



LEDA

LEKSIKOGRAFER
I DANMARK

LEDA-Nyt nr. 53 – April 2012

ISSN 1603-7006

Redigeret af Inger Schoonderbeek Hansen og Liisa Theilgaard

Meddelelser fra bestyrelsen	s. 2
Indkaldelse til medlemsmøde	s. 3
Indkaldelse til generalforsamling	s. 4
Ordbog over dansk tegnsprog – www.tegnsprog.dk	s. 5
Nordisk netordbog for børn	s. 11
Nye ordbøger 2011	s. 16
Rapport fra eLex 2011	s. 17
Mindeord: Finn Køster	s. 21
Kend din leksikograf!	s. 22
Annoncering af LREC, Istanbul	s. 24
Annoncering af ICHLL, Jena	s. 24
Annoncering af SIDG, Wien	s. 24
Annoncering af MUDS 14, Århus	s. 25
Annoncering af Euralex 2012, Oslo	s. 26

Medlemsmøde og generalforsamling

Paneldebat om fremtidens ordbøger

Onsdag 25. april 2012 kl. 15.00

Københavns Universitet Amager

Njalsgade 136

Bygning 27, lokale 27.0.09

Efter generalforsamlingen er der mulighed for at gå ud at spise sammen (for egen regning) på en restaurant i nærheden. Tilmelding på mødet.

Meddelelser fra bestyrelsen

Der indkaldes til generalforsamling onsdag den 25. april 2012 kl. 17.15 på Københavns Universitet Amager, Njalsgade 136, lokale 27.0.09

Se dagsorden mv. på side 4.

Følg med på LEDA's hjemmeside leksikografer.dk.

LEDA-Nyt udgives af Foreningen af Leksikografer i Danmark (LEDA) og udkommer to gange om året. LEDA-Nyt udsendes fast til LEDA's medlemmer og kan desuden – ligesom ældre numre af tidsskriftet – downloades fra LEDA's hjemmeside (www.leksikografer.dk). Bidrag til næste nyhedsbrev sendes til formanden inden 1. oktober 2012. Dette nummer er redigeret af Inger Schoonderbeek Hansen og Liisa Theilgaard. LEDA er en forening der fungerer som forum for praktisk og videnskabeligt ordbogsarbejde og ledes af en bestyrelse, som frem til næste generalforsamling består af:

Liisa Theilgaard (DSL) (formand), lt@dsl.dk / liisa.theilgaard@gmail.com
Sussi Olsen (KU, Center for Sprogteknologi), næstformand, saolsen@hum.ku.dk
Anita Ågerup Jervelund (Dansk Sprognævn), kasserer, aagerup@dsn.dk
Ken Farø (KU, Institut for Engelsk, Germansk og Romansk), kenfaroe@hum.ku.dk
Inger Schoonderbeek Hansen (AU, Jysk Ordbog), jysis@hum.au.dk
Lena Wienecke Andersen (KU, Nordisk Forskningsinstitut), suppleant, wienecke@hum.ku.dk
Jacqueline Levin (Ordbogen A/S), suppleant, jrl@ordbogen.com

Kontingent for medlemskab af Foreningen af Leksikografer i Danmark (LEDA) 125,- DKK (kan indbetales på girokonto 1551 139-1968). Foreningens CVR-nummer er 31193958.
Kontingent for medlemskab af Nordisk Forening for Leksikografi (NFL) 180,- DKK (200,- NOK).
Henvendelse vedrørende indmeldelse og adresseændring sker til foreningens kasserer.
Øvrige henvendelser rettes til formanden.

Medlemsmøde

Findes der ordbøger om 20 år, sådan som vi kender dem i dag?

Paneldebat om fremtidens ordbøger

Københavns Universitet Amager

Njalsgade 136

Onsdag 25. april 2012

Kl. 15.00 i bygning 27, lokale 27.0.09

- Hvordan ser en ordbog ud om 20 år? Indeholder den flere/færre informationer? Er den mono-/polyfunktionel?
- Smelter en ordbog sammen med andre informationskilder/værktøjer?
- Vil en ordbog være et synligt værktøj for brugeren eller en usynlig hjælp?
- Og vil en ordbog i fremtiden være en softwareskal, der henter oplysninger/indhold fra andre steder og præsenterer dem, jf. Wordnik?
- Er det overhovedet nødvendigt at kunne skrive korrekt – på dansk og udenlandsk – eller kan vi overlade det til computeren at rette det til automatisk eller guide os?
- Kan man opretholde copyright på ordbøger eller ordbogsdata?
- Er det brugergenererede/ikkekvalitetssikrede indhold godt nok, så der ikke er brug for egentligt leksikografisk arbejde? Og hvis der er, hvem skal i fremtiden betale for det leksikografiske arbejde, eller bliver det en con amore-beskæftigelse?

Paneldeltagere:

Esben Alfort, erhvervs-ph.d.-studerende inden for ontologisk klassifikation, Ankiro og CBS

Sabine Kirchmeier-Andersen, direktør for Dansk Sprognævn

Anders Søgaard, lektor i sprogteknologi inden for statistisk natursprogsbehandling og maskinlæring, Københavns Universitet

Nanna Vesterdorf, tekster og gymnasielærer i engelsk

Lars Trap-Jensen, ledende redaktør, Det Danske Sprog- og Litteraturselskab

Ordstyrer:

Liisa Theilgaard, formand for LEDA

Indkaldelse til generalforsamling

Onsdag den 25. april 2012 kl. 17.15

Københavns Universitet Amager, Njalsgade 136, lokale 27.0.09

Dagsorden ifølge vedtægterne:

1. Valg af dirigent og referent
2. Formanden aflægger beretning
3. Kassereren fremlægger det reviderede regnskab
4. Fastsættelse af kontingent
Bestyrelsen foreslår at kontingentet fastholdes på 125,- DKK pr. år.
5. Valg af formand
På valg er Liisa Theilgaard (indvalgt 2004).
6. Valg af to bestyrelsesmedlemmer og to suppleanter
På valg er bestyrelsesmedlemmerne Inger Schoonderbeek Hansen (indvalgt 2004) og Sussi Olsen (indvalgt 2010) samt suppleanten Lena Wienecke Andersen (indvalgt 2010).
7. Valg af revisor og revisorsuppleant
På valg er Christian Becker-Christensen og Pernille Neergaard-Dibbern.
8. Forslag
9. Eventuelt

Vedr. pkt. 5:

Liisa Theilgaard har været medlem af bestyrelsen i 8 år og kan derfor ikke genvælges. I stedet modtager Sussi Olsen valg som formand.

Vedr. pkt. 6:

Inger Schoonderbeek Hansen har været medlem af bestyrelsen i 8 år og kan derfor ikke genvælges. Lena Wienecke Andersen modtager genvalg som bestyrelsesmedlem, og som nyt bestyrelsesmedlem modtager Merete K. Jørgensen, DSL, valg.

Som ny suppleant modtager Lotte Follin, Gyldendal, valg.

Endvidere har Ken Farø (indvalgt 2011) ønsket at trække sig fra bestyrelsen. I hans sted indtræder suppleanten Jacqueline Levin (indvalgt 2011) som ordinært bestyrelsesmedlem for en etårig periode. Derfor skal der ekstraordinært vælges en ny suppleant for en etårig periode.

Vedr. pkt. 7:

Christian Becker-Christensen og Pernille Neergaard-Dibbern modtager genvalg som revisor og revisorsuppleant.

Andre kandidater er naturligvis velkomne til at melde sig til formanden, Liisa Theilgaard (lt@dsl.dk), eller på mødet.

Der vil blive serveret et glas og en snack, og efterfølgende vil der være mulighed for at gå ud at spise (for egen regning).

Ordbog over Dansk Tegnsprog – www.tegnsprog.dk

Af Jette H. Kristoffersen og Thomas Troelsgård

Indtil slutningen af 1900-tallet var næsten alle tegnsprogsordbøger udformet som alfabetisk ordnede – eller eventuelt emnegrupperede – lister af ord fra et talt/skrevet sprog med tilhørende fotos eller tegninger af tegnækvivalenter. At en egentlig tegnsprogsleksikografi var stort set fraværende, hang sandsynligvis sammen med at tegnsprog frem til anden halvdel af 1900-tallet i almindelighed ikke blev opfattet som et sprog, men som et kommunikationshjælpemiddel, en visuel tydeliggørelse af ordene, eller i hvert fald de betydningstunge ord fra f.eks. dansk, eventuelt “krydret” med pantomimiske effekter. I sidste halvdel af 1900-tallet skete der imidlertid store fremskridt inden for tegnsprogsforskningen; det blev bl.a. påvist at tegnsprog følger andre grammatiske regler end de omgivende tale-/skriftsprog, at tegn ikke nødvendigvis er ækvivalenter til bestemte ord, og at tegndannelse formelt kan beskrives parallelt med orddannelse – kort sagt at tegnsprog *er* sprog. Og med A Dictionary of American Sign Language on Linguistic Principles af William Stokoe m.fl. fra 1965 grundlagdes en egentlig tegnsprogsleksikografisk tradition.

I 1980’erne dukkede de første elektroniske tegnsprogsordbøger op efterhånden som computerbaserede værktøjer blev lettere tilgængelige, og endnu en teknisk landvinding kom da det blev muligt at bruge digitale videoklip til gengivelse af tegn. Videoklip er i dag den dominerende – i mange ordbøger den eneste – metode til visning af tegn og sprogksempler i elektroniske tegnsprogsordbøger.

Der er dog stadig en række udfordringer inden for tegnsprogsleksikografien, f.eks. i forbindelse med notation af tegn, idet der ikke findes nogen almindeligt udbredt skriftstandard for tegnsprog. Et andet problem er ordbogssøgning med tegnsprog som kildesprog, en facilitet der kræver muligheder for søgning på parametre som håndform, artikulationssted og bevægelse, idet tegnene ikke har nogen standardiseret skriftlig form.

Ordbog over Dansk Tegnsprog – type og indhold

I Ordbog over Dansk Tegnsprog kan der bl.a. søges på artikulationssted og håndform, og ordbogen er således i praksis en tovejsordbog med mulighed for opslag både via ord og via tegn. Ordbogen er dog ikke en typisk tovejsordbog idet alle opslag fører til samme type artikel. Artiklerne har tegn fra dansk tegnsprog (DTS) som opslagsord, og det er tegnenes betydning og brug der beskrives, mens den danske del blot udgøres af ækvivalenter, evt. suppleret med korte, disambiguerende forklaringer ved homografer eller stærkt polyseme ord. Denne løsning er valgt ud fra den formening at de fleste af ordbogens brugere vil have adgang til Nudansk Ordbog, Den Danske Ordbog, eller anden ordbog over moderne dansk, og dér vil kunne finde f.eks. betydningsbeskrivelser og bøjningsoplysninger for de danske ord. Hertil kommer at ordbogsprojektets formål er at beskrive det centrale tegnforråd i DTS, ikke at udfærdige en egentlig tovejsordbog. Ordbogens metasprog er dansk – en beslutning der er truffet idet fraværet af en almindelig udbredt skriftlig form af DTS ville resultere i en meget videotung brugergrænseflade, hvis metasproget var DTS.

Ordbog over Dansk Tegnsprog rummer i sin første udgave knap 2.000 artikler. De fleste af disse beskriver egentlige tegn, mens en lille del er “byggeklodsartikler”, der bl.a. beskriver affikser, bogstavtegn fra håndalfabetet, og de mest produktive proformer, en slags “rødder” (i form af bestemte håndformer) der ved tilføjelse af orientering og bevægelse kan udtrykke hvordan entiteter af en bestemt beskaffenhed (afspejlet af den valgte håndform) bevæger sig eller er placeret.

Artikellayout

Artiklerne i ordbogen rummer fire vinduer (jf. figur 1) der vises side om side:

- Artikelhoved

I hovedet vises tegnlemmaets glosse¹ og et foto af den mest prominente sekvens i tegnet. Desuden vises ikoner for første forekommende artikulationssted og håndform. Endelig er der to knapper til åbning af pop op-vinduer; et med en udskriftsvenlig oversigt med de basale oplysninger om tegnet, et med en konkordans over alle de eksemplarsætninger i ordbogen som opslagstegnet forekommer i.

- Videovindue

Ved åbning af en artikel, dvs. klik på et tegn i resultatlisten eller klik på en henvisning, vises tegnets grundform i videovinduet. Hvis der er registreret variantformer af tegnet, vises der over videoafspilleren særlige play-knapper til visning af disse. Ordbogens øvrige typer af videoklip, sætningseksempler og sammensætninger, aktiveres via særlige play-knapper i betydningsvinduet (se herunder). Videovinduet har herudover knapper til slowmotion-afspilning og fuldskærmsvisning af den aktuelt valgte video.

- Betydningsvindue

Her beskrives tegnenes betydning og brug. Der gives en eller flere danske ækvivalenter for hver betydning. Nogle betydninger har i stedet for danske ækvivalenter, eller som supplement til disse, en beskrivelse af opslagstegnets funktion i den pågældende betydning – en løsning der typisk bruges for funktionstegn, der f.eks. udtrykker tempus eller modus. For en betydning kan der være anført mundbevægelser som typisk bruges sammen med opslagstegnet brugt i den pågældende betydning. De fleste betydninger har herudover en eller flere eksempelsætninger der viser tegnet i brug. Sætningerne er transskriberet glosse for glosse og er forsynet med en dansk oversættelse, og de kan ses i videovinduet ved klik på en særlig play-knap. Endelig kan der være anført særlige brugsrestriktioner der gælder for den pågældende betydning. For nogle tegn er der registreret semantisk uigennemskuelige faste tegnforbindelser. Disse anføres i betydningsvinduet, under de egentlige betydninger, med en play-knap til afspilning af videoklip af tegnforbindelsen, og med en dansk ækvivalent.

- Vindue til ekstra oplysninger

I dette vindue anføres henvisninger til homonymtegn og, i de særlige proform-artikler (se ovenfor), til såkaldt “frosne former”, dvs. leksikaliserede tegn baseret på den pågældende proform. Vinduet bruges også til visning af almindeligt forekommende, semantisk gennemskuelige tegnforbindelser.

¹ En unik betegnelse i form af et eller flere danske ord der associerer til tegnets hovedbetydning.

The screenshot displays the 'Ordbog over Dansk Tegnsprog' (Danish Sign Language Dictionary) interface, divided into four main sections:

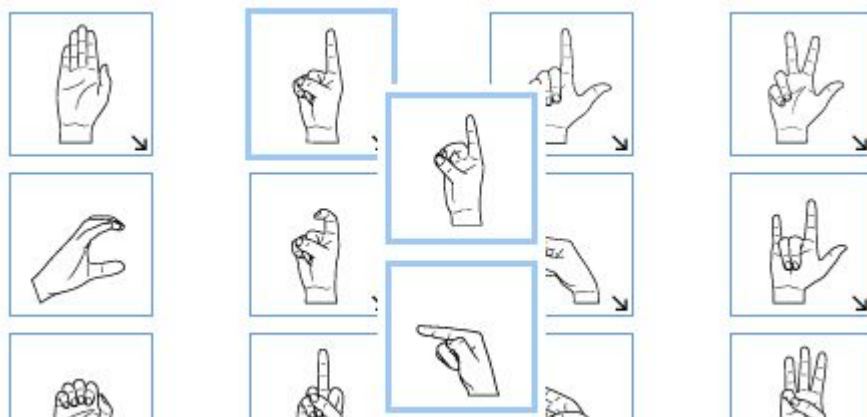
- Window 1 (Top Left):** Shows the sign for 'DEL' (to divide) with a hand diagram and the letter 'K'. It lists related words: *1. dele, del, detalje, fordele, fordeling, andel*. Below this, it shows the sign for 'HALV' (half) with a hand diagram and the letter 'K'. It lists related words: *2. afdeling, afsnit, filial*.
- Window 2 (Top Right):** A video player showing a signer. Below the video, it lists related words: *3. halv*. It also lists related words: *4. speciale, specialist*.
- Window 3 (Bottom Left):** A list of words: *5. gå op i detaljer*. It also lists related words: *6. gå op i detaljer*.
- Window 4 (Bottom Right):** A list of words: *7. gå op i detaljer*. It also lists related words: *8. gå op i detaljer*.

Figur 1. De fire vinduer i artikeldelen af Ordbog over Dansk Tegnsprog.

Opslag i ordbogen

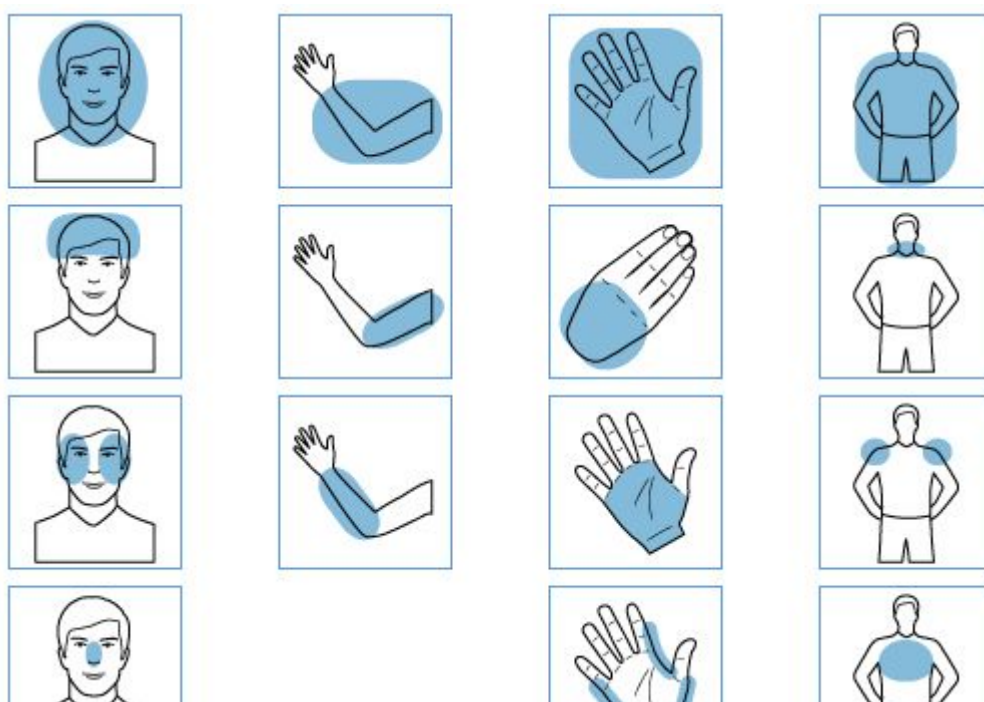
Ordbogens søgefunktion giver mulighed for opslag efter søgekriterier fra forskellige kategorier. Dels kan der søges på tegnets form via de fonologiske kategorier artikulationssted og håndform, dels kan der søges på tekst og på emne. Kriterier fra forskellige kategorier kan kombineres.

Ordbogen giver mulighed for at søge 65 håndformer, der vælges gruppevis via et pop op-vindue med 30 valgmuligheder, hver svarende til en håndform eller til en gruppe af beslægtede håndformer. Hvis man ønsker at søge på bestemte håndformer, kan uønskede håndformer i en valgt gruppe derefter fjernes individuelt fra listen over valgte søgekriterier. Dette valg er truffet for på den ene side at sikre at utrænede brugere søger relativt bredt, idet "nabohåndformer" medtages, mens avancerede brugere på den anden side har mulighed for at søge mere snævert på en eller flere specifikke håndformer uden hensyn til default-grupperingen. Figur 2 viser et udsnit af pop op-vinduet til valg af håndform, med strakt pegefingergruppen aktiveret.



Figur 2. Udsnit af pop op-vinduet til valg af håndform.

Ved søgning på artikulationssted vælges kriterier via et pop op-vindue med 21 valgmuligheder; femten der repræsenterer et specifikt artikulationssted, fire (øverste ikon i hver af de fire første kolonner) der repræsenterer hver sin gruppe af de førstnævnte femten steder, og endelig to der repræsenterer rummet foran kroppen – en overordnet og en der specifikt repræsenterer området foran ansigtet. Figur 3 viser et udsnit af pop op-vinduet til valg af artikulationssted, hvor det f.eks. ses hvordan man kan vælge enten hele armen (øverste ikon i anden kolonne) eller overarm henholdsvis underarm (de to nederste ikoner i anden kolonne). I de underliggende ordbogsdata optræder endnu et niveau.



Figur 3. Udsnit af pop op-vinduet til valg af artikulationssted.

Tekstsøgning foretages via et tekstfelt, og der søges både blandt danske ækvivalenter og tegnglosser, og blandt glosser og danske ord i eksempelsætningerne og deres danske oversættelser. Ved at udbrede søgningen til eksempelsætningerne sikres det at tegn som findes i sætningerne men som ikke har egen indgang, også kan findes, omend blot brugt i løbende tekst. Der kan søges på flere ord vha. anførselstegn, og med trunkering vha. spørgsmålstegn (ét vilkårligt tegn) og stjerne (nul eller flere vilkårlige tegn).

Ordbogen giver også mulighed for emnesøgning, idet ca. 85 % af artiklerne er forsynet med en eller flere emnemarkeringer. Emnerne vælges fra en pop op-liste med 70 emner, grupperet i 9 grupper og 5 undergrupper. Hvis der vælges et emne fra øverste eller mellemste niveau, inkluderes underordnede emner i søgningen.

Ved valg af søgekriterier fra forskellige kategorier gælder en eller-betingelse for kriterierne inden for hver enkelt kategori, og en og-betingelse mellem kategorierne. De aktuelt valgte søgekriterier vises i et panel i venstre side af skærmen. Figur 4 viser således en søgning på håndformen *lige udstrakt pegefinger* og emnet *religion*.

De enkelte valgte søgekriterier kan fravælges individuelt ved klik på selve kriteriet, eller samlet ved klik på Nulstil-knappen (se figur 4). I og med at kriterier fra fem forskellige kategorier kan vælges, kan det naturligvis ske at brugeren mister overblikket, og vi har derfor valgt at vise listen over valgte søgekriterier ved siden af søgeresultatlisten for at give bedre overblik ved justering af søgninger.



Figur 4. Søgekriterielisten (og en del af resultatlisten) for en søgning der kombinerer kriterierne håndform og emne.

Søgeresultater

Brugerne af Ordbog over Dansk Tegnsprog skønnes at udgøre en meget heterogen gruppe, både med hensyn til deres behov og med hensyn til deres lingvistiske forkundskaber og erfaringer som ordbogsbrugere. I et forsøg på at sikre hits ved så mange søgninger som muligt, også lidt upræcise søgninger, søges der derfor relativt bredt. Således returneres der ved søgning på en bestemt håndform alle tegn hvor den valgte håndform optræder, uden hensyntagen til håndformens prominens i de pågældende tegn. Tilsvarende medtages ved tekst-søgning ikke blot de danske ækvivalenter og tegnglosserne, men også glosser og danske ord i eksempelsætningerne og deres danske oversættelser. For at give bedre overblik over søgeresultaterne, der pga. den brede søgning ofte udgør en flere skærmhøjder lang liste, er

resultaterne prioriteret efter deres relevans i forhold til de valgte kriterier. For søgning på de manuelle træk artikulationssted og håndform vægtes placeringen, så fx en tidligt forekommende håndform vejer tungere end en sent forekommende (i et tegn med flere håndformer efter hinanden). For tekstsøgninger optræder match på tegnækvivalenter og -glosser før match på ord og glosser i eksempelsætningerne. I resultatlisten vises denne prioritering i form af 0-3 "relevansstjerner" ud for hvert hit, ligesom den afspejles i defaultsorteringen, hvor hits med tre stjerner placeres først. Hits med lige antal stjerner sorteres som default efter først artikulationssted, siden håndform, men via faneblade i toppen af resultatlisten (se figur 4) kan brugeren vælge sortering efter enten artikulationssted, "STED", eller håndform, "HÅNDFORM". Et sidste faneblad, "RESULTAT", vender tilbage til defaultsorteringen, med relevansstjerne som første kriterium.

Hvad nu? – projektets status og fremtid

Ordbog over Dansk Tegnsprog har været tilgængelig på internettet siden maj 2008 på adressen www.tegnsprog.dk. I efteråret 2011 er der herudover tilkommet en smartphonevenlig udgave af ordbogen på adressen www.m.tegnsprog.dk. Denne udgave giver dog kun mulighed for opslag via de danske ækvivalenter, og dens artikler indeholder kun tegnvideoer og ækvivalenter. I årene efter åbningen af første udgave har ordbogsprojektet kørt på nedsat kraft pga. manglende resurser, men vi har planer om at udvide antallet af artikler med 200 i løbet af 2012, og på længere sigt er håbet at skaffe midler til opbygningen af et DTS-korpus, der dels kan levere bedre konteksteksempler til ordbogens artikler, dels – via frekvensundersøgelser – kan udnyttes til at prioritere lemmakandidater til en yderligere udbygning af lemmabestanden, der, i hvert fald hvis vi skal kunne måle os med de større tegnsprogsordbøger i udlandet, gerne med tiden skulle komme op omkring 5-6.000 tegn.

Jette H. Kristoffersen
Projektleder
Center for Tegnsprog
Professionshøjskolen UCC
jehk@ucc.dk

Thomas Troelsgård
Projektmedarbejder
Center for Tegnsprog
Professionshøjskolen UCC
ttro@ucc.dk

Nordisk børneordbog på nettet

Af Pia Jarvad

Nordisk børneordbog er en ordbog mellem primært dansk, norsk og svensk, og den er beregnet for børn og unge i alderen omkring 12 år. Det er en elektronisk ordbog som bliver tilgængelig på nettet og på mobilen, men ikke i bogform. Danske, norske og svenske ord og fraser vil blive søgbare fra yderligere finsk, færøsk, grønlandsk, islandsk og samisk. Alle ord og ordforklaringer bliver tilgængelige lydligt idet ord og tekst er indtalt på dansk, norsk og svensk. Målet med ordbogen er at tilbyde et redskab til børn for at bedre den nordiske sprogforståelse og træne Nordens børn i at læse, forstå og lytte til nabosprog. Den er gratis og kommer til at ligge på Nordisk Sprogkoordinations hjemmeside: [www.http://nordisksprogkoordination.org](http://nordisksprogkoordination.org). Den åbner ved skolestarten 2012.

Baggrund og organisation

Deklarationen om nordisk sprogpolitik blev vedtaget af de nordiske undervisnings- og forskningsministre i 2006. Baggrunden for deklARATIONEN var behovet for en ny nordisk sprogpolitik som skal sikre helhed og sammenhæng i Nordisk Ministerråds sprogsatsninger (jf. Delsing & Lundin Åkesson 2005). Helt centralt har spørgsmålet om børn og unges sprogforståelse, sprogkundskab og sprogindlæring været, og Nordisk Ministerråds politiske satsningsområde i denne periode er netop børn og unge. I Nordisk Sprogkoordination er børn således den primære målgruppe for arbejdet med at øge den nordiske sprogforståelse, og det sker fx gennem sprogkampagnen og visse andre tiltag. Nordisk sprogkoordination blev oprettet med start i begyndelsen af året 2009 med mig som konstitueret koordinator, og i den egenskab fremlagde jeg ideen om en internetbaseret nordisk børneordbog på sprognævnenes netværksmøde i Borgå, Finland, i september 2009. Netværksmødet vedtog at støtte projektet med midler til at sætte projektet i gang, og efterfølgende blev en projektgruppe nedsat. Det kom til at se således ud: Projektledelse: Norsk språkråd ved Torbjørge Breivik, som også var norsk og nynorsk redaktør, hovedredaktør og dansk redaktør: Pia Jarvad, Dansk Sprognævn, svensk redaktør: Birgitta Lindgren, Språkrådet i Sverige, og yderligere sprogkoordinator fra 1. november 2009: Bodil Aurstad, Nordisk Sprogkoordination, it-konsulent: Jakob Halskov, Dansk Sprognævn. Vi fremlagde en plan for arbejdet, for redaktionsprincipper og økonomi, og det var basis for bevillinger fra Nordplus Sprog- og kulturprogrammet, Nordisk Kulturfond, Språkrådet i Sverige, Språkrådet i Norge, Dansk Sprognævn og Sprogkampagnen i Norden.

Ordbogen

Funktionaliteten i ordbogen vil være sådan at brugeren kan skrive et dansk, norsk eller svensk søgeord, og programmet giver et forslag til ordet på brugerens eget sprog. Brugeren kan søge på sit eget sprog og kan bede om forslag på et andet eller begge de to andre sprog, eller bede om alle tre sprog. Brugeren kan også søge i alle tre sprog samtidigt. Brugeren kan vælge omfanget af information: ord til ord, ord til ord med forklaringer, ord med lyd på, og oplæst tekst med ordet i de tre sprog. Søgeordet, som danner udgangspunkt for opslag, kan være i bøjede form, fx kan man slå op på *blommerne* og finde *blomme*, på *spiste* og finde *spise*. Det er gjort ud fra den antagelse at det ikke altid er nemt at lemmatisere på et fremmed sprog, slet ikke for børn, og slet slet ikke for børn der ikke kan sproget.

Tilkobling af grønlandsk, islandsk, færøsk, samisk og finsk

Projektet vil ikke i første omgang blive en mangesproget ordbog, men vil give indgang for grønlandsk, islandsk, færøsk, finsk og samisk (nordsamisk) til den centrale base. Det gøres

ved at sende den danske liste med forklaringer mv. til oversættelse til grønlandsk, islandsk, færøsk og den svenske og norske til oversættelse til finsk og samisk. Den grønlandske liste vil så rumme opslagsord + forklaring på grønlandsk og en kobling til tilsvarende ord med forklaring og lyd i dansk, norsk og svensk. Hermed giver vi brugere søgemuligheder fra disse sprog ind til kerneordbogen med dansk, norsk og svensk. Det giver alle dem som ikke har dansk, norsk eller svensk som deres førstesprog i hele det samlede Norden mulighed for at vide hvad et dansk, norsk eller svensk ord betyder, og hvordan det lyder som ord og i en sætning. Ved en senere udvidelse af ordbogen vil der kunne være mulighed for at føje lyd til de nye sprogs ord og forklaringer.

Udvalget af ord

Det er målet at ordbogen skal rumme de 3.500 almindelige ord som er centrale for målgruppen børn i alderen omkring 12 år. Til sammenligning indeholder Politikens Lille Danske Ordbog omkring 6.000 ord i 2001-udgaven, Gyldendals Min første røde ordbog, 2009, indeholder 2.600 ord og den norske Barnas store blå ordbok, norsk, 2006 under 2.000, og en sammenligning af ordudvalget viser at forskellige redaktører opfatter det centrale ordforråd ganske forskelligt. Men børn omkring 12 år er også en højst uensartet gruppe af tweens med orientering mod ungdomskultur inden for musik, de sociale medier og forskellige ungdomsgrupperinger, men de er også små børn og små voksne så det synes ikke at være en nem opgave at dække de unges særlige univers med blot 3.500 ord hvoraf de første 1.000 ord dækkes af kerneord fra hyppighedslistor eller som de fx er præsenteret hos Ruus 1995. Som udgangspunkt for udvalget af ord leverede jeg en elektronisk base med knap 3.500 ord med ordforklaringer og for at sikre en vis repræsentativitet for det almene ordforråd for målgruppen er der i særlig grad fokuseret på ord fra disse temaer ved brug af wordnet.dk og tesaussurser:

- kroppen, hovedet, tøj
- familie, højtider, fødselsdag
- huset, lejlighed, værelse, stue, køkken, badeværelse
- skolegård, klasseværelse, begreber i skolen og undervisning
- rollespil, eventyr, fiktion, musik, sport, rejse, camping
- byen, butikker
- skoven, stranden, havet, planter, dyr

Redaktionsgruppen har haft lange drøftelser om ordudvalget, og der er nok også forskellige traditioner for ordbogsarbejde i de tre sprogsamfund. Fx var der et ønske om at medtage 10 akvariefisk, 10 ord fra ridesport osv. for andre hobbyer, men da vi ikke selv ved en hurtig umiddelbar test kunne finde på meget mere end *guld fisk*, *slørhale* og *guppy*, var det blot dem der kom med, og princippet om et bestemt antal blev droppet. Også et ønsket princip om at medtage alle ord i et ordnet, af systematiske årsager, blev kasseret, og argumentet om at når *gedde* er med, så skal *aborre* også være med, blev erstattet af det mere gængse princip om ords og tings udbredelse i de tre sprogsamfund. Vi har været enige om at særord i dansk (fx *bat*, jf. norsk og svensk *racket* som både bruges i bordtennis, tennis og badminton), norsk (fx *gaupa* 'los') og svensk (fx *affisch* 'plakat') skulle være godt repræsenteret, mens kun yderst almindelige fremmedord kunne komme med (fx *adoptere*, *vegetar*), og sjove ord som svensk *aktersnurra* 'påhængsmotor' overlevede ikke, ligesom det danske frisind inden for sex og alkohol heller ikke gjorde det. En særlig forpligtelse er der over for ord som refererer til særlige kulturelle forhold som er forskellige i de tre samfund. Det er ord som dansk *påskehare*, *adventskrans*, eller svensk *midsommarstång*, *valborgsmässoafton*.

Et blandt sprogfolk yndet ordfelt er de lumske ligheder, fx *rolig* i norsk = 'morsom' og i dansk 'stille' har en mere nedtonet rolle i ordbogen ud fra den antagelse at for målgruppen

er alle ord lumske. De er med i ordbogen på samme måde som andre ord i ordforrådet. En lang række ord er ens i sprogene og har samme betydning, fx *bord*, *stol*, *skole*, men det ved den 12-årige jo ikke på forhånd, og de er derfor med. Andre ord der er ens eller kun med lille ortografisk variation, kan have delvis forskelligt betydning, fx dansk *kende*, norsk *kjenne* og svensk *känna*. Også de er medtaget. Forskellige ord i sprogene refererer til samme indhold, fx dansk *løbehjul* svarer til norsk *sparkesykkel* og svensk *sparkcykel*, og ord som kun findes i et eller to af sprogene medtages, fx *bedstefar*, *bestefar* som mangler tilsvarende ord i svensk. I den svenske del forklares ordet så i stedet. Til tider svarer et ord i norsk og dansk til to ord i svensk, fx *svane* 'fugl' svarer til *svan* 'fugl' og verbet *avta* 'bli mindre', og derfor peges der fra et ord i dansk/norsk til to i svensk og omvendt peger to ord i svensk på et og samme ord i norsk og dansk. I et tilfælde som *solskoldet* i dansk, som svarer til *solbrent* i norsk og *solsvedd* i svensk, mens svensk *solbränd* og dansk *solbrændt* svarer til norsk *brun* er der pegninger på kryds og tværs.

Ordforklaringer

At lave en tresprogsordbog kan gøres på flere måder, her valgte vi af praktiske årsager at tage udgangspunkt i den base på ca. 3.500 ord som jeg udarbejdede på basis af Pia Jarvad: Politikens Lille Danske Ordbog, 2001, og dermed blev dansk udgangspunktet. I modsætning til Politikens Lille Danske Ordbog valgte vi at benytte "cobuild"-ordforklaringer, altså ordforklaringer hvor opslagsordet gentages i forklaringen, af typen

Bikube *en bikube er det sted hvor bierne bor*

Bage *når du bager, laver du kager og brød i en ovn*

Denne form for ordforklaring er pædagogisk god, og i oplæsningen – og den senere aflytning – er det en klar fordel at ordet står inde i sætningen. Kun i enkelte tilfælde som fx ved præpositioner er der brugt eksempler på ordet i en kontekst, fx

Ad *vi følges ad til arbejde; der kom tre ad gangen; vi gik hen ad gaden*

Udgangspunktet var som sagt den danske base bestående af opslagsord + bøjningsformer som bruges til søgning + ordklasse + ordforklaring og med et felt hvor den indtalte lydfil indsættes samt et supervigtigt felt så ordene i sprogene kan kædes sammen: et unikt nummer for hvert ord med oplysning om sprog. Basen er i xml i et gratisprogram som nok gav besværligheder for nogle redaktører, og som ikke havde særligt mange funktioner.

Strukturen blev som så:

```
<ARTIKEL SPROG="DA" xml:id="NOB11000">
<OPSLAGSORD>slot</OPSLAGSORD>
<LYD>DA-NOB11000-slot</LYD>
<ORDKLASSE>sb</ORDKLASSE>
<BOEJNING>
<FORM>slottet</FORM>
<FORM>slotte</FORM>
<FORM>slottene</FORM>
</BOEJNING>
<DEFINITION>et slot er en stor, fornem bygning</DEFINITION>
<KOMMENTAR></KOMMENTAR>
</ARTIKEL>
```

Basen sendtes til den norske og svenske redaktør som udfyldte det tilsvarende norske/nynorske/svenske opslagsord + bøjningsformer + en nær oversættelse af den danske ordforklaring, fx dansk *banan* i svensk:

```
<ARTIKEL xml:id="NOB392" SPROG="SE">
<OPSLAGSORD>banan</OPSLAGSORD>
<LYD>SE-NOB392-banan</LYD>
<ORDKLASSE>sb</ORDKLASSE>
<BOEJNING>
<FORM>banan</FORM>
<FORM>bananen</FORM>
<FORM>bananer</FORM>
<FORM>bananerna</FORM>
</BOEJNING>
<DEFINITION>en banan är en lång böjd frukt med gult skal</DEFINITION>
<KOMMENTAR></KOMMENTAR>
</ARTIKEL>
```

Her har det vist sig at det er svært at undertrykke redaktørers kreativitet og ønsker om forbedringer, og da der ikke har været tid til fornyede gennemgange, vil man møde uforklarlige uoverensstemmelser – som dog ikke betyder noget i den nationale liste, men som når man ser alle tre sprog sammen, antyder forskelle i sprogene som ikke er der. I et tilfælde som *banan* hvor princippet for forklaring er at nævne overbegreb (= frugt) eventuelt suppleret med prototypiske træk (krum, gul, skræl), kan nationale forskelle således forklares ved forskelle i oplevelsen af de prototypiske træk, jf. dansk, norsk, nynorsk og svensk *banan*:

en banan er en krum frugt med en gul skræl (dansk)
en banan er en gul, svaktbøydfrukt (norsk)
ein banan er ei gul, svaktbøygdfukt (nynorsk)
en banan är en långböjdfukt med gult skal (svensk)

Ved verberne er der tre hovedtyper, at bruge infinitiv, at bruge *du* + præsens eller *man* + præsens, og her kan det komme til at se forvirrende ud – er det udtryk for strukturelle forskelle sprogene imellem, kan brugeren spørge sig, fx

Bage når du bager, laver du kager og brød (dansk)
Baka å baka er å laga kaker og brød (nynorsk)

eller

bestå at bestå er at klare sig til fx eksamen (dansk)
bestå du kan bestå en prøve (norsk)

eller

aleine når du er aleine, er du ikke sammen med noen (norsk)
ensam när man är ensam är man inte tillsammans med någon (svensk)

Lydfiler er kostbare

At producere lydfiler er kostbart, men højt prioriteret i projektet. Lyden vil blive produceret ved indlæsning, og man tilstræber unge indlæsere, men det ser ikke ud til at målgruppens

alder omkring 12 år også vil kunne indlæse, indlæserne er ældre. Ord + forklaring indlæses på dansk, norsk, nynorsk og svensk, og lydfileerne kobles til ordene i databasen. Syntetisk tale blev overvejet fordi det var en billigere løsning, men resultatet var ikke godt nok. Når målet for ordbogen har været at bidrage til øget sprogforståelse mellem sprogene, måtte lyd-kvaliteten være i orden så der ikke kunne være tvivl om hvad der høres. Valget af den dyrere løsning fik afsmittende effekt på længden af ordforklaringerne som i visse tilfælde, når de var over 10 ord, måtte forkortes.

Litteratur

- Aurstad, B. m.fl. (2006): *Barnas store blå ordbok, norsk*, Kunnskapsforlaget.
- Breivik, Torbjørg (under udgivelse): Nordisk barneordbok på nett. I: *LexicoNordica* 2013.
- DanNet: <http://wordnet.dk>.
- Dansk Sprognævn: <http://www.dsn.dk>.
- Delsing, Lars-Olof og Åkesson, Katarina Lundin (2005): *Håller språket ihop Norden? En forskningsrapport om ungdomars förståelse av danska, svenska och norska*. TemaNord 2005:573, Nordisk ministerråd.
- Jakobsen, Kirsten og Katrine Vissing (2009): *Min første røde ordbog – Nudansk*, Gyldendal.
- Jarvad, Pia (2001): *Politikens Lille Danske Ordbog*, 3. udgave, Politikens Forlag.
- Nordiska Kulturfonden, København
- Nordisk Kulturfond: <http://www.nordiskkulturfond.org>.
- Nordisk Ministerråd: <http://www.norden.org>.
- Nordisk Sprogkampagne 2011: <http://www.norden.org/no/aktuelt/nyheter/spraakkampanje-skallstyrke-det-nordiske-kulturfellesskapet>.
- Nordisk Sprogkoordination: <http://nordisksprogkoordination.org>.
- Språkrådet i Norge: <http://www.sprakradet.no>.
- Språkrådet i Sverige: <http://www.sprakradet.se>.

Pia Jarvad
Seniorforsker
Dansk Sprognævn
jarvad@dsn.dk

Nye ordbøger 2011

Af Liisa Theilgaard

- Bulgarsk-dansk ordbog** af Lise Bostrup og Vladimir Stariradev, 1. udgave, Bostrup, 2011.
- Dansk-thai ordbog** af Donald Shaw og Suphat Sukamolson, 3. udgave, Special-Pædagogisk Forlag, 2011.
- Dansk-ungarsk lommeordbog, magyar-dán zsebszótár** af Claude Kolozs Nagy og Eva Tardoskeddi Nagy, eget forlag, 2011.
- Dansk-ungarsk ordbog, Dán-magyar szótár** af Ágnes Sitkeiné Szira, Margit Nielsen og József Szira, 3. udgave, Special-Pædagogisk Forlag, 2011.
- Den Danske Betydningsordbog** (online på www.ordbogen.com) af Henning Bergenholtz m.fl., 1. udgave, Center for Leksikografi, 2011.
- Den Danske Grammatik- og Staveordbog** (online på www.ordbogen.com) af Henning Bergenholtz m.fl., 1. udgave, Center for Leksikografi, 2011.
- Den Danske Skriveordbog** (online på www.ordbogen.com) af Henning Bergenholtz m.fl., 1. udgave, Center for Leksikografi, 2011.
- Den Danske Synonymordbog** (online på www.ordbogen.com) af Henning Bergenholtz m.fl., 1. udgave, Center for Leksikografi, 2011.
- Ejendomsordbogen fransk-dansk-fransk** (online på www.ordbogen.com) af Liselotte Kruse og Patrick Leroyer, 1. udgave, 2011.
- Engelsk-dansk ordbog, fast ejendom, finansiering** af Lars Kjær, 1. udgave, Jurist- og Økonomforbundet, 2011.
- Holbergordbogen** (online på holdbergordbog.dk) af Aage Hansen og Sv. Eegholm-Pedersen, Det Danske Sprog- og Litteraturselskab, 2011.
- Hvad er det nu, det hedder?** (online på www.ordbogen.com) af Henning Bergenholtz m.fl., 1. udgave, Center for Leksikografi, 2011.
- Min første billedordbog** af Heather Heyworth, oversat af Lena Lamberth, 1. udgave, Lamberth, 2011.
- Min lille ordbog** af Kathleen Amant, oversat af Marie Brocks Larsen, 1. udgave, Turbine, 2011.
- Ordbog for slægtsforskere** af Heini Madsen, 3. udgave, Støvring – Skúvanes, 2011.
- Ordbog over skældsord for dannede mennesker** af Erik Bøegh, 1. udgave, Veivad, 2011.
- Synonymordbog** af Thomas Ingemann, Henrik Andersson, Marianne Holmen og Kirsten Sydendal, 1. udgave, i serien Gyldendals Røde Ordbøger, Gyldendal, 2011.
- Ungarsk-dansk ordbog, Magyar-dán szótár** af Ágnes Sitkeiné Szira, Margit Nielsen og József Szira, 2. udgave, Special-Pædagogisk Forlag, 2011.

Liisa Theilgaard
Seniorredaktør
Det Danske Sprog- og Litteraturselskab
lt@dsl.dk

eLex 2011

Af Jacqueline Levin

I november sendte Ordbogen.com to novicer (min kollega Ane Kristensen og mig) af sted til Slovenien for at deltage i eLex 2011. Konferencen var arrangeret af Trojina-instituttet ved universitetet i Ljubljana og var henlagt til et conferencecenter i det maleriske Bled en times kørsel fra hovedstaden. Med overskriften “Electronic lexicography in the 21st century: new applications for new users” og debattemaet “Will there still be dictionaries in 2020?” var der lagt op til fremadskuen med god plads til nytænkning.

Simon Krek lagde som første hovedtaler ud med sit bud på hvad der kan ske inden for leksikografien det næste årti hvis e-leksikografi, NLP og sprogteknologi kan finde sammen. Visionerne inkluderede for eksempel realtidsundertekstning af tv-programmer og automatisk indeksering af video/tale. Det lyder som ren science fiction, men på nogle punkter er man nået langt. Alene inden for området automatisk ekstraktion af definitioner har der allerede været afholdt flere konferencer og workshops. Drømmen er den automatiserede netsøger som selv kan registrere når eksisterende ord får ændrede eller helt nye betydninger. I forhold til leksikografernes arbejde er det dog et problem at disse fremskridt gøres af og for data-lingvister fordi de arbejder på at producere data der har maskiner, ikke mennesker, for øje.

Den opgave leksikograferne skal skabe for sig selv i fremtiden, må derfor være at sikre at leksikografisk information præsenteres i det lettilgængelige format folk har lært at forvente fra fx deres smartphone hvor datapræsentationen ikke kun kontekstualiseres i forhold til tidligere søgninger, men også til hvor man til enhver tid befinder sig rent fysisk. Da brugerne ikke vil påtage sig opgaven med at fortælle os hvem de er, hvor de er, hvad de laver osv., må vi bruge de muligheder sprogteknologien giver til at påtage os opgaven for dem. Desuden skelner de ikke mellem leksikografisk, encyklopædisk eller grammatisk information, så de finder ikke vores information hvis vi gemmer den i ordbøger – det leksikografiske materiale skal være hvor brugerne søger, ikke omvendt. Krek konkluderede at fremtiden for leksikografer afhænger dels af et brud med traditionelle formater, dels af at vi finder os med at leksikografien blot bliver en mindre klart adskilt underafdeling af overbegrebet “søgning”.

Efter denne inspirerende start på konferencen rendte vi direkte ind i en blindgyde i vores forsøg på at styre efter indlæg om korpuslingvistik fordi det kunne have relevans for vores it-baserede virksomhed. Vi måtte desværre sande at hvis man ikke tidligere har beskæftiget sig med emnet, så er det relativt sort snak. Efter frokostpausen sadlede vi derfor om og gik efter foredrag om ordbøger til undervisningsbrug som vi havde bedre forudsætninger for at få udbytte af.

I denne kategori var Robert Lews foredrag om hans og Roger Mittons undersøgelse af hvordan stavefejl håndteres af de store engelsksprogede lærerordbøger på nettet. Fælles for ordbøgerne var ironisk nok at de klarer opgaven dårligt hvis brugerens fejl skyldes ringe engelskkundskaber. Staver man forkert, foreslår de alternative ord, men da denne funktion er baseret på (relativt fejlfri) engelsksprogede korpora, kan ordbøgerne ikke regne ud hvilket ord der menes hvis det staves forkert. Dette forværres yderligere hvis fejlen begås af en ikke-engelsksproget bruger, som når fx en japaner kan finde på at skrive *crean* for *clean*. Forsøg med at træne ordforslagsfunktionen på korpora over fremmedsproglige brugeres fejl gav væsentlig bedre resultater.

Fra den brugermæssigt stik modsatte grøft holdt Erin McKean dagens afsluttende hovedtale om siden Wordnik for folk der vil gå på opdagelse i deres modersmål, forudsat at det er engelsk. Wordnik kalder sig ikke selv en ordbog da den sammenstykkedes på stedet af eksisterende information fra andre kilder, både via licensrettigheder og i form af ikke-

ophavsretligt beskyttet materiale. Definitioner hentes fra forskellige engelske ordbøger, og eksempler hentes fra korpusmateriale og dagligt opdateret information fra nyhedsmedier. Det betyder at selv hvis et ord (endnu) ikke er defineret i nogen ordbøger, kan man danne sig et indtryk af dets betydning ud fra dets brug. Suppleret med billeder fra Flickr, twitter-beskeder og en hoben brugergenereret indhold, såsom kommentarer og tematiserede lister, kan informationen for et velbeskrevet ord let overvælde, så sidens primære styrke synes at ligge i at den hjælper folk med at forstå ord der ikke er optaget i ordbøgerne (to tredjedele af trafikken stammer da også fra søgemaskiner).

Konferencens anden dag indledtes med Michael Rundells interessante gennemgang af det grundige arbejde der lå bag DANTE-databasen samt hans bud på et svar på det tilbagevendende spørgsmål på konferencen: Hvordan kan man tjene penge på leksikografi når folk ikke vil betale for ordbøger? Med en reference til datalingvistikens primære problem – at informationer er billige, men betydninger dyre – ser Rundell en kommerciel fremtid for leksikografisk gennemarbejdet materiale som råmateriale for datalingvistikken.

Ved aftenens rundbordssamtale skulle han og de øvrige deltagere forsøge at besvare spørgsmålet om hvorvidt der stadig vil findes ordbøger i 2020. Er den elektroniske ordbog bare endnu et skridt på vejen mod udslettelsen, eller vil ældre modeller overleve ved siden af nye former? Den overvejende holdning var at det vi tænker på når vi siger ordbøger i dag, ingen gang får på jord som andet end historisk kuriosum. Rundell pointerede at for det første vil folk ikke vente på at et ord optages i ordbøgerne på den traditionelle måde, og for det andet nægter de at betale for information. Hvis vi vil være med, er det derfor et spørgsmål om hvordan vi kan leve af det. Reklamer giver for lidt, men apps har givet nyt håb. Desuden mente han at der lå uudnyttet potentiale i at integrere ordbøger i andre produkter som fx e-bøger hvor opslag foretages direkte fra bogteksten.

McKean advarede mod at forveksle beholderen med indholdet: De digitale indfødte forventer ikke at få deres data i en bestemt beholder. De forventer en behagelig brugergrænseflade der ikke står i vejen for deres mål om at søge information uanset art og hvis vi ikke tilpasser vores produkter, går folk til andre, lettere tilgængelige kilder. Naturligt nok mente McKean at vi ikke kommer uden om halvautomatisk genereret indhold så vi kan give folk informationerne så snart de har brug for dem. Med udgangspunkt i at leksikale data bruges ikke kun af mennesker men også af programmer, indtog Piek Vossen det modsatte synspunkt af Kreks om at fremtiden ligger i at gøre maskinvenlige data tilgængelige for mennesker. Han argumenterede tværtimod for at hvis bare alle de menneskeskabte kvalitetsdata blev gjort læsbare for maskiner, så sidder der et milliontal af programmører ude i verden som laver apps af hvad som helst til hvem som helst. Hvis bare dataene gøres tilgængelige for dem, skal der nok blive lavet masser af programmer, og de gode vil blive populære. At lave gode apps er en opgave for programmører, ikke for leksikografer. McKean påpegede dog en vigtig ulempe ved denne fremgangsmåde: Hvis vi ikke selv styrer informationsstrømmen til og fra brugerne, ved vi ikke hvilken information de bruger, hvordan, hvor længe osv., og den information skal vi bruge til at forbedre vores produkter.

Magali Paquot mente at ordbøgernes overlevelsesmuligheder er blevet bedre, ikke ringere, fordi internettet har gjort dem trendy. Nye brugere finder vej til velfungerende sider som Wordreference.com, som har lige så mange brugere som de store avisers hjemmesider. Her er ingen smart NLP i spil, bare gode søgemuligheder på fraser og fora for specifikke sprogpar hvor brugerne kan hjælpe hinanden. Problemet med NLP-baseret indhold er også at nogle forlag tror at det sparer penge, men dataene skal stadig valideres. Svaret må derfor være NLP kombineret med brugerdeltagelse som fx på Wordreference og Wordnik. Desuden slipper forlagene ikke for at investere i forskning i hvem der bruger onlineordbøger, og hvordan de bruger dem. Det kan leksikografer hjælpe med at afdække.

Lørdag morgens hovedtaler, Sylviane Granger, talte om at nedbryde barriererne mellem elektroniske ordbøger og computerstøttet læring. Hun forestillede sig et lukket indlæringssystem som ikke blot indeholdt opgaver til eleverne, men som også registrerede de fejl lærerne fandt, lagrede dem i et korpus og fødte dem tilbage i systemet som dermed blev bedre til selv at genkende og advare imod typiske fejltyper og tilbyde skræddersyet hjælp.

I den spanske netbaserede lærerordbog DAELE, hvis udarbejdelse Irene Renaus gjorde rede for, forsøger man at udnytte de mange muligheder der ligger i en korpusbaseret netordbog, men i praksis er konceptet på nuværende tidspunkt så informationsmættet at det bliver uoverskueligt. Serge Verlinde kunne da også i sin hovedtale fortælle at testpersonerne på hans Base Lexicale du français simpelthen opgav at foretage sig noget som helst hvis informationsmængden på en side blev for stor, uanset hvor let den var at finde. Paqout identificerede skrivehjælp som den store mangelvare i de nuværende lærerordbøger og præsenterede sit projekt med at tilpasse hjælpen til teksttypen via en database over kollokationer som er velegnede i fagsprog generelt eller specifikt inden for bestemte fag – en interessant ide som må siges at være relevant for artikelskrivende universitetsfolk kloden over.

Udfordringen er som nævnt at præsentere informationen intelligent så brugerne ikke præsenteres for lister eller tabeller de selv skal tyde, men i stedet automatisk præsenteres for den information der passer på den givne kontekst. Daniël Prinsloo og Ulrich Heids ideoplæg til en grammatisk skrivehjælp var et interessant bud. Hvis man omformulerer tabelinformationerne til en seriel række af beslutningstrin, kommer brugeren i en og samme proces til selv at beskrive den givne situation, få en indføring i de relevante regler og få en løsning på en konkret opgave.

Set ud fra den lærendes perspektiv er der jo dog også den altid truende konflikt mellem hvor let man finder informationen, og hvor godt den fæster sig i hukommelsen. I sin gennemgang af japanske elektroniske multiopslagsværker måtte Yukio Tono indrømme at selv om en japansk gymnasieelev for en pris svarende til en iPhone kan få et apparat med op til 150 værker der dækker det komplette pensum i alle fag, rangerer japanernes engelsk-kundskaber lavt målt på TOEFL-score, og det har årtiers teknologisk tyvstart ikke ændret ved. Derfor indeholder de nyeste apparater også muligheden for at skrive noter eller selv at skrive vendekort til gloser for at styrke indlæringen. Anna Dziemankos undersøgelser af hvilken indflydelse ordbogspræsentationen har på fastholdelse af indlærte gloser bekræftede også at nogle af de elementer som på den ene side kan hjælpe ved tekstproduktion på den anden side kan hæmme reception og fastholdelse.

Der var bred enighed på konferencen om at forestillingen om den ultimative universalordbog for alle er død. Men hvad skal træde i stedet? Henning Bergenholtz' monofunktionelle ordbøger tilpasset specifikke brugersituationer har blandt andet fortalere som Theo Bothma som lukkede konferencen med sin præsentation af projektet med at lave en afrikaans ordbog over faste vendinger. Denne videreudvikling af konceptet bag Bergenholtz' Ordbog over faste vendinger skal desuden indeholde oversættelser mellem afrikaans og engelsk og danne grundlaget for yderligere sprogkoblinger. Ideen er at kunne præsentere databaseindholdet som seks forskellige ordbøger til seks forskellige foruddefinerede brugssituationer som brugerne selv vælger ved at klikke på den relevante knap.

For at vende tilbage til Kreks indledende pointe kan man dog spørge sig selv om folk overhovedet er interesserede i at tage stilling til hvilken brugssituation de befinder sig i? Eller er det vores opgave at regne det ud for dem og så kun præsentere dem for den information der er relevant i den givne sammenhæng? Flere af konferencedeltagerne gav i hvert fald udtryk for, at det vigtigste ikke er om ordbogen i den traditionelle form overlever årtiet ud, men derimod at leksikografisk kvalitetsmateriale når ud til folk, uanset præsentationsformen. Personligt glæder jeg mig meget til at se hvad dygtige designfolk kan stille op med datalingvisternes og leksikografernes bugnende lagre af information mellem nu og

2020.

Fra <http://www.trojina.si/elex2011/> linkes til både *proceedings* og til videooptagelser med tilhørende præsentationsdias for stort set alle foredrag i den store sal, inklusive rund-bordssamtalen.

Jacqueline Levin
Specialkonsulent
Ordbogen.com
jrl@ordbogen.com

Mindeord

Pensioneret universitetslektor og ordbogsredaktør Finn Køster 20. august 1939 – 21. januar 2012

Af Asgerd Gudiksen

Finn Køster blev cand.mag. i dansk og russisk fra Københavns Universitet 1969 og samme år ansat på det daværende Institut for Dansk Dialektforskning, Københavns Universitet (nu Afdeling for Dialektforskning), hvor han var ansat indtil pensioneringen 2006. Finn Køster deltog, lige fra han blev ansat, i udarbejdelsen af *Ømålsordbogen*, en ordbog over dialekterne på Sjælland, Lolland-Falster og Fyn med omliggende øer. Han var med i det hold af unge dialektforskere der fastsatte reglerne for hvordan ordbogsartiklerne skulle se ud, og har skrevet en stor del af de 10 bind der hidtil er kommet, bl.a. ordbogsartikler over store verber som *gøre* og *have*. Ved siden af ordbogsarbejdet har han skrevet en række tidsskriftartikler om ømålene, fx "Da jokket vi med Fyns Tidende" (om datid på fynsk) eller "Træk af sproget i byerne på Fyn". Alle artiklerne er præget af hans indgående viden om dialekter og hans udprægede evner til at gøre skarpe sproglige iagttagelser. Men Finn Køster var ikke nogen stuelærd; han understregede tværtimod altid hvor vigtigt det var at være i kontakt med de mennesker hvis sprog man studerer. 1.000 båndoptagne interview med ældre dialekttalende rundt omkring på øerne blev det til, uvurderlige kilder til sprog og kultur i det 20. århundredes landbosamfund. Vi vil på Afdeling for Dialektforskning mindes ham ikke kun som en indsigtfuld dialektforsker der gavmildt delte ud af sin viden, men også som en god kollega med en finurlig humor.

Asgerd Gudiksen
Lektor
Nordisk Forskningsinstitut
Københavns Universitet
asgerd.gudiksen@hum.ku.dk

Kend din leksikograf – Asgerd Gudiksen, Ømålsordbogen

Denne gang skal vi møde Asgerd Gudiksen fra Nordisk Forskningsinstitut ved Københavns Universitet, som arbejder på Ømålsordbogen, der er en sproglig-saglig ordbog over dialekterne på Sjælland, Møn, Lolland-Falster, Fyn med omliggende øer.

* * *

Hvornår og hvordan kom du til at arbejde med faget?

Jeg redigerede min første ordbogsartikel til Ømålsordbogen – verbet *hakke* – foråret 1994 som led i mit pligtarbejde som ph.d.-studerende ved det daværende Institut for Dansk Dialektforskning, nu Afdeling for Dialektforskning. Verbet var omhyggeligt udvalgt af min vejleder, Karen Margrethe Pedersen, fordi det var ukompliceret mht. udtale og bøjning, men gav en ret alsidig introduktion til arbejdet med de landbrugsfaglige emner.

Baggrunden for at jeg kom ind i arbejdet med dialektleksikografien, var at jeg havde skrevet speciale om suffikset *-vorn* i dansk med hovedvægt på de danske dialekter og dermed havde fået godt kendskab til Jysk Ordbogs og Ømålsordbogens samlinger. Det førte så til endnu et projekt om dialektal orddannelse og ph.d.-afhandlingen *-et-afledninger i dansk med hovedvægt på de danske øsmål*, der igen medførte at jeg blev adjunkt og senere – fra 2001 – lektor ved Afdeling for Dialektforskning med Ømålsordbogen som hovedopgave.

Hvad kan du lide ved leksikografi?

At arbejdet er så alsidigt – og det gælder måske særlig arbejdet ved en stor dialektordbog som Ømålsordbogen. Én dag arbejder man fx med et kompliceret udtaleafsnit hvor man får brug for viden både om fonetik og sproghistorie, en anden dag drejer det sig om at beskrive syntaktisk variation, og en tredje dag gælder det om at rede alle de processor ud man skal igennem når man lægger et lergulv, herunder finde ud af om det virkelig var en almindelig del af processen at jage fårene ind over det nylagte gulv.

Hvad bryder du dig ikke om ved feltet?

Det mindst morsomme ved at redigere Ømålsordbogen er arbejdet med de grønne andetleds-sedler – det er vi alle enige om. Opgaven går ud på at man slår alle de sammensætninger op der har det ord man er i gang med, som andetled – dels for at kontrollere betydningen, dels for at supplere materialet til udtalebeskrivelsen. I det hele taget kan det være noget besværligt at finde frem til materiale der befinder sig i en stor seddelsamling.

Har du nogen leksikografiske kæpheste?

At de samme sproglige forhold skal behandles på samme måde ordbogen igennem, et princip der kan være vanskeligt at opretholde i en ordbog som Ømålsordbogen der i mange tilfælde er baseret på redaktørens skøn. Det gælder fx i meget høj grad for brugen af generaliserende udbredelsesangivelser af typen Ø(spor opt), dvs. ‘sporadisk optegnet i hele ømålsområdet og muligvis almindeligt’ eller øF(alm), dvs. ‘almindeligt udbredt på Østfyn’.

Hvad er du mest stolt af at have lavet?

Mest stolt af? I Ømålsordbogen er det nok de store leksikonafsnit under *høst* og *jul*, eller er det måske *lade* v (23 spalter)? Jeg er under alle omstændigheder godt tilfreds med at have medvirket til at andelstidens landbrug får en mere dækkende saglig beskrivelse i ordbogen end oprindelig planlagt i 1930’erne, hvor redskaber og processer fra slutningen af 1800-tallet blev betragtet som for moderne og derfor uden videre interesse.

Hvad er din største bommert?

Jeg kan ikke komme i tanke om nogen spektakulær bommert der kan mere LEDA-Nyts læsere, men af og til støder jeg jo da på ærgerlige småfejl i de allerede trykte bind. Jeg håber dog i det mindste ikke at have lavet nogen “krumkagejern” (jf. næste svar!).

Fortæl en god leksikografisk anekdote

Det bedste jeg kan finde på, er den om ODS-redaktøren som definerede *krumkagejern* som ‘jern, hvormed krumkagerne krummes’. Det er det (jo) ikke (det er en slags vaffeljern som man bager krumkagerne i over ilden). Hvis det ikke er en anekdote, så er det i hvert fald en historie der kan få det til at løbe koldt ned ad ryggen på en leksikograf der jævnlig beskriver mange redskaber og arbejdsprocessor hun ikke har oplevet i virkeligheden.

Beskriv en typisk leksikografisk arbejdsdag/arbejdsproces for dig

På en typisk leksikografisk dag griber jeg ned i seddelkassen der hvor jeg er kommet til, og begynder at sortere sedlerne op efter udtale – eller hvis jeg har lavet udtaleafsnittet – efter betydning og fortsætter med at strukturere og skrive artiklen. En anden, ret tidkrævende leksikografisk opgave ved siden af redigeringen, som jeg dog ofte udfører i weekenden, er gennemlæsning af medredaktørernes artikler, især er det tidkrævende i de tilfælde hvor jeg læser igennem “med sedler”, altså holder manuskriptet op mod materialet, og desuden kontrollerer citater og eksterne henvisninger. Der er dog også dage hvor ordbogsarbejdet ligger næsten stille, fordi jeg er ved at færdiggøre en tidsskriftsartikel, arbejder med Afdelingen for Dialektforsknings tidsskrift Danske Talesprog som jeg er medredaktør af, eller noget helt tredje.

Hvilket opslagsværk kan du ikke undvære?

Ordbog over det danske Sprog er helt uundværligt. Men det er også svært at forestille sig hvordan man skulle lave Ømålsordbog uden etnologiske håndbøger som Gamle Danske Håndværk, Dansk Fiskeri før Industrialiseringen osv. osv.

Har du en leksikografisk våd drøm?

Ja – en digital dialektordbog med både lyd og sag integreret. Der skulle være lyd til citater fra vores korpus af udskrevne båndoptagelser, autentiske billeder, skematiske tegninger og filmede arbejdsprocesser. Men det er indtil videre kun en vag drøm. Foreløbig må vi koncentrere os om at få Ømålsordbogen færdigredigeret.

Asgerd Gudiksen
Lektor
Nordisk Forskningsinstitut
Københavns Universitet
asgerd.gudiksen@hum.ku.dk





Den ottende LREC-konference (Language Resources and Evaluation) finder sted i Istanbul 21.-27. maj 2012 med hovedkonferencen 23.-25. maj. Den organiseres af ELRA med hjælp fra skiftende organisationer.

LREC har gennem 14 år udviklet sig til den største konference om sproglige ressourcer og evaluering inden for sprogteknologi. LREC har til formål at skabe overblik over status for sprogteknologien, vise den allernyeste forskning og udvikling samt udveksle information om sproglige ressourcer og deres anvendelse. Evalueringsmetoder og værktøjer er også på dagsordenen.

Dette års konference har særligt fokus på taleteknologi og multimodale ressourcer og integrationen af disse med skriftlige ressourcer.

Læs mere på <http://www.lrec-conf.org/lrec2012/>.

6th International Conference on Historical Lexicography and Lexicology (ICHLL) 25th-28th July, 2012

Topics

We welcome contributions on any aspect of historical lexicography, historical dictionary research, or historical lexicology. Particularly papers addressing the following questions:

- How much diachrony is important for dictionaries of present-day language?
- What should a modern dictionary look like – published as printed or online version?
- How can corpus linguistics and dictionary work be combined?
- How can comparative linguistics be used as a starting point for a multilingual dictionary project?
- Which ways are there to a comparative historic lexicography?

Registration

Conference fee: 90 €. It contains the costs for conference documents, refreshments during the coffee breaks, city tour Jena, visit of old prints of the library, 3 lunch in the cafeteria of the university, dinner on Friday and office expenses. Please remit the conference fee until 30.06.2012.

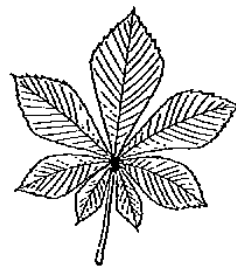
Læs mere på <http://www.saw-leipzig.de/news/ichll>.

7. International Society for Dialectology and Geolinguistics-kongres (SIDG)

Kongressen finder sted 23.-28. juli 2012 i Wien arrangeret af Österreichische Akademie der Wissenschaften. Der vil være flere leksikografiske indslag på kongressen:

- leksikografisk workshop: *Dialektlexikographie 2.0 – Status quo, Herausforderungen und Chancen*
- leksikografiske foredrag
- fejring af *Wörterbuchs der bairischen Mundarten in Österreichs* 100-års-jubilæum

Læs mere på sidg.oew.ac.at.



Kære kollega

Tiden for MUDS-konferencen nærmer sig med hastige skridt, og vi vil derfor minde dig om

14. Møde om Udforskningen af Dansk Sprog

torsdag 11. oktober – fredag 12. oktober 2012

på Aarhus Universitet

med Institut for Æstetik og Kommunikation, Nordisk som vært

Da mødets formål først og fremmest er gensidig inspiration og orientering, modtager vi gerne foredrag inden for **alle** discipliner af dansk sprogforskning. Også beskrivelser af igangværende projekter er meget velkomne, ligesom det vil være muligt at afholde workshops og opstille posters. Der afsættes 20 minutter pr. deltager, fx til et foredrag, plus 10 minutter til diskussion.

I forlængelse af konferencen udsendes en konferencerapport med fagfællebedømte bidrag.

Der vil blive serveret kaffe/te og frokost torsdag og fredag samt festmiddag torsdag aften. Gebyret for konferencen er **850,- DKK**.

Fristen for tilmelding af foredrag, workshop mv. er fredag **21. september** på MUDS' hjemmeside, <http://www.muds.dk>, hvor du desuden løbende kan finde information om konferencens afholdelse. **Betaling af konferencegebyret** foregår kun via Aarhus Universitets webshop; et link til den fremgår af konferencens hjemmeside.

Vi glæder os til at se dig til fagligt og socialt samvær i oktober!

Inger Schoonderbeek Hansen
jysis@hum.au.dk

Tina Thode Hougaard
northth@hum.au.dk

Peter Widell
norwidell@hum.au.dk

15th EURALEX International Congress 7-11 August 2012, University of Oslo

The program offers plenary lectures, parallel sessions on the topics listed below, software demonstrations, pre-congress tutorials and specialized workshops. The program also includes a special session for students and work in progress, a book and software exhibition, and social events for participants and their guests.

Registration

Registration is now open on <https://epay.uio.no/pay/shop.ordercreate.action?project=103290>. Early registration will be open until June 15th 2012.

Academic program

As things look now, there will be approximately one hundred presentations, oral speeches, demonstrations and posters. All the announced fields will be covered:

- Lexicography and national Identity
- Indigenous Languages and Lexicography
- Corpus-driven Lexicography
- Lexicography in Language Technology
- Multilingual Lexicography
- Lexicography and semantic Theory
- Terminology, LSP and Lexicography
- Reports on Lexicographical and Lexicological Projects

It seems that the topics of multiword expressions like proverbs, collocations and idioms are the most popular, as well as reports from lexicographical and lexicological projects, but several have responded to the main topic of the congress, Lexicography and national identity, which will have a special session at the conference. These are topics of current interest to our contemporaries, especially in Scandinavia with two national dictionaries being completed in the near future. Also the related topic Lexicography and language minority is represented by several presentations. All together the presentations will give a good overview of what is going on in European lexicography for the time being.

Plenary speakers

The plenary speakers will be **Ole Henrik Magga, Norway**, Professor of Sami languages, Sámi allaskuvla/The Sámi University College, Guovdageaidnu/Kautokeino; **Arnfinn Muruvik Vonen, Norway**, Director General at The Language Council of Norway, Professor of General Linguistics; **Bolette Sandford Pedersen, Denmark**, Professor in Language Technology, The University of Copenhagen; **Prof. Gilles-Maurice de Schryver, Belgium/South Africa**, Dept of Languages and Cultures, Ghent University, Belgium, Professor of African Linguistics & Xhosa Dept, University of the Western Cape, Cape Town (South Africa), Extraordinary Professor, and **Michael Rundell, United Kingdom**, Editor-in-Chief, Macmillan Dictionaries (www.macmillandictionary.com), Director, Lexicography MasterClass (www.lexmasterclass.com).

www.hf.uio.no/iln/forskning/aktuelt/arrangementer/konferanser-seminarer/2012/euralex