedagog* inngår.

Kapitlet *Helse* står under *Noen yrker og fag* saman med landbruk, jakt og fiske, handverk og industri, økonomi og arbeid, myntnamn, matematikk, måleiningar, fysikk, kjemi, biologi og psykologi. Struktura er parallell med den eg alt har skissert: eit kort historisk utblikk over helse, sjukdom og medisin, og ein språkprofil av feltet som viser eit stort innslag av arveord og elles ei forholdsvis jamn fordeling på nokre nærstående «långevarar» (lågtysk, italiensk, fransk, høgtysk, engelsk og andre). Rolla til gresk og latin som «reservoar» for nylaga fagtermar innanfor dette feltet blir også kort nemnt. Orda «med en særlig spennende etymologi eller historie» er atskillig lengre her enn under *Idrett*, kanskje fordi dette er noko som gjeld alle og ikkje berre entusiastane? Det går frå *angina* og *briller* til *tyfus*, *vaksine* og *verk*. Sjølve etymologilista er på nesten femti sider, dobbelt så lang som under *Idrett*.

3. Nokre enkeltartiklar

Så nokre tilfeldige dykk i sjølve ordartiklane. Eg har plukka ut fire ord og søker dei opp i ordboka, først to vanlege importord, så to sjeldne heimenorske ord. Eg trur det er dei ordkategoriane folk flest helst vil slå opp på. Orda er *admiral*, *bank*, *jabbe* og *talle*. Eg gjengir artiklane med kritiske kommentarar der eg meiner det er naturleg; det gjeld særleg den første av artiklane. Det kan stå som eit eksempel på korleis desse etymologiane kan problematiserast og borast i, utan at det nødvendigvis finst opplagte svar.

Admiral står i kapitlet *Det militære*, i hovudbolken *Samfunnet* (s. 1061–2). Artikkelen er slik:

admiral: sjøoffiser av høyeste rang. Av fransk *amiral*, avledet av arabisk *amīr* «fyrste, befalingsmann» (jamfør *emir* i

kapitlet *Politikk og forvaltning* [...] + kanskje *al 'ali* «meget stor» eller *ar-rahl* «av transport» (men ikke *bahr* «hav», som omtalt i flere ordbøker). Da skal ordet egentlig bety «øverste fyrste, befalingsmann» eller «fyrste, befalingsmann av transporten». Tittelen er kommet til europeiske språk via normannernes kongedømme på Sicilia og betydde først «hærleder» og mye seinere «eskadreleder» [her følger seks referansar til ordbokskjelder].

Det er vanskeleg å skjønne avvisinga av *amir al-bahr* 'havets kommandant' som mogleg opphav. Det som (etter vanleg oppfatning) har overlevd i våre språk, er ein del av eit lengre uttrykk: *amir al-xxx*, altså 'xxx's kommandant', og reint språkleg kan i prinsippet nesten kva som helst ha stått etter *al*. I arabisk har ordet *amir* 'kommandant' vore brukt i mange ulike samanhengar, alt etter kva vedkommande kommanderte over. Dermed framstår *amir al-bahr* stadig som eit mogleg opphav til *admiral*. Dei andre alternativa YdC nemner, er faktisk mindre plausible. Men det finst andre alternative forklaringar, som at ordet kanskje berre kjem av *amir* med den latinske endinga *-alis* (etter analogi med *general*). Dei eldste kjeldene der ordet er nemnt (frå tidleg mellomalder) gir ikkje grunnlag for ein konsistent hypotese.¹

Bank står i kapitlet Økonomi og arbeid i hovudbolken *Noen yrker og fag* (s. 1181):

bank: institusjon som tar imot innskudd, gir lån, driver handel med verdipapirer og utfører pengetransaksjonstjenester. Gjennom fransk *banque* av italiensk *banco* (nå oftast *banca*) som egentlig betyr «pengevekslerdisk» siden de første banker var benker der penger ble vekslet og seinere lånt. *Banco*, som er avledet av frankisk **bank* og beslektet med norsk *benk* (se dette ordet i kapitlet *Husets innredning*

1 Takk til Knut S. Vikør for oppklaringar kring dette.

og utstyr [...]), har på norsk også gitt **banko** «bankopost, verdipost» og *bankerott* (se dette ordet) [tre referansar].

Bank er altså egentleg eit germansk ord (frankisk var eit germansk språk som låg nært dagens nederlandsk og tysk) som er «lånt ut» til verda gjennom italiensk.

Jabbe ‘snakke mykje, fort, snakke tull, skravle’ står ikkje i *Norsk etymologisk ordbok*. Det er eit folkeleg norsk ord som ein må til *Norsk Ordbok* for å få beskrive; Torp (1919) har det også med, og viser til parallelle ord i færøysk og shetlandsk. Det har YdC altså ikkje fått med. Det gjeld også eit anna ord *jabbe* ‘gå tungt, gå tregt, labbe, trippe’. Begge orda er med i *Norsk Riksmålsordbok* som dialektale. Truleg er det lydord i utgangspunktet. Ivar Aasen og Hans Ross har forteneasta av å ha registrert desse orda i ordbok, og dei har høyrte det i fleire landsdelar.

Talle er derimot eit ord YdC har fått med, i kapitlet *Landbruk, jakt og fiske*, hovudbolken *Noen yrker og fag* (s. 1134). Han skriv om det:

talle: (sammentrampet) gjødsel (særlig av geit og sau). Av norrønt *tað* «gjødsel» beslektet med *taða* «gjødslet eng», derav nynorsk *tøde* «talle; gjødslet eng; frodig eng». Det er foreslått at *talle* går tilbake til indoeuropeisk **dā-* «dele» eller den utvidede formen **dat-* «det utbredte» og kan være beslektet med gresk *daíomai* «jeg deler» og sanskrit *dáyate* «dele opp» (Watkins, de Vries, Torp, Hellquist).

I det siste utdraget har eg ført opp referansane, som døme på den måten YdC gjer det på gjennom heile boka. Som ein ser, viser han berre til ei uspesifisert mengd forfattarnamn. Han har ei imponerende, ja, enorm, mengd kjelder, som litteraturlistene hans viser, og fagfolk med god tid kan nok kontrollere opplysningane hans, men for andre er det vel i praksis umogleg (eg kjem tilbake til det

nedanfor). Det er likevel vanskeleg å sjå korleis han kunne gjere det annleis om han i det heile skulle få realisert prosjektet. Ein unødvendig komplikasjon har han likevel tillate seg: Ordbøker med to forfattarar er oppgitt med initialar, til dømes BL for Bjorvand & Lindeman (2007). Men i bibliografien står berre dei fulle forfattarnamna som oppslag; her burde kombinasjonar som BL (og andre, mindre kjente) ha stått på alfabetisk plass med tilvising til tittel.

4. Mangslungne artiklar

Mange artiklar går ut over det å gi etymologiane til oppslagsordet; vi kan altså få ganske lange og omfattande artiklar med ulike typar informasjon om ordet og begrepet. Nokre få døme: Under *Kroppens funksjoner og aktiviteter* (hovudbolk *Mennesker*) finn vi verbet *ligge*. Artikkelen er for så vidt ikkje så lang, men det blir plass til ei liste over avleiingar av opphavsrota **legh-* i ulike indoeuropeiske språkgreiner: latin *lectus* ‘seng’, gresk *lékhos* ‘seng’, gælisk *laigh* ‘ligge’, russisk *ležat* ‘ligge’ og eit par til. I same bolk finn vi *mareritt*, som byr på andre utfordringar. Her kjem folkloristikken inn gjennom ei liste over ulike vesen som ligg til grunn for dette begrepet: *mare* i svensk *mardröm*, engelsk *nightmare* og fransk *cauchemar*, *alv* i tysk *Albtraum*, «noko som trykker på brystet» i spanske og italienske ord for dette, og andre begrep i gresk, walisisk og katalansk.

Ordet *palme* (under *Planter og frukter*, hovudbolk *Naturen*) blir innleidd med etymologien som går til latin *palma* ‘handflate’, basert på likskap mellom eit palmeblad og ei hand med utstrekte fingrar. Så får vi ulike kulturelt baserte avleiingar, bl.a. *palmesøndag*, nokre indoeuropeiske avleiingar av rota **pelə-* ‘flat’, særleg med tydinga ‘handflate’ (gammalhøgtysk *folma*, gresk *paláme*, walisisk *llaw*), og så kjem nokre avleidde stadnamn (*Palma* på

Mallorca, *Las Palmas* på Kanariøyane og *Palm Springs* i California). Alt dette på under ei spalte.

Eit døme på korleis ein ordartikkel kan nærme seg sjangeren leksikonartikkel, er **bjørn** (under *Pattedyr*, hovudbolk *Dyrelivet*). Den fyller ei heil side i boka. Etter sjølve hovudetymologien (indoeuropeisk **bher-* ‘brun’) får vi ei drøfting av tabu-problematikken og det russiske *medved* ‘honningetar’ (av YdC feilaktig transkribert *medvėd*) som uttrykk for den, og så går vi også her over i folkloristikken, både den gresk-latinske, den keltiske og den germanske. Eg kan her berre antyde dei «stasjonane» YdC passerer gjennom i utgreiinga si av begrepet: Vi blir presentert for den gammalgreske bakgrunnen for stjernebilda *Storebjørn* og *Veslebjørn*, nemningane *Arktis* og *Antarktis* (av gresk *árktos* ‘bjørn’), det norrøne *berserk*, ymse ordlag og forestillingar om bjørnen (*ti manns styrke, tolv manns vett; bjørneteneste; selje skinnen før bjørnen er skoten*), personnamna *Bjørn* og *Ursula*, og stadnamn knytte til bjørnen – eller ikkje: *Bern* og *Berlin*, vanleg assosiert med *Bär* ‘bjørn’, antydar YdC kan komme av indoeuropeisk *ber-* ‘sumpmark’ – basert på kjelder som han gir opp, men han markerer samvitsfullt uvisse kring dette.

Dette var nokre døme på kor mangslungne desse artiklane kan vere, og på korleis YdC kan sno seg mellom det språklege og det ensyklopediske på ofte uføreseielege måtar. Det er mogleg at « amatørposisjonen » gjer dette lettare for han enn det ville ha vore for ein «rein» språkforskar, som ofte vil vere meir innstilt på å halde seg innanfor etablerte faglege skiljeliner.

5. Hovudbolken Språket

I eit skrift som *LexicoNordica* kan det høve å sjå nærare på den siste hovudbolken, *Språket*. Den inneheld kapitla *Lingvistikk og grammatikk, Språknavn, Talespråket, Skriftsystemer, Prefikser og suffikser, Banning, og Noen øvrige ord*, og omfattar ca. 150 sider. I kapitlet

Lingvistikk og grammatikk kjem latin og gresk sterkt inn, både som kjelder til terminologien, og som leverandørar av døme (mest latin, da). Lista omfattar alle nivå i språket. Det er ikkje mange stader YdC går skikkeleg i baret, men det gjer han etter mitt syn når han gir ordet *kasus* fleirtalsforma *kasi*. Det må vere ein hyperkorreksjon, og altså heilt gale. Det heiter *kasus* også i fleirtal, ikkje berre på norsk, men også på latin: *casus* er ein u-stamme og får altså ingen *-i* i fleirtal. *Hyperkorrekt* er for øvrig eit ord YdC ikkje har funne plass til (sjølv om han har med prefikset *hyper-*, men med andre orddøme).

Svært interessant er kapitlet *Språknavn*, fordi språknamna jo utgjer mykje av beingrinda i eit verk som dette. Her har han til dels lange omtalar av dei språka han behandlar, der han både inkluderer beskrivingar av den allmennhistoriske bakgrunnen for språka og rekker av døme på ord vi har fått frå dei respektive språka. Dei fungerer som leksikonartiklar om språka, men her er relasjonen mellom det språklege og det ensyklopediske meir uanstrengt enn når ein behandlar andre emne, kan ein vel seie. Dei lengste artiklane handlar om engelsk, latin og tysk, og det er jo ikkje unaturleg. Artikkelen om tysk (s. 1495–6) inneheld både drøftingar av etymologien som er basert på *þeudi- ‘folk’ og ei oversikt over dei mange andre (til dels stammebaserte) nemningane for Tyskland i ymse grannespråk: engelsk *Germany*, fransk *Allemagne*, polsk *Niemcy*, finsk *Saksa*, litauisk *Vokietija*.

Eit problematisk punkt generelt i boka er at YdC konsekvent bruker *tysk* for det høgtyske skriftspråket, også når han set det i kontrast med *lavtysk*, som har vore så viktig for det norske ordforrådet, til dømes i artikkelen **brøle** (i kapitlet *Talespråket*, s. 1500): «Trolig av lavtysk **brölen*, som likesom nederlandsk *brullen* og tysk *brüllen*, går tilbake til germansk **brōlian-*». Det hadde tatt seg betre ut å gjennomføre *høytysk*, i alle fall når det står i direkte motsetning til *lavtysk* – som, uansett korleis ein snur og vender på det, også er tysk.

Kapitlet *Talespråket* inneheld kanskje meir allmennspråklege ord om språkbruk enn faglege (*akklamasjon, bable, deklamere, fritte, ljuge* (alfabetisert på den forma), *retorikk, si, stemme, svar, ytre* osv.). Under *Skriftsystemer* har YdC ein særskild bolk om enkeltbokstavane i det latinske alfabetet, med ein figur over utviklinga av bokstavane frå fønikisk til latin. Prefiks og suffiks er samla i eitt kapittel, og så kjem ei svært kort liste over bannord og skjellsord (berre to sider, grunnlagt med at dei fleste slike ord har «en usikker opprinnelse»). Til slutt *Noen øvrige ord*, i stor grad grammatiske funksjonsord (*av, bli, den, gjøre, hvem, jeg, og, opp, selv, til* osv.), men òg ein del innhaldsord, både vanlege og meir spesielle framandord, som han av ein eller annan grunn ikkje har fått plassert i systemet sitt (*bagatell, blande, diffundere, enorm, filateli, forske, full, hobby, irre, kavalkade* osv.). Med den omfattande emneinndelinga som heile boka bygger på, ser det merkeleg ut at ikkje mange av desse orda kunne ha vore plassert i ein «naturleg» omgivnad elles i boka, men eg tar ikkje opp det i detalj her.

6. Tilhøvet til kjeldene: Ord frå malayisk

YdC bygger altså på ei imponerande mengd kjelder, men utan å spesifisere kor han har dei enkelte opplysningane frå ut over å referere generelt under kvart oppslagsord, som eg har vist eksempel på ovanfor. For å finne ut litt meir om dette, har eg sett på ei ordgruppe frå eit språk YdC ikkje er kompetent i, nemleg malayisk (som også er nasjonalspråk i Indonesia under namnet indonesisk; det er snakk om to nært beslektet varietetar som er gjensidig forståelege, men likevel har ein del ulikskapar i ordbruk og idiomatikk). Eg har sjølv studert dette språket og kan dermed «kike YdC i korta», som vi seier på norsk. I Vikør (2012) har eg beskrive etymologien til ein del malayisk/indonesiske importord i norsk, og sidan metoden min var annleis enn den YdC har brukt, går det an å samanlikne.

På s. 1485 (under *Språknavn*) har YdC ei liste over malayiske ord i norsk (som eksempel på ord han har med i ordboka). Nokre av dei er (meir enn) tvilsame (*beriberi, djunke, mangrove*), men det spelar mindre rolle her. Eg har gått gjennom og registrert alle kjeldene han gir opp, og konstaterer at det i all hovudsak er snakk om allmenne etymologiske ordbøker for vesteuropeiske språk, som altså ikkje bygger på spesialkompetanse i malayisk og andre indonesiske språk. I Vikør (2012) har eg tatt utgangspunkt i *Bokmålsordboka* og *Nynorskordboka* med tanke på førekomsten i norsk, og deretter søkt opp orda i malayiske og indonesiske ordbøker. Eg vil ikkje gå i detalj i jamføringa, bl.a. fordi eg ikkje har hatt høve til å sjekke YdC sine kjelder direkte. Eg viser berre til artikkelen min for dei som vil jamføre måtane våre å gå fram på.

Eg kommenterer to av orda, der YdC har hypotesar eg ikkje har sett andre stader. *Amok*, seier han på s. 1420, kan vere avleidd av «*amar-kkan* «kriger», der *amar* betyr «krig» på malayalam (et dravidisk språk fra Kerala i Sørvest-India)». Dette verkar langhenta på meg. Om sjølve ordet *malayisk* skriv han at det stammar frå sanskrit og «kanskje er avledet av dravidisk *malai* «fjell»». Det siste skal da også gjelde *malayalam*. Eg er også her skeptisk og har ein annan etymologi i Vikør 2012, men det er interessant at YdC i begge tilfelle viser til dravidisk. Det er eit faktum at Malaya og Vest-Indonesia i mange hundreår før kolonitida hadde mykje kontakt med det dravidiske Sør-India, og framleis finst det ein tamilsk minoritet i Malaysia (tamil er det største dravidiske språket). Det er også eit faktum at denne kontakten i stor grad gjekk tapt i nyare tid, og dei vitskaplege miljøa som i moderne tid har danna seg kring dravidisk og indonesisk lingvistik har vore heilt åtskilde, med engelsk som einaste gjensidig forståelege kontaktspråk. Eg trur ikkje det har funnest ein einaste person som har hatt ekspertkunnskap i begge språkgruppene, i alle fall var det slik da eg studerte indonesisk på syttitalet, og det er heilt nødvendig for at ein skal kunne gjenoppdage gamle sambandsliner. Eg kan altså ikkje bastant av-

vise YdC sine teoriar, sjølv om eg er skeptisk; den same mangelen på kontakt kan jo lett føre til ukritisk overtaking av teoriar ein ikkje har innsikt nok til å vurdere. Så der står saka.

Ein siste merknad er at YdC gjengir opphavsforma til *orangutang* (s. 427) feil, det er *orang (h)utan* 'skogmenneske'. *Hutan* tyder 'skog', mens *hutang* tyder 'gjeld, skuld' på indonesisk, så denne feilen kunne gi grunnlag for ein humoristisk kommentar. Lyden *-ng* i vår form av ordet har nok oppstått i «importprosessen» i eitt av mottakarspråka.

7. Sluttvurdering

På dette stadiet i meldinga er det vel på tide med ei generell vurdering av sjølve det overordna grepet som gjer denne boka til eit slikt originalt verk. Som prestasjon er det imponerende og beundringsverdig å hanskast med ei slik overveldande stoffmengd på denne kompliserte, men godt gjennomførte måten. Frå eit brukarsynspunkt blir nok vurderinga meir ambivalent. Det er ikkje lett å ta seg fram i boka, ein må ha god tid og vere innstilt på litt plundring – trass i at ein har den alfabetiske lista på nesten 300 sider bak i boka. Dei fleste som slår opp i ei etymologisk ordbok, er interessert i enkeltordet der og da, ikkje i samanhengane – som stort sett ikkje er etymologiske – med det semantiske feltet ordet høyrer til i. Likevel får boka fram visse dimensjonar som ein ikkje lett ser med den tradisjonelle (alfabetiske) disposisjonen. Det er såleis fascinerande – og kulturhistorisk relevant – å sjå korleis ulike felt har ulike språkprofilar, altså at somme språk dominerer som långevarar på somme felt, andre på andre, arveorda på atter andre. Mykje av dette kan ein jo tenke seg til ut frå vanleg allmennkunnskap, men det er likevel fascinerande å få det demonstrert for eit stort tal enkeltord slik det blir gjort her.

Det er nok også slik at somme semantiske felt er meir saman-

hengande heilskapar for brukarar flest enn andre, og der kan ein ha ei særskild glede av å sjå enkeltorda i samanheng med kvarandre og jamføre dei utan å måtte bla alle stader i boka. Det gjeld slike felt som fargar, tal, dyregrupper, myntar, bokstavane i alfabetet, og andre. Likevel trur eg ikkje YdC får særleg mange etterføljarar i denne spesielle sjangeren «onomasiologisk etymologisk ordbok». For dei fleste føremål vil ein tradisjonell «glattalfabetisk» struktur vere mest praktisk og handterbar. Der kan ein søke ordet ein vil ha greie på direkte opp i staden for å ta det i to omgangar – først i den alfabetiske lista, så med leiting etter eitt av over 1500 sidetal. Ein ekstra komplikasjon er det at mange ord høyrer heime i ulike semantiske felt, så YdC må lage mange tilvisingar mellom ulike oppslag av same ordet. Til dømes: *banne* står under *Talespråket* med tilvising til *Banning*; *alpakka* og *astrakan* står under *Klær, stoff og sko* med tilvising til *Pattedyr* osv. Det ligg i sakas natur at det må bli ganske mykje av slikt. Her ville ei glattalfabetisk ordning føre til mindre plassbruk og ein meir greiskoren struktur. I den alfabetiske lista til slutt er det rett nok berre vist til hovudomtalen av det aktuelle ordet.

YdC er seg bevisst at han er amatør, og han lener seg tungt på dei mange kjeldene han har konsultert. Han skriv i innleiinga (s. 21) at han berre har tillate seg ein einaste personleg teori om eit ord; det gjeld *tårn* i sjakk – i kapitlet *Spill og leker* (s. 726). Her handlar det om sjølv brikka, meir enn ordet *tårn* (som blir forklart ein annan stad, i kapitlet *Arkitektur og urbanisme*, s. 589). Brikka heitte på persisk *rokh*, foreslått avleidd av *rah* ‘krigsvogn’, og YdC spør seg «hvorfør denne høyst bevegelige brikken nå er representert ved et stillestående tårn». Så skriv han etter å ha referert eit par andre teoriar: «Personlig er jeg fristet til å tro – men har ikke noe belegg for det – at persisk *roh* og middelalderlatin *rochus* er blitt forbundet med italiensk *rocca* ‘fjellfestning’».

Uttrykt så varsamt må det vere lov å antyde ein teori også for

ein amatør. Men etymologiske ordbøker verkar ofte ganske skråsikre der ein som interessert lesar kunne tenke seg meir uvisse og atterhald. YdC er fullstendig avhengig av det ekspertbatteriet han bygger heile verket på. Han verkar svært samvitsfull i måten sin å behandle stoffet på, kanskje nettopp fordi han veit at han lett vil bli avfeidd om han kjem med bombastiske fikse idear. Det finn eg altså ikkje her, men ingen kan lage eit slikt verk sånn at det blir absolutt sikkert og rett. Når eg ovanfor har funne ting å pirke på, må ein sjå det i det lyset. Også på det reine korrekturplanet er eit sånt verk ei gedigen utfordring, ikkje minst bibliografiane, og igjen må eg seie at YdC har komme godt frå det i det store og heile, men enkelte feilskrivne forfattarnamn som «Linström» for *Lindström* (s. 1627), «Sandness» for *Sandnes* (s. 1632) og «Torby» for *Torvik* (s. 1627) har han ikkje unngått. Det finst fleire døme, men eg kan ikkje gå meir inn på det her.

Ei reell kvalitetsvurdering av eit slikt verk kan i det heile tatt ikkje gjerast i éi melding, det må skje i det små gjennom (regel-messig) bruk over lang tid. Det vil bli oppdaga feil, enten på grunn av ny forskning, eller på grunn av kunnskap som var tilgjengeleg alt tidlegare. Eit eksempel på det har *Lexico-Nordicas* norske redaktør, Sturla Berg-Olsen, gjort meg merksam på. Han er kompetent på baltiske språk, noko eg ikkje er, og har merka seg ei rekke feil i bruken av diakritiske teikn i litauiske ord. Her burde ein fagperson på dette språket vore konsultert.

Likevel: Denne meldaren er sjølv ordboksredaktør, og slike har nok eit meir illusjonslaust forhold til dette enn «meinige» brukarar ofte har. Nivået verkar i det store og heile så høgt her, både når det gjeld ambisjonar og prestasjonar, at verket vil få status som det sentrale arbeidet for vår tid i sitt slag. Det tar ikkje vekk behovet for andre etymologiske ordbøker med andre vinklingar, som Bjorvand og Lindemans, der ei ny utgåve igjen er i arbeid. Den loddar djupare, men femner mindre breitt enn YdC si bok, og vi treng begge måtane innanfor eit så samansett emnefelt.

Litteratur

- Bjorvand, Harald & Fredrik Otto Lindeman (2000): *Våre arveord. Etymologisk ordbok*. Oslo: Novus forlag. Revidert og utvidet utgave 2007.
- Bokmålsordboka* (2005). Boye Wangensteen (red.). 3. utgave. Oslo: Kunnskapsforlaget.
- Falk, Hjalmar & Alf Torp (1903–06): *Etymologisk Ordbog over det norske og det danske Sprog I–II*. Kristiania: Aschehoug.
- Falk, Hjalmar & Alf Torp (1911): *Norwegisch-Dänisches etymologisches Wörterbuch I–II*. Heidelberg: Carl Winthers Universitätsbuchhandlung.
- Norsk Ordbok V: harm–jåttut* (2005). Lars S. Vikør (hovudred.). Oslo: Det Norske Samlaget.
- Norsk Riksmålsordbok I,2: for- –låvevegg* (1937). Trygve Knudsen & Alf Sommerfelt (red.). Oslo: Aschehoug.
- Nynorskordboka* (2006). Marit Hovdenak et al. (red.). 4. utgåva. Oslo: Det Norske Samlaget.
- Ross, Hans (1895): *Norsk Ordbog*. Kristiania: Alb. Cammermeyers Forlag.
- Torp, Alf (1919): *Nynorsk etymologisk ordbok*. Kristiania: Aschehoug.
- Vikør, Lars S. (2012): Indonesiske ord i norsk. I: Unn Røyneland & Hans-Olav Enger (red.): *Fra holtijar til holting. Språkhistoriske og språksosiologiske artikler til Arne Torp på 70-årsdagen*. Oslo: Novus forlag, 399–412.
- Aasen, Ivar (1873): *Norsk Ordbog med dansk Forklaring*. Christiania: Alb. Cammermeyers Forlag.

Lars S. Vikør
 professor
 Norsk Ordbok 2014
 Institutt for lingvistiske og nordiske studium
 Universitetet i Oslo
 Postboks 1021 Blindern
 NO-0315 Oslo
 l.s.vikor@iln.uio.no

MEDDELANDEN

Nyt fra bestyrelsen for Nordisk Forening for Leksikografi

Henrik Hovmark

1. 13. Konference om Leksikografi i Norden – København 19.-22. maj 2015

13. Konference om Leksikografi i Norden afholdes denne gang i København og arrangeres af Ømålsordbogen, Nordisk Forskningsinstitut, Københavns Universitet i samarbejde med Nordisk Forening for Leksikografi (NFL).

Konferencen finder sted på Københavns Universitet på Amager. Det faglige program afvikles 20.-22. maj 2015, men allerede 19. maj om aftenen er der en uformel sammenkomst med mulighed for registrering.

Alle oplysninger om konferencen – praktiske såvel som indholdsmæssige – annonceres på konferencens hjemmeside (<nfi.ku.dk/nfl2015>), som opdateres løbende. Indsendelse af abstracts, tilmelding og betaling vil ligeledes foregå via hjemmesiden. Fristen for indsendelse af abstracts er 15. januar 2015 (jf. afsnit 1.2).

13. Konference om Leksikografi i Norden er i lighed med tidligere konferencer støttet af Nordplus Nordiske Sprog.

1.1. Fagligt program

Lige siden starten i 1991 har de nordiske konferencer i leksikografi, arrangeret af NFL, samlet deltagere fra alle dele af det leksikografiske miljø. 13. Konference om Leksikografi i Norden holder denne tradition i hævd og indkalder derfor bidrag fra alle områder af leksikografien og tilgrænsende discipliner. Som sædvanlig vil der være mulighed for at deltage med såvel foredrag som poster og

softwarepræsentationer. Som noget nyt introduceres denne gang også et tema: Den digitale ordbogsbruger, se afsnit 1.1.2. Der udgives en fagfællebedømt konferencerapport efter konferencen.

1.1.1. Plenarforedragsholdere

- Maria Koptjevskaja Tamm, professor, Stockholms universitet
- Oddrun Grønvik, hovedredaktør, Norsk Ordbok 2014, Oslo
- Esben Alfort, lingvist og forskningsleder, Ankiro, København
- Holger Hvelplund, direktør, IDM Danmark

1.1.2. Tema: Den digitale ordbogsbruger

I dag udgives ordbøger digitalt – og ældre ordbøger digitaliseres – men:

- Hvordan udgiver man egentlig digitalt?
- Hvordan skal traditionelle ordbogsdata præsenteres i en digital applikation?
- Hvordan bruger brugerne de digitale ordbøger?
- Hvilke brugere er der tale om, og hvordan får man viden om deres adfærd?
- Hvordan tager man højde for, at en digital ordbog ofte vil have mange forskellige brugere?

Vigtigheden af at interessere sig for, hvordan sproglig viden op-søges, tilegnes og bruges på digitale medier, understreges fx af den stærkt øgede brug af nye platforme og kommunikationsprogrammer som fx smarttelefoner, tablets/pads, Facebook, Twitter, Whatsapp osv.

Vi opfordrer alle dele af det leksikografiske miljø og tilgrænsende discipliner til at komme med indlæg om dette emne. Der

lægges vægt på erfaringsudveksling og på diskussion af problemer, udfordringer og løsninger på alle niveauer.

To af plenarforedragsholderne, Esben Alfort og Holger Hvelplund, bidrager til konferencens tema:

Esben Alfort er leder af forsknings- og analyseafdelingen i den sprogteknologiske virksomhed Ankiro, som bl.a. udvikler søgeteknologi. Han har i sin ph.d.-afhandling forsket i søgeadfærd i et forsøg på at forstå, hvad der bestemmer brugeres valg af søgestreng, og hvad der gør et dokument relevant.

Holger Hvelplund har i mange år arbejdet med digitale ord-bøger og ordbogsredigeringsværktøjer, i de seneste år i det franske softwarefirma IDM (Interactive Digital Media), hvor han også har arbejdet med søgemaskiner og søgemaskineoptimering (SEO).

1.2. Indsendelse af abstracts

Der åbnes for indsendelse af abstracts til foredrag og poster 1. oktober 2014, og sidste frist er 15. januar 2015. Forslagene bedømmes anonymt af et fagligt panel, og alle får besked om accept af bidrag 1. marts 2015. Foredragssproget er dansk, norsk eller svensk.

1.3. Vigtige datoer

- | | |
|------------------|--|
| 1. oktober 2014 | Indsendelse af abstracts til foredrag og poster åbner. |
| 15. januar 2015 | Sidste frist for indsendelse af abstracts. |
| 15. januar 2015 | Tilmelding til konferencen åbner. |
| 1. marts 2015 | Besked om accept af foredrag, poster med videre. |
| 15. marts 2015 | Sidste frist for tilmelding med nedsat konferenceafgift. |
| 1. april 2015 | Sidste frist for tilmelding og betaling. |
| 19.-22. maj 2015 | Konference. |

1. september 2015 Sidste frist for indlevering af artikel til konferencerapporten.

Vel mødt i København i 2015!

Asgerd Gudiksen, Henrik Hovmark, Lena Wienecke Andersen og Liisa Deth Theilgaard

Ømålsordbogen, Nordisk Forskningsinstitut, Københavns Universitet

2. NFL og Nordisk Sprogkoordination

NFL har i en del år haft et frugtbart samarbejde med Nordisk Sprogkoordination (NOSK), som bl.a. har vist stor velvilje ved at huse NFL's hjemmeside. 1. januar 2014 overgik rollen som forvalter af NOSK fra Dansk Sprognævn til Foreningerne Nordens Forbund (FNF), der er ansvarlig for det nordiske sprogsamarbejde i perioden 2014-18. NFL er glad for at kunne videreføre samarbejdet med NOSK i den kommende periode og dermed fortsat bidrage til det nordiske sprog- og kultursamarbejde. NFL vil således fortsat optræde på NOSK's hjemmeside, der i forbindelse med projekt-overdragelsen har gennemgået en omfattende modernisering.

Digitaliseringsprojektet, der også er blevet administreret af NOSK og bl.a. omfatter digitalisering og webpublicering af NFL's to skriftserier: LexicoNordica og konferencerapporterne, fuldføres derimod af den tidligere projektleder Dansk Sprognævn, desværre ikke i 2013 som nævnt i LexicoNordica 20, men først i 2014.

Henrik Hovmark
lektor, ph.d.
Nordisk Forskningsinstitut
Afdeling for Dialektforskning
Københavns Universitet
Njalsgade 136
DK-2300 København S
hovmark@hum.ku.dk

REDAKTIONELLT

1. LexicoNordica **udkommer** hvert år i november. Tidsskriftet indeholder leksikografiske bidrag som er skrevet på et af følgende nordiske sprog: dansk, finsk, færøsk, islandsk, norsk (bokmål eller nynorsk) og svensk. Bidrag på engelsk, fransk eller tysk kan også optages hvis særlige forhold taler for det.
2. **Bidrag** sendes til det medlem af redaktionskomitéen som repræsenterer bidragerens land:
 - Christian Becker-Christensen, Børglumvej 39, st., DK-2720 Vanløse. <cbcvanlose@gmail.com>.
 - Sturla Berg-Olsen, Språkrådet, Postboks 8107 Dep, NO-0032 Oslo. <sturla.berg-olsen@sprakradet.no>.
 - Annika Karlholm, Institutet för språk och folkminnen, Dialekt- och folkminnesarkivet i Uppsala, Box 135, SE-751 04 Uppsala. <annika.karlholm@sprakochfolkminnen.se>.
 - Mariann Skog-Södersved, Vasa universitet, Tyska språket och litteraturen, PB 700, FI-65101 Vaasa. <mss@uwasa.fi>.
 - Ásta Svavarsdóttir, Árni Magnússon-instituttet for islandske studier, Neshaga 16, IS-107 Reykjavík. <asta@hi.is>.

Seneste tidspunkt for aflevering af bidrag er **den 1. april** hvis artiklen skal kunne trykkes i det nummer af tidsskriftet som udkommer i november samme år. Bidraget indleveres elektronisk i både tekstbehandlingsformat og i PDF-format.

3. **Illustrationer** der skal medtages i artiklen, indsættes i manuskriptet og vedlægges som separate grafikfiler, helst i JPG-format og minimum 300 dpi.

4. **Manuskript:** Bidraget forfattes i LexicoNordicas stilark, der kan rekvireres ved henvendelse til redaktionen. Manuskriptet **indledes** med titel på artiklen og forfatterens navn. For tematiske og ikke-tematiske bidrag følger et **abstract** på engelsk på op til 10 linjer og dernæst selve artiklen, som opdeles i kapitler. Bidraget afsluttes med angivelse af forfatterens navn, titel samt post- og e-mailadresse. Bidrag kan normalt have et omfang på højst 20 sider.
5. **Citater:** Kortere citater (op til 3 linjer) bringes som en del af teksten med dobbelte anførselstegn omkring, mens længere citater eller fremhævelser af større vigtighed gives i et afsnit for sig selv **uden** anførselstegn.
6. Vi anbefaler en tilbageholdende brug af **fodnoter**. Evt. nødvendige noter gennemnummereres i teksten med højtstillet angivelse uden parentes.
7. **Litteraturhenvisninger** foretages i teksten efter følgende model:

som det fremgår af Herbst (2009)
 som det fremgår af Borin & Forsberg (2011:18)
 (se Herbst 2009:158ff.)

I den løbende tekst angives ikke hele internetadresser, men et forfatternavn eller en angivelse af titlen på internetbidraget, som bruges i litteraturlisten. Her angives internetadresser uden understregning, men omgivet af < >.

8. Særlige angivelser: Vær tilbageholdende med brug af **fede** typer; **sprogksempler** markeres med kursiv, fx: ordet *ungkarl* har synonymet *alenemand*; **betydninger** af sproglige enheder

angives ved hjælp af enkelte anførselstegn, fx: ‘en ugift mand’; dobbelte anførselstegn bruges ved citater eller forbehold, fx: De er vokset op i de “glade” tressere. Tegnsætningsreglerne, bl.a. for brug af komma, tankestreg, bindestreg (i betydningen ‘fra ... til’), er forskellige i de nordiske lande, og forfatterne bør naturligvis følge reglerne for det sprog som bruges i artiklen.

9. Litteraturangivelser

I tilfælde af en længere litteraturliste kan den inddeles i to dele i lighed med nedenstående eksempel. Hvad angår angivelser som *red.*, *eds.*, *Hrsg.*, anbefales det så vidt muligt at bruge originalsproget. Det vigtigste er dog konsekvens inden for samme liste.

Litteratur

Ordbøger

ALD (1948) = A.S. Hornby, E.V. Gatenby & H. Wakefield: *A Learner's Dictionary of Current English*. London: Oxford University Press.

COBUILD (1987) = *Collins COBUILD English Language Dictionary*. Editor in Chief: John Sinclair, Managing Editor: Patrick Hanks. London/Glasgow: Collins.

ELEXIKO = Annette Klosa m.fl.: *elexiko*. Mannheim: Institut für Deutsche Sprache. <www.elexiko.de> (maj 2008).

Jarvad, Pia (1999): *Nye Ord. Ordbog over nye ord i dansk 1955-1998*. København: Gyldendal.

Norstedts stora engelska ordbok (2000). Stockholm: Norstedts.

Oxford-Hachette French Dictionary (1994). Oxford: Oxford University Press.

STO = Anna Braasch m.fl. (red.): *Sprogteknologisk Ordbase*. København: Center for Sprogteknologi. <www.cst.dk/cgibin/defisto> (april 2007).

Anden litteratur

Haiman, John (1980): Dictionaries and Encyclopedias. I: *Lingua* 50, 329-357.

Lakoff, George & Mark Johnson (1980): *Metaphors we live by*. Chicago/London: The University of Chicago Press.

Mugdan, Joachim (1985a): Grammatik im Wörterbuch: Wortbildung. I: Herbert Ernst Wiegand (Hrsg.): *Studien zur neuhochdeutschen Lexikographie IV*. Hildesheim/Zürich/New York: Olms, 237-308.

Nikula, Kristina (2012): Samspelet mellan text och bild i enspråkigt svenska ordböcker. I: *LexicoNordica* 19 (dette bind).

NRG = Jan Terje Faarlund, Kjell Ivar Vannebo & Svein Lie (1997): *Norsk referansegrammatikk*. Oslo: Universitetsforlaget.

Svenonius, Peter (1996): The Verb-Particle Alternation in the Scandinavian Languages. <<http://ling.auf.net/lingbuzz>> (september 2011).

Zgusta, Ladislav (1971): *Manual of lexicography*. The Hague: Mouton.

10. LexicoNordica udkommer både som trykt tidsskrift og i en **internetudgave**. Ved indsendelse af et bidrag til redaktionen erklærer forfatterne sig derfor indforstået med både trykt og elektronisk publicering.