

LEDA-Nyt nr. 42 – Nov. 2006

ISSN 0908-355

Redigeret af Henrik Lorentzen, Sanni Nimb og Liisa Theilgaard

Indhold

Meddelelser fra bestyrelsen	s. 2
Medlemsmøde: ZooLex: Det er for vildt!	s. 3
DanNet – et leksikalsk-semantisk ordnet for dansk	s. 3
Anmeldelse: Svenska Akademiens ordlista, 13. udgave	s. 12
Referat fra 8. Nordiske Dialektologkonference	s. 17
EURALEX XII i Torino – Referat og anmeldelse	s. 20
Rapport fra LEDA-udflugt til Svenska Akademiens ordbok	s. 28
Ny dansk disputats	s. 31
Festskrift til Pia Jarvad	s. 33
Årsberetning for LEDA for året 2005-2006	s. 34
Referat af generalforsamling	s. 36
Annoncering af 9. Konference om Leksikografi i Norden	s. 38

Medlemsmøde ZooLex: Det er for vildt!

Ved Henrik Køhler Simonsen

Onsdag 22. november 2006 kl. 15.00

**CBS / Handelshøjskolen i København, Dalgas Have 15, 2000 Frederiksberg,
lokale 2V071**

se side 3

Meddelelser fra bestyrelsen

Portoudgifterne til dette nummer af LEDA-Nyt er finansieret af Lexitext. I den forbindelse følger et indstik om Lexitexts nye udgivelse på forlaget Alinea: SKOLEORDBOG TYSK. Ordbogen er designet og skrevet af Carsten Hansen og Katrine Wohllebe.

Kommende medlemsmøder

Ud over det medlemsmøde der annonceres i dette nummer af LEDA-Nyt, har bestyrelsen planlagt et møde hvor Kjeld Kristensen vil præsentere den store svensk-danske ordbog der er under udarbejdelse i Det Danske Sprog- og Litteraturselskab. Mødet finder sted onsdag den 7. februar 2007. Nærmere detaljer om tid og sted vil blive sendt ud til medlemmerne og desuden blive annonceret på LEDA's hjemmeside.

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.cst.dk/leda).

Bidrag til næste nyhedsbrev sendes til formanden inden 1. oktober 2006.

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:

Henrik Lorentzen (DSL) (formand), hl@dsl.dk, tlf.: 33 13 84 20-9637

Henrik Hovmark (KUA, Ømålsordbogen) (næstformand), hovmark@hum.ku.dk

Christian Becker-Christensen (Politikens Forlag) (kasserer), cbc@pol.dk

Inger Schoonderbeek Hansen (AU, Jysk Ordbog), jysis@hum.au.dk

Jørgen Nørby Jensen (Dansk Sprognævn), jn.jensen@dsn.dk

Sanni Nimb (KUA, CST), sanni@cst.dk

Liisa Theilgaard (Gyldendal), Liisa_Theilgaard@gyldendal.dk

Kontingent for medlemskab af Foreningen af Leksikografer i Danmark (LEDA) 125,- DKK (kan indbetales på girokonto 1551 139-1968).

Kontingent for medlemskab af Nordisk Forening for Leksikografi (NFL) 180,- DKK/200,- NKR.

Henvendelse vedrørende indmeldelse og adresseændring sker til foreningens kasserer.

Øvrige henvendelser rettes til formanden.

Medlemsmøde

ZooLex: Det er for vildt!

af Henrik Køhler Simonsen

ZooLex: Det er for vildt!

ZooLex er et virksomhedsleksikografisk værktøj, som er specielt udviklet til virksomheden København Zoo og dens medarbejdere. Det overordnede formål med ZooLex er at understøtte København Zoos kommunikationsindsats, hvilket gøres ved at integrere og konvergere en række leksikografiske og ikke-leksikografiske data.

De overordnede leksikografiske funktioner, som ZooLex er udviklet til at understøtte, er de kommunikationsorienterede funktioner L2-produktion og L1-L2-oversættelse samt den vidensorienterede funktion faglig tilegnelse. ZooLex blev med andre ord udviklet til at hjælpe medarbejderne i København Zoo til at skrive tekster på engelsk, at oversætte danske tekster til engelsk og at tilegne sig viden om zoologiske emner.

Foredraget indeholder en diskussion af de metaleksikografiske og systemarkitektoniske overvejelser, som jeg har gjort mig i forbindelse med udvikling og implementering af ZooLex, og diskussionen vil blive ledsaget af en præsentation af de forskellige funktioner i ZooLex.

DanNet – fra ordbog til et leksikalsk-semantic Word-Net for dansk

Af Bolette Sandford Pedersen og Jørg Asmussen

1. WordNets – leksikalsk-semantic ordnet for alverdens sprog

Tilstedeværelsen af leksikalsk-semantic ordnet, såkaldte WordNets (Fellbaum 1998, Vossen 1999), er en af de forudsætninger der skal til for at sætte en computer i stand til indholdsbaseeret håndtering af sprogligt materiale. Fordelene ved fx at kunne finde tekster med et bestemt indhold snarere end tekster der indeholder en bestemt ordform er åbenlyse, og det samfundsmæssige potentiale for applikationer der kan udnytte dette til forskellige former for informationsnavigering, er uvurderligt, ikke mindst når det gælder brugen af dem på internettet. Derfor undrer det næppe hvor omsiggribende arbejdet med WordNets er på verdensplan. Af Global WordNet Associations hjemmeside www.globalwordnet.org fremgår det at der for tiden enten findes eller udarbejdes WordNets for omkring 40 sprog. Dansk er et af dem, idet det daværende Statens Humanistiske Forskningsråd (nu Forskningsrådet for Kultur og Kommunikation) i 2005 bevilgede midler til etableringen af et WordNet for dansk, DanNet, der nu er på vej.

Det danske WordNet-projekt er et samarbejdsprojekt mellem Center for Sprogteknologi, Københavns Universitet (CST) og Det Danske Sprog- og Litteraturselskab (DSL). Hvert af disse to miljøer har for nylig afsluttet et omfattende leksikalsk projekt, hvis resultater nu tilsammen udgør et væsentligt udgangspunkt for udviklingen af det danske WordNet: Den Danske Ordbog (DDO: Hjorth et al. 2005) udgør en omfattende resurse der på baggrund af korpusundersøgelser beskriver ords betydninger, primært ved hjælp af definitioner og brugseksempler (Lorentzen 2004). CST har deltaget i det internationale projekt SIMPLE (Semantic Information for Multifunctional, Plurilingual Lexica, jf. Lenci et al. 2000), hvor ontologiske modeller for formel betydningsbeskri-

velse er blevet afprøvet og konsolideret på et bredt ordmateriale; en afprøvning på 10.000 begreber i dansk er udført i den danske del af projektet, SIMPLE-DK (Pedersen og Paggio 2004). Erfaringerne fra disse to projekter i henholdsvis det leksikografiske og det sprogteknologiske miljø har givet optimale muligheder for at udarbejde et WordNet for dansk.

2. Hvad er et WordNet?

Et WordNet er en samling af begreber eller “betydninger” (fx ‘apparat der modtager tv-signaler og omsætter dem til bevægelige billeder på en skærm’) med tilknyttede leksikaliserede udtryk (fx *fjernsyn*, *tv*, *tossekasse*). En sådan helhed bestående af et begreb og dets leksikaliserede udtryk betegnes et synonym sæt eller kort *synset*. Disse synsets udgør ordnettets byggestene. De enkelte synsets forbindes med hinanden ved at definere de semantiske relationer der hersker mellem dem, fx over- og underbegreber (fx ‘kop’ – ‘tekop’), del-helhed (fx ‘hank’ – ‘kop’) og funktionsrelationer (fx ‘beholder’ – ‘indeholde’). Gennem disse relationer etableres således en eksplicit leksikalsk-semantic beskrivelse i form af et semantisk net.

3. Den Danske Ordbog – hovedkilden til DanNet

Den Danske Ordbog udgør i praksis hovedkilden til DanNet. Da DDO har en nøje beskrevet og XML-opmærket mikrostruktur og blev redigeret i semantiske og – for funktionsordenes vedkommende – i ordklassebestemte “felter”, kan man gå ud fra at den er forholdsvis konsistent i sin opbygning og derfor lader sig udnytte datamatisk. Tanken om ligefrem at kunne omsætte ordbogen til et WordNet ved tryk på en knap er ganske besnærende, og nogle overvejelser om hvor langt man muligvis kan komme i den retning, gives i afsnit 6. I figur 1 vises DDO-artikelstrukturen af sememet med betydningen ‘apparat’ fra artiklen *fjernsyn*.

```
<⊞ Semem >
  <⊞ Restspec >
    <⊞ Sysfag > fje
    <⊞ Denbet > kasseformet apparat der kan modtage tv-signaler og
    omsætte dem til bevægelige billeder på en skærm og tilhørende lyd i
    apparatets højtalere
    <⊞ Genprox > apparat
    <⊞ Onym >
      <⊞ Syn >
        <⊞ txt > tv
        <⊞ Syn >
          <⊞ KomOnym > mere formelt
          <⊞ txt > fjernsynsapparat
    <⊞ Rel >
      <⊞ Typsam >
        <⊞ txt > slukke for fjernsynet
      <⊞ Typsam >
        <⊞ txt > tænde (for) fjernsynet
    <⊞ Dok >
      <⊞ Citat >
        <⊞ txt > Efter at det først var "fint" at have fjernsyn, blev det fløvt:
        Apparatet skulle skjules
      <⊞ Kilde >
        <⊞ DDOkilde > BT91
        <⊞ Kildeid > CPTn
```

Figur 1: DDO-mikrostruktur for et semem

Som det fremgår, er den semantiske beskrivelse opslittet i en række funktionelle elementer. Således er der en (i dette eksempel desværre upræcis) angivelse af domænet (i elementet Sysfag): fje, som egentlig er domæneangivelsen for 'fjernsyn som institution', mens der i denne semantiske beskrivelse er tale om det tekniske apparat. Betydningsdefinitionen står i elementet Denbet ("denotativ betydning") som har et underelement Genprox der angiver et leksikaliseret udtryk for det genus proximum som danner udgangspunktet for definitionen. Der er endvidere to synonymangivelser i Syn-elementet, den ene med indledende stilistisk kommentar i et tilknyttet KomOnym-element. Dette efterfølges af eksempler på statistisk påfaldende kookkurrenser i Typsam-elementerne. Og endelig gives der i en dokumentationsdel (Dok) et korpusbelæg (Citat) med tilhørende kildeoplysninger.

Da et WordNets grundlæggende enheder som beskrevet er såkaldte synsets, altså grupper af leksikaliserede udtryk (typisk lemmaer) med tilknyttet fælles betydning, er det nærliggende at bruge synonymoplysningerne fra DDO-artiklerne umiddelbart til en første gruppering af leksikaliserede udtryk i synsets. Tilsvarende kan artiklernes antonymoplysninger anvendes direkte til at etablere tilsvarende semantiske relationer mellem synsets, mens domæneoplysningen fx giver en indikation af den ontologiske type. Herudover er det dog begrænset hvad de forskellige funktionelle elementer i artiklens mikrostruktur, sådan som de kommer til udtryk via XML-opmærkningen, faktisk kan tilføre DanNet af semantiske oplysninger om pågældende semem, idet disse oplysninger primært er "kodet" i selve betydningsdefinitionen. Det er derfor DanNet-redaktørernes opgave at omsætte informationen indeholdt i definitionerne til oplysninger der er nødvendige i et WordNet. Først og fremmest skal DDO's hyponymstruktur omsættes til en der er gangbar i et WordNet.

4. SIMPLE-ordbasen – udgangspunktet for en ontologibaseret ordbeskrivelse

Mens DDO altså udgør den væsentligste basisresurser for DanNet, så udgør SIMPLE det eksperimentelle udgangspunkt for en ontologibaseret beskrivelse af det danske ordforråd. I SIMPLE-ordbasen har vi eksperimenteret med en på forhånd givet ontologisk struktur med et fast vokabular og med en kortlægning af hvilke semantiske relationer der kan og bør angives i en formel leksikografisk tilgang. Et andet element fra SIMPLE som vi har overført til DanNet, er de såkaldte templates eller skabeloner. En skabelondrevet metode går i sin enkelthed ud på at der for hver ontologiske type angives en ramme for hvilke elementer – fx hvilke relationer – synsettet bør indeholde. For begreber af den ontologiske type *Artifact* (kunstigt frembragt genstand) er det fx relevant at angive hvad en ting typisk bruges til, mens det for den ontologiske type *Part* er vigtigt at angive hvad noget er en del af. Ved at opstille skabeloner for typiske egenskaber for de enkelte ontologiske typer oparbejdes en højere grad af konsistens imellem de forskellige synsets. Skabelonerne fungerer således som en slags redaktionsregler for udviklingen af DanNet og vil også kunne bruges som referencepunkt for den senere anvendelse og eventuelle udvidelse af nettet.

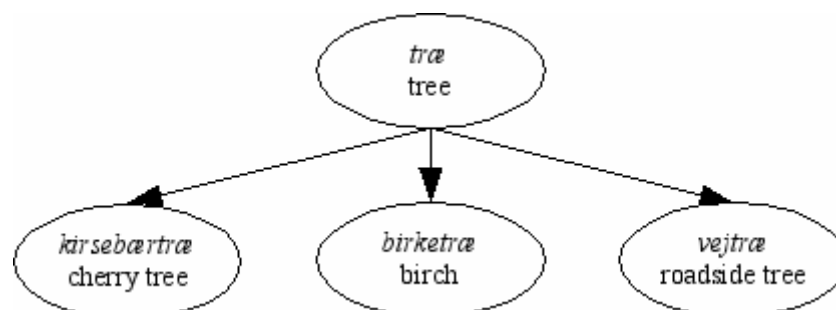
5. Hyponymirelationerne – ordnettets skelet

Hyponymirelationerne, dvs. angivelsen af over- hhv. underbegreber, udgør WordNettets skelet, og i projektets første fase står udviklingen af disse relationer i fokus. Disse relationer udtrækkes semi-automatisk på baggrund af DDO's definitioner som overvejende er givet ud fra den klassiske *genus-differentia*-model, hvor man først angiver det nærmeste overbegreb og dernæst de adskillende træk. Således gælder der for eksemplet *fjernsyn*, at 'apparat' er genus proximum og (nogle af) de øvrige oplysninger i definitionen differentia-oplysninger (en mere detaljeret beskrivelse følger i afsnit 6). Oplysningen om nærmeste overbegreb, som er angivet i det særlige *Genprox*-element

som ikke vises i den trykte ordbog, udnyttes med andre ord til automatisk at foreslå over- og underbegreber for DanNet-redaktørerne.

Selv om en sådan opbygning letter arbejdet med at placere begrebet i en taksonomisk struktur, resterer der meget efterarbejde i form af dels at håndtere flertydighed (da nærmeste overbegreb angives som et leksikaliseret udtryk i DDO – ikke med reference til en specifik betydning af dette udtryk), dels at balancere WordNettet således at den taksonomi der opbygges, bliver så konsistent og anvendelig som mulig til datamatiske formål. Udfordringen ligger bl.a. i at de anvendte overbegreber i DDO's betydningsdefinitioner ikke er taget fra et i forvejen etableret ontologisk system, men er valgt af den enkelte redaktør i overensstemmelse med nogle generelle anvisninger i redaktionsreglerne for ordbogen. Det gør fx at én redaktør har valgt *lære* for at udtrykke overbegrebet for *informatik* og *bromatologi*, mens en anden har valgt *fag* for at udtrykke overbegrebet for *samfundsfag* og en tredje overbegrebsudtrykket *videnskab* for *datalogi*. I disse tilfælde er det DanNet-udviklernes opgave at forsøge at harmonisere nettet, og dette sker ofte ved at samle udtryk som fx *fag*, *lære* og *videnskab* i samme synset i de tilfælde hvor der ikke synes at være konsistente principper for at bibeholde en opdeling. Ligeledes synes det at bero på et tilfældigt valg af synonymier når *gadespejl* anses for at være en indretning, mens fx *støttefod* er en anordning. Eller *spisekort* beskrives som en oversigt, mens *menukort* anses for at være et kort. Hvis WordNettet fx skal anvendes til at gøre informationshåndtering mere fleksibel, er en sådan harmonisering vigtig. Derimod kan det være problematisk at indarbejde andre og mere finkornede hierarkier end dem der er antydnet i DDO; generelt stræbes ved udarbejdelsen af DanNet blot efter at eksplicitere allerede angivet information.

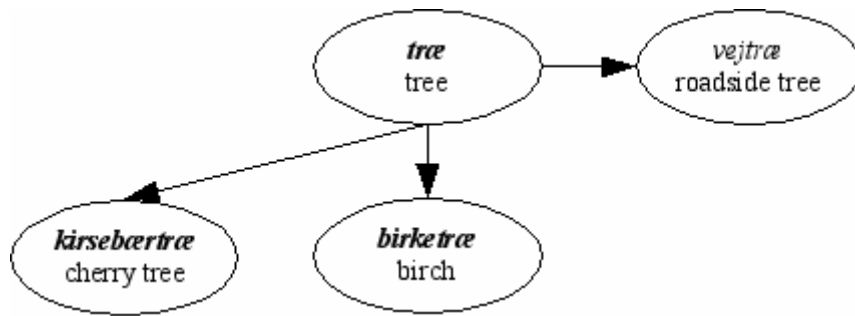
Håndteringen af den ontologiske struktur som det leksikalsk-semantiske WordNet uundgåeligt vil få, betragtes således som en særlig forskningsmæssig udfordring. Særligt hyponymrelationen har en central ontologisk status, og når man udtrækker store grupper af hyponymer fra Den Danske Ordbog, begynder der at danne sig nogle interessante mønstre. Det viser sig tydeligt at hyponymrelationen dækker over forskellige undertyper af relationer hvoraf kun nogle har en egentlig taksonomisk status. Hvor hyponymi kan defineres ud fra sætningen: *X er en Y*, så kan taksonomi forstås som en undertype til hyponymi med definitionen: *X er en slags/type Y* (Cruse 1991 & 2002). Taksonomier er vellidte i ontologier bl.a. fordi nedarvningsstrukturen er nogenlunde klar; anden slags hyponymi er derimod mere kompliceret at håndtere fra et ontologisk synspunkt. I figur 2 ses tre hyponymer til 'træ' (i 'plante'-betydningen) udtrukket fra DDO:



Figur 2: Hyponymer til 'træ'

Da *vejtræ* adskiller sig fra de to øvrige hyponymer ved ikke at være 'en slags træ', men derimod 'en hvilken som helst slags træ der står i vejkanten', bør den få en anden status i hierarkiet. Den opfattes i DanNet som "ortogonal" til den øvrige taksonomi (jf. figur 3), altså som et hyponym der angiver en anden betydningsmæssig dimension end

de øvrige to hyponymer. For en dybere redegørelse af opdelingen af hyponymirelationerne på basis af Cruses klassifikation henvises til Pedersen og Sørensen 2006.



Figur 3: Den ortogonale hyponymirelation

Hvor taksonomi altså defineres som: *X er en slags/type Y*, kan ikke-taksonomiske hyponymer defineres som: *X er en hvilken som helst slags Y, for hvilket det gælder at Z*. Både birketræer og kirsebærtræer kan altså i princippet fungere som vejtræer. I figur 4 ses hvordan denne skelnen rent praktisk udtrykkes i DanNet-basen.

The screenshot shows the DanNet interface with two main panels highlighted in red:

- taxonomic hyperonymy:**
 - {ahorn_1} (Plant+Object): løvtræ med store ferner
 - {ambratræ_1} (Plant+Object): løvfældende træ
 - {bøg_1} (Plant+Object): løvtræ med slank, lige stamme
 - {eg_1} (Plant+Object): op til 30 m højt løvtræ med
 - {hickory_1} (Plant+Object): højt, løvfældende træ
 - {kamfertræ_1} (Plant+Object): stedsegrønt løvtræ
- orthogonal hyperonymy:**
 - {beplantning_1_1} (Material+Natural+Substance): træ
 - {gavntræ_1} (Plant+Object): træ der er af så god kvalitet
 - {hænge træ_1} (Plant+Object): træ med nedhængende
 - {klatre træ_1} (Plant+Object): træ der er godt at klatre i, f
 - {livstræ_1} (Plant+Object): træ som menes at give me
 - {skovtræ_1} (Plant+Object): træ el. træart der oftest vok
 - {vejtræ_1} (Plant+Object): træ der står i vejkanter, plan

Below these panels, there are fields for 'constitutive', 'formal', 'synonymy', and 'telic' relationships, each with a dropdown menu and a 'Fill' button.

Figur 4: Ortogonal og taksonomisk hyponymi i DanNet-interfacet

6. Udtræk af andre semantiske oplysninger fra ordbogsdefinitioner

Hyponymirelationerne er forholdsvis nemme at udtrække semi-automatisk fra DDO-definitionerne da der til hver definition er knyttet et overbegrebsudtryk (*GP-udtryk* i det

følgende) i det særlige *Genprox*-element. Herudover indeholder definitionerne dog også differentia-oplysninger som det ville være fordelagtigt også at kunne udtrække på en forholdsvis automatiseret måde for sidenhen at indarbejde dem i DanNet. I det følgende skal det derfor skitseres hvilke problemer man støder ind i hvis man ønsker også at kunne omsætte andre dele af DDO's betydningsdefinitioner til semantiske oplysninger i DanNet. En detaljeret redegørelse for problemerne og en beskrivelse af deres leksikografiske implikationer kommer i Asmussen 2007.

Princippet der ligger til grund for betydningsdefinitionerne i DDO, er som nævnt i afsnit 5 at anføre *genus proximum* og *differentia specifica*. For valg af GP-udtrykket kunne redaktørerne støtte sig til nogle anvisninger i redaktionsreglerne, mens der for differentia-oplysningerne var meget vide rammer for hvad der var muligt at skrive i definitionen: Der var således ingen fastlagt semantisk beskrivelsesmodel der fx foreskrev hvilke træk eller relationer der skulle beskrives for hvilke typer ord. Endvidere var hverken vokabularet eller syntaksen for definitionerne lagt fast på forhånd. Alt sammen forhold der normalt ikke er afgørende så længe en ordbogs målgruppe er mennesker, men forhold der så også gør det særdeles komplekst – om ikke umuligt – sidenhen at udnytte definitionerne maskinelt.

En første tilnærmelse til en mulig maskinel udnyttelse af materialet kan ske ved at betragte samtlige definitioner i DDO som et *korpus* hvorpå der så kan udføres gængse korpusanalytiske undersøgelser der kan hjælpe til med at give et indblik i definitionernes opbygning. Dette kan illustreres på baggrund af den definition der er gengivet i figur 1. Her er GP-udtrykket *apparat* mens det attributive adjektiv til GP-udtrykket *kasseformet* og *kan* med infinitivkonstruktionerne *modtage tv-signaler* samt *omsætte dem [...]* er differentia-oplysninger. På baggrund af disse observationer kan der opstilles følgende struktur-hypoteser for ordbogsdefinitioner for artefakter:

1. De(t) adjektiv(er), der står foran GP-udtrykket, beskriver definiendums generelle egenskaber, her dets udseende eller fremtoning.
2. Infinitivkonstruktioner efter *som/der kan* beskriver definiendums teliske rolle (formål eller funktion).

Hypoteserne kan omformuleres til nogle forespørgsler til definitionskorpusset. Beder man således om samtlige definitioner med GP-udtrykket *apparat* som har præcist ét ord foran sig (et formodet attributivt adjektiv), fås følgende eksempler: *elektronisk*, *kasseformet*, *lille*, *mekanisk*, *mindre*, *optisk* – ikke mange i betragtning af at der i DDO er 209 definitioner med GP-udtrykket *apparat*. Så den struktur som hypotese 1 postulerer, anvendes åbenbart kun lidt i DDO selvom der principielt ikke er noget der taler imod en mere konsekvent brug af den.

Mere interessant er resultatet af hypotese 2 omsat til en korpus-forespørgsel: Figur 5 viser et udsnit af den resulterende konkordans sammen med lemmaet der knytter sig til hver enkelt definition.

line	left context	match	right context	Lemma
000026		Apparat der kan opvarme	vand fx til brug i hush...	vandvarmer
000025		Apparat der kan sprøjte	vand ud med stor kraft	vandkanon
000024	Kasseformet	apparat der kan modtage	tv-signaler og omsætt...	tv
000023	Elektronisk	apparat der kan efterligne	forskellige tromme- og ...	trommemaskine
000022		Apparat der kan sende	el overføre signaler til ...	transmitter
000021	Elektronisk	apparat der kan måle	tiendedele el hundred...	timer
000020		Apparat der kan besvare	et telefonopkald med e...	telefonsvarer
000019	Elektrisk	apparat der kan frembringe	en summetone fx som ...	summer
000018		Apparat der kan måle	rotationshastighed	stroboskop
000017	...l elektrisk	apparat som kan skrive	bogstaver og andre te...	skrivemaskine
000016	...ammerbart	apparat som kan optage	og afspille digitalt lagr...	sequencer
000015		Apparat der kan frembringe	og udsende radiosign...	sender
000014		Apparat der kan modtage	radiobølger og gengiv...	radio
000013		Apparat der kan bringe	en væske el luftart i b...	pumpe
000012	Optisk	apparat som kan fremvise	et forstørret billede af t...	overheadprojektor
000011		Apparat der kan vise	elektriske svingninger	oscilloskop
000010		Apparat som kan omdanne	energi el elektroniske ...	omformer
000009		Apparat som kan opfange	og gengive radio- el t...	modtager

Figur 5: Telisk rolle udtrykt i DDO-definitioner

Ved hjælp af det strukturelle mønster som er udtrykt i hypotese 2, lader det til at være muligt at isolere oplysninger om telisk rolle i DDO-definitionerne. Imidlertid fanger man desværre ikke alle angivelser af telisk rolle i definitionerne på denne måde, idet angivelsen kan være indlejret i en række forskellige strukturelle mønstre som de følgende eksempler viser:

- GP som bruges til at VP-inf med: *apparat som bruges til at spinde garn med*
- GP til at VP-inf med: *apparat til at afspille cd'er med*
- GP der kan VP-inf: *apparat der kan modtage tv-signaler ...*
- GP der VP-fin: *apparat der måler og viser et køretøjs hastighed*
- GP til NP: *apparat til optagelse og afspilning af lyd*
- GP som er specielt beregnet til at VP-inf: *apparat som er specielt beregnet til at optage og afspille tale*

Selvom man kan argumentere for at der i definitionerne kommer semantiske nuancer til udtryk, og at den store udtryksvariation dermed måske kan være berettiget, kan alle disse strukturer i princippet godt vikariere for hinanden uden de helt store konsekvenser for definitionernes korrekthed. Det viser sig altså at et bestemt semantisk træk – her telisk rolle – kan udtrykkes på mange forskellige måder i definitionerne hvilket gør et automatisk udtræk af dem vanskelig, ikke mindst fordi de selv kan komme til udtryk ved forskellige kategorielle udfyldninger: VP-inf, VP-fin, NP. Et præcist udtræk nødvendiggør derfor i princippet en syntaktisk opmærkning af definitionskorpusset.

En anden, betydelig mere grovkornet måde at nærme sig DS-udtryk fra ordbogsdefinitionerne der måske oplyser om telisk rolle, er at anvende en ren statistisk fremgangsmåde hvor man bestemmer hyppigheden af forekomster af ordformer i definitioner med fx GP-udtrykket *apparat* i forhold til deres hyppighed i definitions-korpusset som helhed og registrerer overhyppige ordformer ved hjælp af en statistisk test som fx *log likelihood* (jf. Dunning 1994). Ved at bruge en sådan fremgangsmåde

(i det følgende eksempel faktisk en særanvendelse af *mutual information*, jf. nærmere beskrivelse i Asmussen 2007) fås for apparaternes vedkommende, blandt mange andre, følgende DS-ordformer, der muligvis indikerer telisk rolle sammen med de lemmaer der udtrykker pågældende betydning:

- *afspille*: *grammofon, cd-afspiller, afspiller, sequencer, diktafon*
- *afspilning*: *kassettespiller, hjemmevideo, kassettebåndoptager, båndoptager*
- *måle*: *stroboskop, måler, timer, løgnedetektor, ekkolod*
- *måler*: *gasmåler, speedometer, omdrejningstæller, benzinmåler, fotofælde*
- *måling*: *elmåler, trykmåler, luxmeter, spirometer, gyrometer, alkometer, newtonmeter, magnetometer, instrument, kalorimeter*
- *målinger*: *måleinstrument, radiosonde, satellit, fartskriver*

Med udgangspunkt i sådanne automatisk opstillede lister bestående af overhyppigt forekommende ordformer i definitioner med GP-udtrykket *apparat* med tilknyttede lemmaer kan redaktøren sættes på sporet af hvilke synsets der skal udstyres med angivelse af samme teliske rolle.

Eksemplerne viser at man ganske vist kan bruge bestemte metoder fra korpuslingvistikken til at få et første indblik i ordbogsdefinitioners opbygning, men selve fortolkningen af hvilke elementer i definitionen der udtrykker hvilke semantiske træk, må overlades til redaktøren. En fuldautomatisk omsætning af ordbogsdefinitioner til et WordNet er derfor næppe mulig, dog kan visse korpusanalytiske metoder være et brugbart værktøj der kan gøre det lettere for DanNet-redaktørerne at navigere i DDO's definitioner – dette kan både gavne definitionerne idet det giver en vis indsigt i hvordan de er skruet sammen (og hvordan de eventuelt kan harmoniseres), og det skulle først og fremmest gavne DanNets konsistens.

Følgende korpusanalytiske metoder kan formentlig med fordel anvendes på et korpus over ordbogsdefinitioner:

- Ved at opstille frekvensprofiler over GP-udtryk kan lavfrekvente GP-udtryk opspores og erstattes af mere generelle. Ligeledes kan der opstilles frekvensprofiler over definitionsvokabularet med henblik på at reducere eller forenkle det.
- Frekvensundersøgelser kan også udføres på ordgrupper. Ved at sammenligne med tilsvarende resultater fra et almensprogligt referencekorpus kan signifikante syntagmatiske mønstre i ordbogsdefinitionerne afsløres.
- Ved at bruge forskellige statistiske lighedsmål kan definitioner der strukturelt ligner hinanden, fremfindes. Dette kan gøre det lettere for redaktørerne at samle udtryk med lignende betydning i et synset.
- En række statistiske metoder kan bruges til afsløring af bestemte træk-udtryk i definitionerne.
- KWIC-konkordanser kan i øvrigt understøtte mange af de nævnte undersøgelser.

7. Konkluderende bemærkninger

DanNet forventes afsluttet i 2008 med et omfang af omkring 40.000 begreber. Vi forventer at ordbasen vil blive stillet mere eller mindre frit til rådighed til forskningsformål, mens der til kommerciel anvendelse skal betales hvad der svarer til prisen for de øvrige europæiske WordNets der forhandles igennem organisationen ELRA (European Language Resources Association), jf. www.elra.info.

Man kan følge arbejdet med DanNet på www.wordnet.dk.

Referencer

- Asmussen, J. (2007). Korpuslinguistische Verfahren zur Optimierung lexikalisch-semanticser Beschreibungen. I Kallmeyer, W., udgiver, *Jahrbuch des Instituts für Deutsche Sprache 2006*. de Gruyter, Berlin / New York.
- Cruse, D. A. (1991). *Lexical Semantics*. Cambridge University Press, Cambridge.
- Cruse, D. A. (2002). Hyponymy and Its Varieties. I Green, R., Bean, C., og Myaeng, S., udgivere, *The Semantics of Relationships: An Interdisciplinary Perspective, Information Science and Knowledge Management*. Springer-Verlag.
- DDO: Hjorth, E., Kristensen, K., Lorentzen, H., Trap-Jensen, L., Asmussen, J., et al., udgivere (2005). *Den Danske Ordbog 1-6*. DSL og Gyldendal, København/Copenhagen.
- Dunning, T. (1994). Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics*, 19(1):61-74.
- Fellbaum, C., udgiver (1998). *WordNet: An Electronic Lexical Database*. The MIT Press, Cambridge, MA.
- Lenci, A., Bel, N., Busa, F., Calzolari, N., Gola, E., Monachini, M., Ogonowski, A., Peters, I., Peters, W., Ruimy, N., Villages, M. og Zampolli, A. (2000). SIMPLE – A General Framework for the Development of Multilingual Lexicons. *International Journal of Lexicography*, 13:249–263.
- Lorentzen, H. (2004). The Danish Dictionary at large: presentation, problems and perspectives. I *Proceedings of the 11th EURALEX International Congress*, volume 1, pages 285–294, Lorient. Euralex.
- Pedersen, B. S. og Paggio, P. (2004). The Danish SIMPLE Lexicon and its Application in Content-based Querying. *Nordic Journal of Linguistics*, 27(1):97-127.
- Pedersen, B. S. og Sørensen, N. H. (2006). Towards sounder taxonomies in wordnets. I *Proceedings from the OntoLex Workshop in association with LREC 2006*, 9-16, Genova.
- Vossen, P., udgiver (1999). *EuroWordNet. A Multilingual Database with Lexical Semantic Networks*. Kluwer Academic Publishers.

Anmeldelse: Svenska Akademiens ordlista, 13. udgave

Af Kjeld Kristensen

I Svenska Akademiens vedtægter fra 1786 står der at det nye akademis vigtigste opgave er “att arbeta uppå Svenska Språkets renhet, styrka och höghet”. Derfor skulle akademiet bl.a. udarbejde “en Svensk ordabok och Gramatica”. Der skulle gå over 200 år før grammatikken forelå i form af det imponerende firebindsværk *Svenska Akademiens grammatik* (1999). En mindre og mere populær grammatik med både beskrivende og vejledende sigte udkom i 2003: Tor G. Hultmans *Svenska Akademiens språklära*. Altså to grammatikker – men hele tre ordbøger! For det første *Svenska Akademiens ordbok* (SAOB), én af verdens største historiske modersmålsordbøger, hvis første hæfte udkom i 1893, og som ventes afsluttet i 2017 med i alt ca. 40 bind. For det andet *Svensk ordbok utgiven av Svenska Akademien*, en kommende ordbog over nutidssvensk der bygger på *Nationalencyklopedins ordbok* (1995-96) og før den *Svensk ordbok* (1986). Og for det tredje *Svenska Akademiens ordlista över svenska språket* (SAOL), der i april 2006 udkom i 13. udgave.

SAOL udkom første gang i 1874, 10. udgave i 1973, 11. udgave i 1986, 12. udgave i 1998 og 13. udgave altså i 2006. SAOL er således kommet i ny udgave i hvert af de seneste fire årtier, og det kan ifølge forordet i SAOL12 være en passende model for fremtidige udgaver også.

Mens vi i Danmark har haft officielle retskrivningsordbøger lige siden Svend Grundtvigs *Dansk Haandordbog* (1872), har SAOL aldrig opnået en tilsvarende status i Sverige. Den er “till sin natur en samling rekommendationer” (fra forordet til SAOL12), som dog tilsammen udgør en almindeligt accepteret norm for stavning og bøjning af svenske ord. Der er tale om en uofficiel norm, en situation der måske skyldes Svenska Akademiens historiske statsuafhængige autoritet. Reelt har SAOL samme position og funktion i Sverige som Retskrivningsordbogen (RO) har i Danmark, men den har tilsyneladende større udbredelse, med 300.000 solgte eksemplarer af SAOL12 og et førsteoplæg af SAOL13 på 80.000 eksemplarer.

Indhold

SAOL13 er en tyk bog i stort oktavformat (ca. 19 × 12 cm) med 1130 ordbogssider og 41 sider omtækt: indholdsfortegnelse, forord, indledning og kort redegørelse for ordbogens opbygning. Bogen har ikke noget godt opslag; den bliver ikke liggende efter opslag, i hvert fald ikke mit eksemplar. Et større format og et dermed følgende lavere sideantal kunne afhjælpe dette problem og i det hele taget gøre bogen mere handy.

Der er ikke som i RO nogen afdeling med retskrivningsregler. Dem finder man i Svenska språknämndens håndbog med “skrivregler”, seneste udgave i 2000. Indtil 1. juli 2006 var Språknämnden et halvstatsligt foretagende. Derfor har SAOL’s ortografiske norm og Språknämndens retskrivningsregler uofficiel eller halvofficiel status, en tilstand der set fra en dansk synsvinkel må forekomme besynderlig og uholdbar.

Som nævnt normerer SAOL (altså uofficielt) svenske ords stavning og bøjning, dog sådan at dobbeltformer tillades og endda prioriteres indbyrdes; det vender jeg tilbage til i afsnittet Normering og rekommendation. Kun simpleksord og sammensatte verber har oplysninger om bøjning. Bøjningsoplysninger for øvrige sammensætninger og for afledninger skal findes under ordets sidste del, der i sammensætninger er den del af ordet der følger efter en hel lodret streg i opslagsordet. Mere om denne markering i afsnittet Markering af morfemer – orddeling. Man kunne, ikke mindst som dansk bruger, ønske at alle opslagsord havde bøjningsoplysninger på stedet sådan som det har

været tilfældet i RO siden 2001. Det ser dog også ud til at være redaktionens mål, efter en passus i indledningen s. XXVII at dømme.

SAOL indeholder andre oplysningstyper, også flere end RO, og ganske naturligt får oplysningerne et autoritativt præg og en normerende effekt. Det gælder bl.a. oplysninger om sammensætningsfuge, der dog kan siges at høre under stavning.

I mange tilfælde gives der oplysning om udtale, dog ofte kun for den udtalemæssigt interessante del af ordet. Det drejer sig bl.a. om tryk, vokallængde og afvigelser fra de generelle regler i svensk for sammenhængen mellem skrift og lyd; *giraff* har fx udtaleangivelsen [ʃirafʔ] og *kör* (da. *kor*) [k-].

SAOL har langt oftere end RO ordforklaring til ordene. Hertil benyttes forskellige virkemidler: egentlige definitioner ved svære eller sjældne ord, identificerende betegnelser som fagmarkering eller ordfelt (fx “växt” eller “fågel”), og synonymer. I mange tilfælde benytter ordforklaringerne sig af forskel i skriftstørrelse til at angive eventuelitet, præciserende oplysning e.l., som når opslagsordet *manke* forklares med “ryggens främre del på t.ex. häst”. Jeg tror at mange ordbogsbrugere er glade for at få disse korte, men nyttige oplysninger i forbindelse med reception eller produktion af svensk tekst. Men oplysningerne om betydning får uvægerligt normerende effekt, hvad man må være klar over hvis man vil gå denne vej.

En oplysningstype som SAOL anvender i modsætning til RO, er brugelighed. Fagmarkeringer kan angive begrænsninger i brugelighed, ligesom etiketter som <åld.> (gammeldags), <vard.> (dagligsprog), <prov.>, <finl.> (finlandssvensk), <litt.>, <skämts.> (spøgende) og <kan uppfattas som nedsättande> (denne sidste etiket står fx ved ordet *neger*). Også disse oplysninger er gode at få, ikke mindst for indvandrere i Sverige og for udlændinge der interesserer sig for eller arbejder med svensk sprog.

Endelig indeholder mange artikler oplysning om udtryk som opslagsordet kun eller ofte forekommer i, undertiden endda med forklaring af udtrykket. Et eksempel er opslagsordet fra før, *manke*, der i ordbogen er eksemplificeret med det hyppigt forekommende udtryk *lägga manken till* ‘anstränga sig’.

Alt i alt fremtræder SAOL13 som en ordbog der groft sagt er en mellemtung mellem Retskrivningsordbogen og Nudansk Ordbog, og det er virkelig positivt ment, for den forener begge disse ordbøgers bedste egenskaber.

Ordforråd

SAOL13 indeholder over 125.000 opslagsord, ca. dobbelt så mange som i RO. Af en eller anden grund giver ordbogen ikke selv denne oplysning, men den fremgår af et par af de kilder der er anført til sidst i anmeldelsen, bl.a. Reuter (1999). En stor del af forklaringen på de mange flere ord i SAOL13 finder man ved at sammenligne antallet af sammensætninger i de to ordbøger. En stikprøve for tre ord i bogstav M viser følgende billede:

sms. med	SAOL12	SAOL13	RO
<i>miljö-</i>	42	85	12
<i>mjök-</i>	39	42	13
<i>muskel-</i>	15	26	8

SAOL indeholder langt flere sammensætninger end RO gør, og man kan spørge hvor

meget dette øger SAOL's brugsværdi. Vel ikke som afgørende hjælp ved stavning og bøjning af en bestemt sammensætning som står i SAOL13 (de enkelte led kunne jo ellers slås op i ordbogen og sættes sammen efter det i ordbogen angivne mønster), snarere ved afgørelsen af om ordet har en tilstrækkelig udbredelse i svensk til at virke upåfaldende i en tekst, hvis det nu er meningen.

Ordforrådet i SAOL13 er nemlig i høj grad bestemt af ordenes frekvens i de store tekstsamlinger (Språkbanken) af især avistekster fra 1965 til 2004 som ordbogsfolkene på Institutionen för nordiska språk ved Göteborgs universitet har til deres rådighed. Med en morfologisk database som hjælpemiddel har man samkørt ordmaterialet i SAOL12 med de store tekstmængder som man i stigende omfang har til sin disposition, og derved har man fået en kæmpemæssig liste med ord der findes i teksterne, men ikke i SAOL12, og omvendt en noget mindre liste med ord der findes i SAOL12, men ikke i teksterne. Den første ordliste har man reduceret ved at sætte en nedre frekvensgrænse for optagelse i SAOL13. Denne automatiske metode, der er beskrevet i indledningen til SAOL13, er forklaringen på at man uden det store besvær har kunnet optage 10.500 "nye" ord i SAOL13 og kassere 5.500 "gamle" ord. Opstillingen ovenfor giver et fingerpeg om hvor vækstområderne i SAOL's ordforråd ligger.

SAOL-redaktionens udstrakte brug af it i ordbogsarbejdet må give inspiration til udarbejdelsen af den næste danske retskrivningsordbog, men metoden rejser også spørgsmålet om en retskrivningsordbog kun skal tjene til produktion af nutidstekster med nutidsord om nutidsemner, for sådan bliver det let med den valgte måde at lave ordbog på.

Det skal retfærdigvis siges at redaktionen også har trukket på andre kilder end Språkbanken, bl.a. avisdatabaserne PressText og Mediarkivet, søgemaskinen Google samt forslag fra offentligheden via Svenska Akademiens hjemmeside. Der er bestemt også andet end mere eller mindre banale sammensætninger blandt de mange nye ord i SAOL13, fx ord der spejler samfundsudviklingen og verdenshistorien siden SAOL12 udkom i 1998: *blogg, burka, jihad, mingla, pixel, sms:a, sudoku* og *tsunami*.

Markering af morfemer – orddeling

Opslagsordene i SAOL13 står i alfabetisk orden, i øvrigt med W som nyt selvstændigt bogstav. Den alfabetiske orden er ikke nogen selvfølge, for først fra og med SAOL12 opgav man opstilling efter samhørighedsprincippet, der indebærer at ord der ord-dannelsesmæssigt hører sammen, altså har samme simpleksord som udgangspunkt, også står sammen i ordbogen i det der kaldes redealfabetisk makrostruktur. Ulempen ved det er at det kan være svært at finde ordene. Fordelen er at det semantiske og etymologiske fællesskab mellem ord i samme gruppe betones.

Delvis som konsekvens af opgivelsen af samhørighedsprincippet har man indført markering af de morfemer der opbygger ordene: hel lodret streg for stor grænse mellem morfemer, halv lodret streg for lille grænse. Princippet kan illustreres af eksemplet *miljö, kon, sekv, ens|be, skriv, ning*. Markeringen giver som omtalt tidligere mulighed for at finde ud af hvordan ordet bøjes; bøjningen finder man under det ord der står efter den hele lodrette streg.

Men morfemmarkeringen tjener nogle videre formål. Den kan være en nøgle til at forstå ordets betydning, ligesom i eksemplet, hvor placeringen af den hele lodrette streg viser at ordet betyder 'beskrivelse af miljøkonsekvenser' og ikke 'konsekvensbeskrivelse af/for miljøet'. En anden gevinst er efter SAOL-redaktionens opfattelse af pædagogisk art: Morfemmarkeringen demonstrerer hvordan svenske ord er dannet og opbygget. Endelig – og det er kontroversielt – anbefaler SAOL-redaktionen morfemopdelingen som udgangspunkt for "avstavning", orddeling. Eksempelvis kan ordet *konsekvens* altså deles to steder: *kon-sekvens* og *konsekv-ens*. Man har fra flere

sider, bl.a. Svenska språknämnden, kritiseret dette eneste princip for orddeling for at kræve for megen etymologisk viden af den skrivende, som fx skal kende den latinske baggrund for ordet *konsekvens* for at kunne dele ordet, medmindre han eller hun da ulejliger sig med at slå ordet op i SAOL. Og man har påpeget at princippet kan resultere i orddele med konsonantgrupper der er umulige i svenske stavelser, som fx *(kon)sekv-* i *konsekvens*. Endelig har man anført at flerstavede ord uden morfemgrænse, fx *badminton*, *mahogany* og *vatten*, slet ikke kan deles efter den anbefalede model.

Over for denne radikale model står den form for orddeling som Svenska språknämndens retskrivningsregler anbefaler. Den kombinerer morfemprincippet og det såkaldte enkonsonantprincip, der indebærer at en konsonant mellem to vokaler føres til den sidste stavelse, fx *ju-len* og *bru-kar*. Ofte giver de to principper samme resultat, fx *ro-lig*.

Det er en underlig situation at de to sproglige autoriteter Svenska Akademien og Svenska språknämnden er uenige om en så forholdsvis grundlæggende ting som ord-deling. Det forlyder at Dansk Sprognævn omgås med planer om orddeling i næste retskrivningsordbog. Ideen er god, og i givet fald bliver det interessant at se hvordan man vil prioritere de eksisterende principper for orddeling i dansk i forhold til hinanden.

Normering og rekommendation

Uanset at SAOL formelt er en samling rekommendationer af svenske ords stavning, bøjning m.m., opfattes ordbogen som normerende af sine brugere. Men i visse tilfælde tilbyder den også brugerne valgmuligheder. Ligesom RO har SAOL dobbeltformer, ofte endda med indbyrdes prioritering. Dobbeltformer som redaktionen betragter som lige gode, anføres i én artikel med et *el.* mellem formerne. Eksempler fra bogstav M er *majsena/maizena* ‘majsmel’ og *mor/moder*. Variant nr. 2 står med mindre skrift end nr. 1, og der henvises fra variant nr. 2 til nr. 1 med formuleringen *variantform till ...* Det får nu let den konsekvens at brugeren regner variant nr. 2 for lidt mindre god end nr. 1. RO giver derimod begge former akkurat samme behandling. I andre tilfælde er det netop meningen at anbefale variant nr. 2 som et alternativ i anden række. Så står der *äv.* (dvs. *även* ‘også’) mellem de to former, igen med forskellig skriftstørrelse og henvisning fra variant nr. 2. Eksempler er *mascara/maskara* og *mazurka/masurka*. Ved valgmulighed for genus og bøjning bruges samme praksis med *el.* og *äv.* Det er et kritisabelt punkt ved behandlingen af dobbeltformer at førsteledssammensætninger med et ord der har dobbeltformer, ikke anføres med dobbeltformer i ordbogen. Eksempel: For dobbeltformerne *id/ID* ‘identitetspapir’ er der tre førsteledssammensætninger: *id-bricka*, *id-handling* og *id-kort*, spredt ud over halvanden side i afstand fra selve artiklen *id/ID* så man ikke kan se at også sammensætningerne kan skrives med *ID-*.

Med den anvendte graduering af normering og rekommendation har SAOL-redaktionen et smidigt redskab til afspejling af skiftende strømninger i sprogbrugen og modsat til forsøg på regulering af sprogbrugen i en ønsket retning. På den anden side kan ordbogsbrugerne få et flimrende billede af normen, især fra udgave til udgave af SAOL. Spørgsmålet om dobbeltformer har været aktuelt i adskillige år i Dansk Sprognævn (og for sprogbrugerne!) og er det fortsat. Derfor kan det være interessant at sammenligne omfanget af dobbeltformer i SAOL og RO. Som stikprøve har jeg optalt dobbeltformerne i bogstav M i de to ordbøger og fundet at der er 37 *el.*-par og 33 *äv.*-par i SAOL og 92 *el.*-par i RO. I betragtning af at der antagelig er dobbelt så mange m-ord i SAOL som i RO, ser det ud til at RO har betydelig flere dobbeltformer end SAOL.

Mens medtagelsen af dobbeltformer vel kan opfattes og bruges som et sprogliberalt tiltag, forholder SAOL sig puristisk regulerende til behandlingen af engelske låneord. Ofte anbefales et svensk alternativ til det engelske ord, fx ved *interface* “Använd hellre *gränssnitt*” og ved *backslash* “Använd hellre *bakstreck*”. Der er også puristiske indslag

i stavningen af engelske låneord, ofte med den engelske form som 2. (ligestillede) variant, fx *dejt/date, dejta, dejt/fight, fejka, mejl, tajma, tajts/tights* og *tejp/tape*. Og de engelske ords bøjning søges tilnærmet svensk, fx *interface -t, =* og *backslash -en, -ar*. Hvor pluralisbøjning med *-s* overhovedet er anført som en mulighed, anbefales en svensk bøjningsform “hellre än *-s*”. Der er tale om en sprogrøgt der er forståelig ud fra Svenska Akademiens vedtægter citeret først i anmeldelsen, men som muligvis kommer til kort over for sprogbroen i visse tilfælde.

En sprogrøgt af tvivlsom karakter møder man på den politiske korrektheds slagmark. I indledningen til SAOL13 står der at der fra udgave til udgave udgår ord der er blevet “socialt olämpliga, fx *negerby* och *hednamission*”. I SAOL13 har ordet *neger* etiketten <kan uppfattas som nedsättande>, lad gå med det, og der er tre førsteleds-sammensætninger: *negerslav* <hist.>, *negerboll* <vard.> ... “Använd hellre *chokladboll*”, og *negerkyss* <vard.> ... “Använd hellre *chokladkyss*”. Det fortælles at SAOL blev meldt til den svenske diskrimineringsombudsmand på grund af de to sidste opslagsord i SAOL12; der stod bare <vard.> *sötsak*. Jeg synes redaktionen har lyttet for meget til de røster der har hævet sig for at få anstødelige ord fjernet fra SAOL og dermed forhåbentlig fra sproget. I forordet hedder det ellers at “Sådan makt har inte denna ordlista och är den inte avsedd att ha”. Men bevares, de anstødelige og politisk ukorrekte ord bidrager måske ikke ligefrem til det svenske sprogs “höghet”.

Med det kendskab jeg har fået til SAOL og til Svenska Akademiens øvrige værker, synes jeg at Gustav III, Akademiens stifter, i sin himmel kan glæde sig over virkeliggørelsen af sine fremsynede planer til fremme af det svenske sprog, det sprog som er ét af flere officielle sprog i EU og officielt sprog i Finland sammen med finsk, men som efter den svenske riksdags mening kun må være hovedsprog i Sverige, dog under protest fra Svenska Akademien. Det er imidlertid en helt anden historie.

Kilder

Mikael Reuter: SAOL12 i: *LexicoNordica* 6-1999, s. 211-226.

Hæftet *Svenska Akademien. Fakta om Svenska Akademien*.

Avisartikler fra *Aftonbladet* 19. april 2006 og *Bohusläningen* 20. april 2006.

Referat fra 8. Nordiske Dialektologkonference Århus 2006

Af Asgerd Gudiksen

I midten af august (fra den 15. til den 18.) var der nordisk dialektkonference på Århus Universitet, nærmere bestemt i de smukke og funktionelle Søauditorier ved søen i Århus Universitetspark. Konferencen var arrangeret af Peter Skautrup Centret for Jysk Dialektforskning, med andre ord Jysk Ordbog. Det er kun en lille institution, så organisationskomiteen omfattede samtlige ansatte: Torben Arboe, Viggo Sørensen, Inger Schoonderbeek Hansen, Nina Grøftehaug og Kirsten Lyshøj.

Traditionen tro mødtes deltagerne aftenen før konferencen officielt blev åbnet, denne gang til en overdådig modtagelse på Århus Rådhus, hvor et medlem af byrådet fortalte om Århus før og nu og ønskede os en god konference. Og det fik vi. Noget kunne tyde på at ordbogsarbejde er en god skoling til konferencearrangering. I hvert fald var det en helt igennem velorganiseret konference både hvad angår mikro- og makrostrukturen. Så det var fuldt fortjent at konferencedagene var velsignet med godt vejr til trods for at det vistnok har været den mest regnfulde august siden 1964.

Konferenceudflugten er vi jo mange der altid ser frem til. Og den var som alt det andet godt planlagt. Vi startede med en kort køretur sydpå til oldtidsmuseet Moesgaard. På forhånd kunne man jo være bekymret for om det overhovedet kunne lade sig gøre at få noget ud af en time på et museum som Moesgaard, men det kunne det. Der blev ikke tid til at studere detaljer, men inddelt i tre "roterende" grupper fik vi tre meget interessante rundvisninger om udvalgte genstande: Grauballemanden, stykker fra våbenofringerne i Illerup Ådal og fra museets udstilling om vikingetidens Århus. Næste oplevelse var Ebeltoft og fregatten Jylland. Igen blev vi delt i tre hold, men denne gang havde vi samme guide under hele rundvisningen. Hvordan det gik de øvrige hold, kan jeg derfor ikke oplyse noget om. Men mit hold havde en meget energisk dame (nogen syntes lovlig energisk) der myndigt jog os rundt på skibet, alt imens hun fyldte os med oplysninger om krudt og kanoner, om skibsbygning og skibsførplejning, om slaget ved Helgoland og toiletforholdene på sejlskibene. Noget kunne tyde på at også de andre guider har været effektive, for vi var hjemme igen nøjagtigt på det klokkeslæt der var angivet i programmet, kl. 19.30. Om vi slet ikke fik noget at leve af hele den lange eftermiddag? Jo, det gjorde vi rigtignok. Inden rundvisningen på fregatten var der kaffe med hjemmebagt, varm kringle og lagkage med syltetøj, så det var ikke kun en tiltrængt forfriskning vi fik, men også et godt eksempel på traditionel dansk kagekultur når den er bedst.

Men selv for de mest inkarnerede udflugtsgæster gør udflugten det ikke alene. Der skal også være et fagligt udbytte hvis konferencen skal være vellykket. Og det fik vi også med to og en halv dages foredrag, og en aftensammenkomst der ikke kun omfattede "selskabeligt samvær over chips og øl/vin", men også Peter Klitgaards dokumentarfilm om danskerne og deres dialekter: *I Danmark er jeg født*.

Peter Trudgill og Kurt Braunmüller holdt de to plenarforelæsninger. Emnet for Peter Trudgills foredrag var sprogforandring. Det er en almindelig antagelse at sprogforandringer altid medfører forenklinger, men denne forestilling hidrører fra at man inden for (socio)lingvistikken især har studeret nyere sprogsamfund med høj grad af sprogkontakt. Peter Trudgill fremlagde en række modeksempler på forandringer der har medført øget kompleksitet, hentet fra isolerede sprogsamfund, bl.a. Færøerne. Førhen

har isolerede sprogsamfund været meget mere almindelige, og man må derfor være forsigtig når man vil “use the present to explain the past”. Også Kurt Braummüller beskæftigede sig med sprogforandring. I foredraget med titlen *Dialects, varieties and historical linguistics* viste han, ligeledes med skandinaviske eksempler, at man kan risikere at nå frem til misvisende resultater inden for analyser af sprogforandring hvis man ikke tager dialekterne i betragtning.

Sektionsforedragene var tematisk opdelt og blev afviklet i to parallelle sektioner. Den første foredragsdag indeholdt bl.a. en sektion med foredrag inden for dialektleksikografi. Lars Vikør fra Norsk Ordbok 2014 præsenterede ordbogens reviderede regler for hvordan dialektal variation skal beskrives i ordbogen: Hvilke typer af variation skal medtages, og hvordan kan udtaleformerne bedst forenkles. Inger Schoonderbeek Hansen redegjorde for sin redaktion af pronomenet *hvad* til Jysk Ordbog, som et eksempel på overvejelser redaktører af en dialektordbog må gøre sig ved redigeringen af ord der er fælles for dialekterne og standardsproget. Karin Hallén, Svenskt Dialektlexikon, fremlagde inventaret af franske låneord i ordbogsarkivet under bogstav *e*, gjorde rede for ordenes brug og udbredelse i svenske dialekter og i de øvrige skandinaviske dialekter (ved hjælp af Jysk Ordbog og Ømålsordbogen). Afslutningsvis konkluderede hun at det var velmotiveret at medtage mange af låneordene i en dialektordbog, men med SDL’s begrænsede omfang ville det nok kun blive få franske lån der i sidste ende ville blive medtaget. Tor Erik Jenstad og Gunhild Aamodts foredrag handlede ligeledes om dialektal brug af låneord, idet de gjorde rede for et forsøg med at udnytte Norsk Ordbok til at undersøge udbredelsen af nedertyske og hollandske lån. I femte og sidste foredrag i den leksikografiske sektion fortalte Elisabeth Lien, Øyvind Eide og Lars Jørgen Tvedt (ligeledes fra Norsk Ordbok) om opmærksomheden af lokale ordsamlinger dels til brug for Norsk Ordbok, dels med henblik på varig lagring og formidling til andre sprogfolk.

Termen *dialekt* har i de senere år fået en udvidet betydning. Tidligere brugte man (i hvert fald i Skandinavien) dialekt om de traditionelle dialekter på landet, men nu kan dialekt også bruges om nyere og mere standardnære varieteter både i byerne og på landet. Den udvidede betydning af dialekt afspejlede sig i conferenceprogrammet. Ud af i alt 40 sektionsforedrag handlede de 14 om undersøgelse af nyere varieteter ud fra især sociolingvistisk teori og metode. Nogle få eksempler. Malene Monka holdt et foredrag om sprogalternation mellem rigsdansk og sønderjysk på grundlag af optagelser med fem sønderjyder i servicesituationer og diskuterede i hvor høj grad en sprogforsker har mulighed for at forklare hvorfor sprogbrugerne skifter mellem varieteterne. Agnete Nesse fremlagde en undersøgelse af radiotalesproget i Nordnorge mellem 1936 og 1983. Hovedresultatet af undersøgelsen var at brugen af dialekt var øget, mens brugen af formel tiltale (med *De*, brug af titler osv.) næsten var faldet bort. Margareta Svahn talte om “tendenser i skarborgskt ungdomsspråk”, og Anna Greggas Bäckström præsenterede unge finlandssvenskeres brug af dialekt i sms’er.

Konferencen var også præget af at der var en del foredrag om syntaktiske fænomener, for en dialektkonference at være usædvanlig mange. Syntaks har været et noget forsømt område inden for dialektologien, bl.a. fordi det inden båndoptageren kom til, var vanskeligt at fastholde større mængder “tekst”. For at rette op på denne skævhed har en række forskere dannet en “projektparaply”, Nordisk Dialektsyntaks (ScanDiaSyn), der dels har som formål at samordne og stimulere udforskning af dialektal variation inden for syntaks, dels at oprette en database med transskriberet og tagget materiale som skal være alment tilgængelig. Projektet havde en workshop der indledtes med at lederen af projektet, Øystein Vangsnes, fortalte om hvor langt man var nået (se

<http://uit.no/scandiasyn>) og fortsattes med otte sektionsforedrag om syntaktiske fænomener i nordiske dialekter.

Ud over de allerede nævnte sektioner var der en lille sektion med foredrag om prosodi og accentuation og en sektion med meget blandet indhold, men med det fællestræk at foredragene især beskæftigede sig med traditionelle dialekter. Henrik Hovmark præsenterede fx et projekt om retningsadverbier i dansk, hvor han specielt fokuserede på hvordan et lokalsamfunds interne og eksterne relationer kan være bestemmende for brugen af retningsadverbier. Therese Leinonen præsenterede et projekt om at måle den lingvistiske afstand mellem svenske dialekter ved hjælp af datalingvistiske metoder udviklet af det dialektmetriske projekt *Determinant of Dialectal Variation* ved universitetet i Groningen. Og Anna Westerberg fremlagde sine planer for en ny lærebog i dialektologi.

Jeg må nøjes med disse få eksempler, men foredragene fra konferencen vil blive publiceret i en konferencerapport udgivet af Peter Skautrup Centret.

Konferencemiddagen blev vi naturligvis heller ikke snydt for. Den foregik i Stakladen, endnu en af Århus Universitets kønne bygninger. Der var tre retter dejlig mad, underholdning ved det dygtige kor *Vocalline* og taler. Og så blev der naturligvis snakket både fagligt og ufagligt inden vi gik ud i silende regn, en tydelig markering af at nu var konferencen helt slut.

EURALEX XII i Torino – Referat og anmeldelse

Af Jens Laursen Pind, cand.mag., Euralex-medlem

‘Er det ikke lidt for nemt at kritisere store ordbøger ud fra en *forbedringsstrategi*, når det må være muligt at udarbejde systematiske kriterier?’

Dette præcise, her let modificerede, faglige spørgsmål fra ordstyreren Pius ten Hacken i “Aula 5” var nok det første, der overhovedet blev stillet under det netop afsluttede leksikograf-træf i Torino – og dermed er vi allerede midt i denne spændende kongres, afholdt 6.-9. september 2006 i den italienske storby. For så vidt kunne flere foredragsholdere have spurgt sig selv om det samme langt tidligere, idet der allerede ifølge de udsendte abstracts var optræk til intense detailstudier af kendte ordbøger, sommetider måske uden det nyttige matchende overblik.

Hvor *turisten* nærmest oplevede bymiljøet Torino som et betagende vidunder af skønhed og atmosfære, kom *kongresdeltageren* hurtigt ned på jorden igen: Palazzo Nuovo, det massive og dynamiske hovedsæde for det humanistiske fakultet skråt bag Mole Antonelliana, stod kun til rådighed for et par “pre-congress tutorials”: Pfister/Ronco + Eurac – hvorimod hovedkongressen var henlagt til mere beskedne rammer i det lidt isolerede Palazzina Einaudi.

Af de 333 realiserede kongrestilmeldinger kom under 100 fra italienere. Antallet af italiensktalende udlændinge er ukendt, men næppe kæmpestort. Alligevel var et af problemerne i Torino den mindre tilfredsstillende kommunikation, specielt vilkårene for engelsktalende.

Rundt omkring i det trykte kongresprogram dukker der umotiveret uoversatte lokale elementer op, og det kunne indimellem være vanskeligt at få et brugbart svar på væsentlige spørgsmål. Grunden var nok den simple, at ikke mange arrangørrepræsentanter følte sig trygge ved mundtligt engelsk.

I den situation er det naturligtvis nemmere at holde sig til sit modersmål, så ca. 75 % af den generelle taletid ved den officielle åbning af konferencen i Aula Magna del Rettorato var helliget indlæg på lokalsproget, en urimelig skævvridning og forskelsbehandling af de mange udenlandske deltagere uden sikre italienskunderskaber. Som en logisk konsekvens valgte mange af gæsterne at sive fra den langstrakte åbning, i afventen af den forjættede buffet.

Bygningen i Via Verdi var ellers klassisk karakterfuld, og da åbningshøjtideligheden blev en del forsinket, fik nogle af de tilstedeværende mulighed for i små hold at komme til forpremiere på biblioteksudstillingen “The Taming of the Word – A Selection of Dictionaries from the 16th to the 20th centuries” sammesteds. Det var spændende, med tankevækkende ordbøger – og samtidig dunkelt og utilnærmeligt, som den slags støvede samlinger vel i et eller andet omfang forventes at være det. Mon de dog ikke fandt ud af at gøre det lidt mere lyst og indbydende de kommende to dage, hvor offentligheden var lovet adgang?

Ved onsdagens åbningsceremoni kanaliseredes eventuel offentlig interesse dog primært gennem de snurrende kameraer, der også lige nåede frem til afslutningen lørdag middag i Palazzina Einaudi.

Af en eller anden mærkværdig grund havde initiativtagerne til Torino-forløbet helt fra de første forskræp i sommeren 2004 lagt sig fast på, at Euralex XII skulle være en dag kortere end forgængeren – mens nogle måske vil erindre, at de fortræffelige arrangører

af Euralex XI i Bretagne undervejs, i planlægningsfasen, overvejede mulighederne for en hensigtsmæssig *forlængelse* af kongressen.

Lad os sammenligne: Lorient 2004 (5 dage) rådede over 34 slots til ca. 105 oplæg, mens vi i Torino 2006 (4 dage) var tvunget til at presse omkring 150 bidrag ind på kun 28 slots. Det kan ikke undgå at give problemer. For i det hele taget at kunne gennemtvinge planen måtte arrangørerne gøre store dele af kongressen *femsporet*, en struktur, der – selv om man i princippet jo altid kan kontakte en tilstedeværende forfatter direkte, eller blot læse de trykte versioner i beretningen – ubestrideligt er egnet til at samle opmærksomheden om de(t) kendte og sikre. Således kunne mange mere snævert anlagte foredrag i Torino *højest* påregne 15-20 tilhørere. Og de allermest populære? Vel omkring 150.

I Lorient indgik en “Poster Session” med 12 bidrag, som kunne glæde sig over at have 80 min. til rådighed; under Euralex XII havde man i det tilsvarende slot kun 100 min. tilovers for op til 31 bidrag (enkelte lod sig overføre til selvstændig fremlæggelse i auditorierne). Det siger vist sig selv, at både forfatterne i denne kategori og deres betydelige publikum blev budt langt dårligere vilkår i Italien 2006 end i Frankrig 2004.

Torsdag eftermiddag måtte **Janet DeCesaris** foredrage om “Lexicography as [R]eported: the EURALEX Conference Proceedings Bibliography (1983-2004)”, et værk hun skulle udarbejde sammen med Elisenda Bernal. Bibliografien var for længst lovet færdig til kongressen, hvilket havde vist sig at være alt for optimistisk, og dette gav DeCesaris anledning til i trettende time at indhente forslag til det passende format (medie) for værket (!). På baggrund af den lejlighedsvis forbløffende besvarelse af forsamlingens spørgsmål (bl.a. om praktiske søgefunktioner) virkede det ærlig talt ikke som om den normalt meget overbevisende DeCesaris var helt *parat* til opgaven. Nu må vi sætte vores lid til, at den i anden sammenhæng oplyste frist for færdiggørelsen på ca. to måneder holder nogenlunde.

Avancerede og nyttige søgefunktioner en masse finder leksikografen til gengæld i den digitale version af tidsskriftet “Der Brenner”, som udspringer af de enorme tekstsamlinger hos AAC (Austrian Academy Corpus), en *forskningsenhed* under Österreichische Akademie der Wissenschaften, Wien – med et elektronisk tekstkorpus over det 20. århundredes tyske sprog på 1 milliard ord som et af de (foreløbige) perspektiver for AAC’s virke, til dels i samarbejde med andre akademier. “AAC” anvendes dog også som nøgtern betegnelse for det *korpus*, der forventes at komme ud af anstrengelserne, og som det er tanken at stille til rådighed online.

Genoplivningen af det førende østrigske litterært-kulturelle tidsskrift, udgivet 1910-1954, er kort sagt en demonstrations- og testmodel, en *filologisk* digital udgave af den ærværdige publikation.

Dr. **Hanno Biber** fra AAC præsenterede onsdag eftermiddag Euralex-tilhørerne for det velgennemtænkte og veldesignede interface, de præcise søgefunktioner og den gigantiske informationsmængde, der efterhånden bliver tilgængelig for eksterne forskere, men både han og vi kunne såmænd sagtens have brugt det dobbelte af de tilmålte 20 minutter.

Bibers engelsksprogede gennemgang i *Proceedings* [395] giver et solidt indtryk af projektets tyngde og rækkevidde.

Efterhånden som den menneskelige viden og specialiseringen successivt vokser, stiger også behovet for kompetente og pålidelige forsker-formidlere. En sådan er **Willy Martin**, professor i leksikologi fra Amsterdam, der onsdag glædede gæsterne i auditorium 5 med en engageret, velstruktureret oversigt over, hvordan ‘Minskyan frames’ kan bidrage til forbedrede Advanced Learner’s Dictionaries. En faglig intro-

duktion blev kombineret med præcise og facetterede definitioner og åbne, hensigtsmæssige teser på vejen frem mod det teoretiske ideal – den perfekt brugertilpassede monolingvale ordbog. Resultatet bliver ifølge leksikologen bl.a. eksakte beskrivelser af ord som realiserede typer og en produktiv indsigt i kollokationer og deres mulige mønstre.

Martin nåede næppe at få fremlagt meget mere end *halvdelen* af alt det, der optog ham, men *Proceedings* [281] byder på det komplette manuskript: “Frame-based Lexicons and the Making of Dictionaries”.

Mens EURALEX XII lakkede mod enden, havde **Nilda Ruimy** fra Istituto di Linguistica Computazionale i Pisa afsat sin lørdag formiddag til at orientere en begrænset skare (20 heldige ...) om “A Computational Multi-Layered Italian Lexicon for HLT Applications”, et oplæg, der her kun skal fremhæves for sit ekstremt vellykkede og æstetisk tiltalende grafiske præsentationsniveau. Den blændende demonstration af interaktive diagrammer mv. var ganske enkelt i særklasse blandt alt, hvad i hvert fald jeg fik lejlighed til at opleve i de italienske universitetslokaler!

I kongressens sidste slot, lørdag fra 12.00-12.20, redegjorde **Andrea Abel** fra Europäische Akademie Bozen for en tysk-italiensk valensbaseret online-ordbog for *begyndere* (ELDIT), hvor verbet betragtes som sætningens strukturelle centrum, en analyse, der er uproblematisk og korrekt. Hun var på den ene side overbevist om online-tilgangens fordele, men savnede selvsagt indtil videre grundige brugerundersøgelser og -feedback, som kunne hjælpe ordbogskonstruktørerne med at finpudse den didaktisk vanskelige fremstilling af valensstrukturerne.

For de tyskinteresserede rummer *Proceedings* [1091] en række systematiske overvejelser over valensvinklens faldgruber, udarbejdet af Andrea Abel og Claudia Richter.

Janet DeCesaris – der muligvis får endnu mere travlt under EURALEX XIII i Barcelona – havde sin første torsdagsopgave lige efter frokost med en fremragende analyse af “A New Kind of Dictionary: REDES [...]”, en *diccionario combinatorio* fra 2004, baseret på et stort spansk journalistisk korpus og skabt af professor Ignacio Bosque, der fortsat heller ikke selv betragter sig som leksikograf, men som grammatiker. Taleren placerede specialordbogen i dens rette nationale og faglige kontekst, nævnte påvirkningen fra Fillmore, og man kunne mærke hendes beundring for ophavsmandens innovative præstation. Alligevel fremhævede hun loyalt værkets relative svagheder, det ensidige korpus, den restriktive lemmaselektion og en ikke særlig brugervenlig eller konsekvent struktur. DeCesaris gjorde en del ud af professorens ekstraordinært lange introduktion(er) til ordbogen, som nok også er et af dens største aktiver, og som sammen med værkets omfattende datamængde givetvis vil inspirere kolleger og leksikografer til at raffinere den spanske lingvists koncept.

I *Proceedings* [399] finder man det tilhørende meget læseværdige manuskript af foredragsholderen og Paz Battaner, begge fra Universitat Pompeu Fabra, Barcelona.

I betragtning af at 2006 vist hidtil har været et usædvanligt begivenhedsrigt år for Torino, kunne man vel forvente, at byens borgere efterhånden måtte være mætte, men der indløb skam stadig nye aktiviteter, herunder “Torino Settembre Musica – 29th edition”. Organisationskomitéen for Euralex XII var blevet så begejstrede for denne vistnok rent klassiske musikfestival, at de havde medtaget *dette års udgaves komplette program* på kongressens hjemmeside www.euralex2006.unito.it – hvor der ellers kun fandtes relevante oplysninger om tilmelding, transport, indkvartering, foredragsindhold ...

“Lancia Club Turin International Meeting 2006” (med talrige lettere bedagede

eksemplarer af dette mærke) var også en mulighed for de oplevelseshungrende, eller hvad med de mange initiativer, der kunne udspringe af UNESCO's dobbelte italienske hædersbevisning: "TORINO CAPITALE MONDIALE DEL LIBRO CON ROMA" (april 2006 til april 2007)? Både bilmødet og bogåret havde dog større held til at unddrage sig leksikografi-arrangørernes opmærksomhed end den festlige musik.

Fredag middag havde karismatiske **Patrick Hanks** (Berlin-Brandeis-Brno) tilsagt fans og forskere i "Aula 4", hvor **Thierry Fontenelle** sad klar som den kompetente og tålmodige ankermand for en turbulent session, der skulle blive præget af dårlig forbedelse, praktiske problemer og utilstrækkeligt forelæsningsmateriale.

Efter ordstyrerens sikre introduktion begyndte kosmopolitten sin uorganiserede odysse, med universitetets italienske laptop som modvillig følgesvend – frem til forestillingens første forstyrrende mellemspill fra en bekymret Patrick Hanks (frit citeret) med blikket på den bærbare projicerede skærbillede: "Oh! That's an old version ...", hvorefter foredraget slingrede videre – indtil næste forhindring, og en bestyrtet Patrick Hanks: "Oh! That's a *very* old version ...".

Som en bedre rygsækturist gav Hanks sig nu til at rode i oppakningen, hvilket snart bragte både en laptop og en USB-nøgle for dagen. Begge dele blev overladt til Fontenelle, der fik til opgave at opspore en aktualiseret version (den nyeste viste sig faktisk at være fra samme dag, 8. september) af Hanks-oplægget på den senest ankomne harddisk.

Mens Microsoft-manden kæmpede for at få liv i foredragsholderens computer, stred hovedpersonen selv sig møjsommeligt igennem det lagrede forlæg, holdt oppe af udsigten til snart at blive undsat.

Under de efterfølgende forviklinger afgav Hanks i et uselvisk øjeblik sin mikrofon til lidelsesfællen ved bordet, mens han samlede tanker og zoomede ind på dagens tekst "The Organization of the Lexicon: Semantic Types and Lexical Sets".

Da alle tilstedeværende således omsider var på plads og i orden, ændrede fremlæggelsen karakter og blev lige så inspirerende som talerens engagement i det forventeligt innovative projekt omkring en "Pattern Dictionary of English" i NLP-regi.

Denne løbske session overskred uundgåeligt alle rimelige tidsfrister, til skade for dem, der havde planlagt at bruge auditoriet bagefter, og for potentielle tilhørere til kongressens første *round table*, om *terminology and terminography*.

Næsten samtidig med Euralex XII var der betydelig møde- og hjerneaktivitet andetsteds på kloden. På ICCG4 i Tokyo (1.-3. september) var **Charles J. Fillmore** i aktion som hovedtaler med "Articulating Lexicon and Construction", og ved samme lejlighed kunne **Collin Baker**, i identisk funktion, berige sin forsamling med "Frame Semantics in Operation: The FrameNet Lexicon and Text Understanding". Endelig var **James Pustejovsky** efter planen keynote speaker på TSD 2006 i Brno (11.-15. september) med "Unifying Semantic Annotation for Linguistic Description" – og i Torino var fredagens hovedtaler Jean Aitchison, der rutineret og kompetent causerede over avishistoriens sproglige finurligheder.

Kunne Torino-arrangørerne virkelig ikke finde noget bedre at bruge en indbudt professor og plenary speaker til? Masser af de almindelige bidragsydere til det sammenpressede kongresprogram ville også have fået noget fornuftigt ud af den tid, der her blev spildt.

Det må principielt antages, at alle regulære foredragsholdere ved Euralex-kongresser har gjort et stort forarbejde, og de fortjener derfor respekt fra deres potentielle publikum. Som kongresdeltager bør man således overveje, om ikke også et oplæg, der f.eks. kun tangerer ens fagområde, alligevel kunne blive udbytterigt. Et godt eksempel

fra Torino på de gensidige fordele ved denne strategi var fredagens sene gennemgang ved **Aurelia Martelli** fra det lokale universitet, som har analyseret et italiensk segment af “The International Corpus of Learner English”, hvor viderekomne italienske studerende også kommer til at producere forkerte engelske ordkombinationer, “collocational errors”. Det fremgik heldigvis, at den behandlede datamængde er blevet stærkt forøget siden den offentliggjorte tekst i *Proceedings* [1005]: “A Corpus[-]Based Description of English Lexical Collocations Used by Italian Advanced Learners”, men Aurelia Martelli holder fast ved sin definition af det som bekendt meget changerende nøglebegreb “collocation”. Det gik over stok og sten med speedsnakkeren Martelli, men alt var veltilrettelagt, inkl. et glimrende eksempelmateriale, og netop den kvikke præsentation gav fyldig plads til efterfølgende spørgsmål og kommentarer fra salen, mange flere end normalt. En nærmest eksemplarisk foredragssession, med en velfungerende og kommunikerende hovedperson i samspil med et sagkyndigt publikum!

Og hvor mange retfærdige havde så fundet vej til dette mønstergyldige forløb? Cirka 22 – af de langt over 300 tilmeldte kongresdeltagere.

Lørdag morgen havde **Sue Atkins** et problem, hun stod og manglede sin *chair*: “Has anyone seen Doctor Temmerman – for the last three days?”. Det var der ikke rigtig nogen, der havde, og så måtte den nærmest uundværlige Geoffrey Williams uden varsel springe ind og styre ordene under den moderat besøgte forelæsning om “Lexicographic Profiling: an Aid to Consistency in Dictionary Entry Design”.

I sin oversigt over de indledende stadier af et nyt (“Phase 1”: 2003) korpus-baseret projekt, *New English[-]Irish Dictionary*, demonstrerede Atkins, hvordan moderne ordbogsredaktion kan struktureres effektivt og professionelt, og hvor langt man i den sammenhæng kan komme med en metodisk tilgang samt en enorm ballast af leksikografisk viden og erfaring. Derved bliver også en central støttefunktion under ordbogsforberedelsen, udvælgelsen af problemsensitive og -repræsentative mønsterartikler, videnskabeligt systematiseret.

Med behørig respekt for de to oplægsholderes vidt forskellige udgangspunkter og temaer kunne det alligevel være fristende at sammenligne bidragene fra Rundell (se nedenfor) og Atkins (begge repræsentanter for Lexicography MasterClass Ltd), der synes at supplere hinanden fint mht. metode og målgruppe. Således skøjter Sue Atkins elegant hen over en komplikation, som det havde været rart at høre meget mere om, nemlig præcis på hvilket grundlag hendes og Valerie Grundys LPD (Lemma Properties Database) egentlig fandt ud af, hvilke af de – undertiden rivaliserende eller omstridte – koncepter fra fremtrædende lingvistiske teoretikere som Apresjan, Fillmore, Mel’čuk m.fl. den skulle anvende, integrere, ja måske endda koordinere, og hvordan?!

I *Proceedings* [1097] kan interesserede kaste sig over en detaljeret, punktopdelt fremstilling af sagerne, specielt hvis de formår at forlige sig med den slørede optiske gengivelse af illustrationer (figurerne), der generelt skæmmer tobindsværket noget.

Som den store programtaler ved Euralex XI i Lorient langede en veloplagt **John Sinclair** ud efter nyere generationer af leksikografer, ikke mindst dem fra konkurrerende forlag, fordi de syntes utilbøjelige til at tilpasse sig og udbrede Collins COBUILD’s Full Sentence Definitions.

Til årets kongres havde **Michael Rundell** taget denne væsentlige udfordring op. Ved hjælp af en række konkrete – og problematiske – eksempler på den COBUILDske definitionspraksis søgte han at formidle en almen oversigt over divergerende definitionstypers relative egnethed og konkluderede, at satsninger på en *bred* vifte af definitionsstrategier bedst fører til en fagligt og kommercielt vellykket MLD, Monolingual Learner’s Dictionary. Alligevel var foredraget ikke helt overbevisende.

COBUILD's definitionstilgang består i et kompliceret formelsæt med et netværk af in- og eksterne referencer, som det ikke altid er lykkedes de involverede leksikografer at præsentere optimalt, og som især da uheldigvis kan forøge risikoen for fejltolkninger og forvirring hos nogle brugere af de avancerede ordbogsopstillinger. Hvis man *systematisk* vil angribe eller blot forbedre grundlaget for COBUILD, bør man tage udgangspunkt i det abstrakt-strukturelle plan og analysere de anvendte formler. Senere kan man så fylde konkret eksempelmateriale på de fundne strukturer. Englænderen antydede under gennemgangen, nok med et glimt i øjet, at han ikke ønskede at fremstå som urimeligt kritisk over for en kollega og konkurrent ("for fear of retaliation": Rundell var jo blandt meget andet hovedredaktør på den prisbelønnede Macmillan English Dictionary for Advanced Learners fra 2002), og han kan vel også have undervurderet sit publikum. Det lidt skuffende oplæg præsenterede ikke altid de mindre heldige COBUILD-definitioner loyalt, ligesom dets ambitionsniveau ikke ydede foredragsholderen selv fuld retfærdighed.

Den publicerede version i *Proceedings* [323] kan dog bestemt anbefales.

Her finder man tilmed masser af andre bemærkelsesværdige trykte bidrag:

Lucie Barque og Anne-Laure Jousse: Analyzing Combinatorial Properties of Polysemic Lexical Units to Characterize Metaphorical Links [1109];

František Čermák: Collocations, Collocability and Dictionary [929];

Sabine Citron og Thomas Widmann: A Bilingual Corpus for Lexicographers [251];

Stephen Coffey: 'Delexical Verb + Noun' Phrases in Monolingual English Learners' Dictionaries [939];

Anna Dziemianko og Robert Lew: *When You Are Explaining the Meaning of a Word: The Effect of Abstract Noun Definition Format on Syntactic Class Identification* [857];

Laura Elliott og Sarah Williams: Pasadena: A New Editing System for the Oxford English Dictionary [257];

Ulrich Heid og Rufus H. Gouws: A Model for a Multifunctional Dictionary of Collocations [979];

Elisabetta Jezek: Argument Structures, Verb Patterns and Dictionaries [1169];

Adam Kilgarriff: Collocationality (and how to measure it) [997];

Karin Kipper, Anna Korhonen, Neville Ryant og Martha Palmer: A Large-Scale Extension of VerbNet with Novel Verb Classes [173];

Geart van der Meer: On the Status of Word Combinations in the English Monolingual Learner's Dictionaries: Entry or Collocation? [1061];

Margarita Alonso Ramos: Towards a Dynamic Way of Learning Collocations in a Second Language [909];

Gerardo Sierra, Rodrigo Alarcón, César Aguilar og Alberto Barrón: Towards the Building of a Corpus of Definitional Contexts [229];

Vasyl Starko: Entry Selection for a Bilingual Phraseological Dictionary [1045];

Petra Storjohann: New Lexicographic Approaches to the Description of Sense Relations [1201] ...

Ovenstående var blot en lille appetitvækker; det samlede udbud er nærmest overvældende.

Kongresberetningen breder sig denne gang over to tykke bind (ca. 1300 sider i alt) og er født med ISBN: 88-7694-918-6. Sektionen "Work in Progress" fra Euralex 2004 er gledet ud igen, mens oplæggene til den pragmatiske "Poster Session" denne gang er fordelt på de relevante emne-kapitler (1-12). Logisk – men der er behov for endnu en indholdsfortegnelse, hvor *hver enkelt* forfatter opføres *strengt alfabetisk* med artikel, sidetal og eventuelle medforfattere, frigjort fra den tematiske opdeling af værket.

Pga. stor lokal ekspertise og interesse er der denne gang tilkommet en ekstra indholdskategori: "Dialect Dictionaries and Linguistic Atlases" med 10 tekster. Nogen nationalistisk sproglig bølge synes dog ikke umiddelbart at være skyllet ind over det akademiske plan i Torino: 18 % af de regulære bidrag i 2006 var på værtens modersmål, mod ca. 23 % i 2004.

En anden kongresnyhed var to selvstændige, tematisk motiverede *round tables*, om hhv. "Bilingual Lexicography" (med vægt på mindre gængse sprogpar, og i programhæftet angivet som "Bilingual Dictionaries") og "Lexicography for Specialised Languages – Terminology and Terminography".

Bedømt ud fra førstnævnte (som var det eneste jeg overværede) er idéen principielt lovende. En stor håndfuld forskere deles om ét slot, hvor hver især fremlægger sit stof på sit foretrukne sprog, typisk i 8-9 minutter. Først når alle seks oplæg er afleveret, gives der (god) plads til spørgsmål. For at kunne videreudvikle konceptet, skal man nok undgå *for* stor heterogenitet, og den samlende introduktion bør ikke misbruges til at promovere arrangøren på bekostning af gruppemedlemmerne. Nogle indlæg kunne tydeligvis fint have klaret sig alene, mens andre måske aldrig ville have fået chancen ellers. Endelig må man ikke lade sig forlede af betegnelsen til at tro, at der foregik en eller anden form for drøftelse, eller samarbejde, mellem de involverede forskere. Dét er dog et af de (langsigtede og optimistiske?) perspektiver, som man kunne øjne i dette initiativ.

Forskningssamarbejde er generelt et dynamisk felt; på Torino-kongressen havde 37 % af de almindelige indlæg mindst to forfattere, en mindre stigning i forhold til 2004, men et mindre fald sammenholdt med 2002.

Rigtig mange foredragsholdere havde glædeligvis en masse på hjerte, og havde forberedt skriftligt materiale, som det i praksis ville blive umuligt at komme igennem, selv på 20 mundrappe minutter. Nogle valgte i den situation at skære en del mellemregninger og detaljer væk for trods alt at kunne levere et veldisponeret og afrundet oplæg, mens enkelte andre synes at sværge til en bare mas på-strategi. Begge løsninger fungerede egentlig udmærket.

Ved fremlæggelsen spiller de ofte undertallige *handouts* derfor en vigtig rolle. Disse betragtes normalt blot som forhåndenværende støttemateriale under gennemgangen i auditoriet, men hvor de rummer nye data eller en revideret version af (dele af) ophavsmandens tekst, bør alt sådant som minimum stilles til rådighed for samtlige interesserede senest dagen efter hovedindlægget. Den opgave må det organisatoriske sekretariat kunne påtage sig.

Med sig hjem fra en mangesidet kongres tager man naturligvis en mængde informationer, samt en række mentale øjebliksbilleder af de *mennesker*, der deltog.

Det gælder f.eks. Reinhard R. K. Hartmann, der lyste af glæde over den organisation – Euralex – som han selv tog initiativ til for snart et kvart århundrede siden; Pius ten Hacken, der præsenterede en rå side (som ordstyrer) og en poleret (i pauserne); Anthony P. Cowie, der skabte en behagelig stemning overalt hvor han var nærværende, beredvilligt delte ud af sin viden, og viste sig som en god lytter; og sidstnævntes temperamentsmæssige modstykke, de flygtige, hektiske, der kun kom til syne i korte, klare glimt: Adam Kilgariff, Sylvie De Cock, Gilles-Maurice de Schryver, Janet De-Cesaris – men som netop derved bidrog til det flow, der også fører til fortsat fremdrift.

Hvordan klarede Torino-arrangementet sig så sammenholdt med forgængeren i nabolandet Frankrig?

Umiddelbart kan 2006-udgaven synes langt mindre farverig end kollegaen. Sidst, i Bretagne, var vi således tvunget til at bryde ind på det lokale universitet og smadre dørlåsen, hvilket uvægerligt fik tyverialarmen til at gå i gang ... Slige spændingsmomenter måtte vi helt undvære i år, på linje med fransk syndflod, excentrisk bygningsindretning og timelange strømafbrydelser.

Lidt anonyme, undertiden fabriksagtige – sådan var min personlige oplevelse af de italienske kongres-lokaliteter. Det gælder det såkaldte Palazzina Einaudi, der normalt vist mest bruges af det juridiske fakultet, inkl. statskundskab, samt den noget afsidesliggende kantinebygning. Der var ganske enkelt slet ikke den *esprit* over Torino-komplekset, som Lorient helt naturligt bare besad. Hvor Bretagnes auditorier bød velkommen med smukke, klangfulde navne som “Lecture Room Massiac” og “Soleil d’Orient”, blev vi i Piemonte dirigeret ind i “Room 4” (“Aula 4”), “Room 5”, og så videre.

Underdimensioneringen var et gennemgående problem i den italienske metropol, begyndende med registreringen onsdag formiddag. Allerede på dette meget tidlige tidspunkt savnede jeg uvilkårligt faciliteterne fra “Maison de la Recherche” i den fjerne franske provins.

Dém havde vi så tid til at drømme om, mens vi stod i kø – ved buffeten onsdag aften, til den daglige frokost i kantinen (om torsdagen hang bunker af uheldige kongresdeltagere så fast i madkøen, at de gik glip af mindst ét foredrag), og i *hver eneste* kaffepause med den garanterede voldsomme trængsel omkring den rytmisk dunkende mokka-maskine ude på fløjen. Til gengæld trækker den flotte kongresberetning med ikke mindre end 147 + 5 bidrag selvsagt op.

Organisatorisk og praktisk formåede Euralex XII altså ikke helt at matche forgængerens standard. De problemer der opstod i Lorient – hvor spektakulære de end måtte være – skyldtes enten force majeure eller usystematiske fejl; svaghederne i Torino-konstruktionen, derimod, var *strukturelle* og *holdningsmæssige*.

Undervejs blev *Anna Braasch* – der allerede fra kongressens start måtte fungere som leder – officielt kåret til ny *President*, mens *Lars Trap-Jensen* lige så smertefrit gled ind som *Assistant Secretary-Treasurer*. Sammen med den øvrige bestyrelse og de lokale catalanske arrangører skal de således føre Euralex frem til 2008-kongressen i Barcelona.

Set med danske solbriller var det altså nok et vellykket forløb, men alligevel kan EURALEX XII blandt medlemmerne generelt risikere snarere at blive husket for sin kaotiske generalforsamling.

Selv vil jeg gerne kreditere kongreshelheden 2006 for det faglige niveau i auditorierne, for de to store bind med *Proceedings* og for den helt usædvanligt tiltalende værftsby.

Rapport fra LEDA-udflugt til Svenska Akademiens ordbok (SAOB) i Lund

Af Henrik Hovmark

Onsdag den 20. september 2006 drog 25 LEDA-medlemmer af sted mod Lund for at aflægge Svenska Akademiens ordbok (SAOB) et besøg. Forventningerne var store til besøget på denne kolossale og næsten sagnomspundne ordbog. Medarbejderne på SAOB bød generøst på lunsj kl. 12, der var lagt et detaljeret program for eftermiddagen – og vejret viste sig fra sin smukkeste side.

Men ingen udflugt uden vanskeligheder. Regionaltoget til Malmö afgår fra spor 26 på Københavns Hovedbanegård. Eller også fra spor 6 ... Undertiden også fra spor 5 ... DSB's veje er uransagelige og svære at hitte rede i, selv for hærdede leksikografer. Denne formiddag afgik toget fra spor 5, og det var desværre ikke alle LEDA-medlemmer der bestod DSB's test. Tre af os endte med at danne bagtrop i et senere tog, herunder undertegnede. Stor var bekymringerne hos bagtroppen – ville vi mon nå frem til tiden (dvs. lunsj)?! Det gjorde vi heldigvis. Vi nød maden, inklusive letøl og kaffe, og følte os på alle måder velkomne.

Det faglige program startede med oplæg ved ordbogschef Anki Mattisson, der fortalte om den bygning som SAOB siden 1998 har haft til huse i, om finansieringen af SAOB og om kriterierne for udvælgelse, eller rettere: udeladelse af opslagsord. Derefter fortalte førsteredaktør Birgit Eaker om de forskellige typer af flerordsforbindelser og deres behandling i ordbogen. Endelig var der rundvisning på redaktionen og introduktion til ordbogens nye elektroniske redigeringsystem.

Redaktionen, der består af 20 fuldtidsstillinger fordelt på 21 personer, føler sig hjemme i de nuværende lokaler, hvor de agter at blive indtil SAOB efter planen er færdigredigeret i 2017. Rent faktisk er huset oprindeligt et privathjem, tegnet af arkitekten Sten Samuelsson til sig selv og sin store familie i 1960'erne. Efter at familien solgte villaen, blev den lavet om til kontordomicil i en periode, men SAOB og Svenska Akademien har efter overtagelsen forsøgt at føre den tilbage til sin oprindelige plan. Sten Samuelsson har bl.a. også tegnet Karlskrona Kunsthall, og ligesom denne er SAOB's hjem en atriuminspireret bygning. Atriumrummet, der bruges til frokoststue og mødelokale, havde skydetag, og da vejret som sagt var smukt, blev der åbnet på klem og et gyldent blad eller to dalede yndefuldt ned mellem stolerækkerne under de to foredrag. I SAOB's bygning mødes gammelt og nyt: gulnede blade og nyt redigeringsystem; på gangen står et gammelt solidt træskrivebord med specialdesignede dueslag til SAOB-*lappar* – og ovenpå en anelig antik blyantspidser, vistnok af støbejern og med drejehåndtag. Antik, men absolut duelig, viste det sig da førsteredaktøren foretog en spontan efterprøvning.

SAOB finansieres som bekendt af Svenska Akademien, der også tager sig af SAOL (Svenska Akademiens ordlista, den svenske retskrivningsordbog) og af den litterære nobelpris. Svenska Akademien og dermed SAOB har én vigtig, fast indtægtskilde, nemlig indtægterne fra Post- och Inrikes Tidningar, den svenske pendant til Statstidende (eller også er det omvendt, for Post- och Inrikes Tidningar er udkommet uafbrudt siden 1645, hvilket er verdensrekord!). Dette privilegium blev imidlertid anfægtet for nylig af et privat patentfirma, og dermed blev SAOB's finansiering pludselig alvorligt truet. Efter en retssag er Svenska Akademiens rettigheder til denne indtægtskilde dog heldigvis blevet stadfæstet, og dermed er finansieringen af ordbogens færdiggørelse sikret.

SAOB er en meget stor ordbog, på størrelse med Oxford English Dictionary, Deutsches Wörterbuch og Woordenboek der Nederlandsche Taal. Ligesom disse har den været længe undervejs. Beslutningen om at lave en ordbog blev taget helt tilbage i 1787, men først i 1883 kom arbejdet for alvor i gang. Første bind udkom i 1898 (første hæfte i 1893), og siden er der udkommet 34 bind, det seneste til og med *tojs* i 2005. Op igennem 1970'erne og 80'erne blev redaktionen mere og mere udførlig, detaljeret og langsom – ordbogen er vel næsten “berømt” for bogstav S der begynder med bind 24 i 1965 og strækker sig helt ind i bind 34 der først udkom i 2002. I midten af 1980'erne forventedes projektet således ikke færdigt før ca. 2045! I 1993 blev der derfor lavet en revurdering af projektet: Det redaktionelle arbejde blev digitaliseret, der skete en opstramning af redaktionspraksis, og der blev fastsat en dato for arbejdets fuldførelse: 2017. For at nå dette mål er der blevet sat tal på produktivitet og omfang: Hver redaktør skal pr. uge behandle 200 sedler, og 1000 sedler må højst resultere i 8 spalter færdigredigeret tekst.

Anki Mattisson fortalte at de stramninger der blev gennemført i 1990'erne, faktisk ikke indebar en ændring af redaktionsreglerne, men “bare” en tilbagevenden til de redaktionsregler der var blevet fastlagt omkring 1920. Altså en opstramning af redaktionspraksis. Som eksempel gennemgik hun reglerne for hvilke ord der ikke har ret til optagelse i ordbogen, og jeg vil i det følgende gengive de vigtigste.

SAOB beskriver, som nogle vil vide, det svenske skriftsprog helt tilbage til 1521. Ordbogen er altså i allerhøjeste grad en historisk ordbog, og beskrivelsen af det ældre sprog er meget vigtig. Det betyder dog ikke at alle ord medtages ukritisk. Man følger således den hovedregel at et ord skal være belagt i mindst to uafhængige kilder for at kunne blive medtaget i ordbogen. Denne hovedregel kan i princippet kun fraviges ved ord fra 15-1600-tallet der også findes i Söderwalls ordbog, men reglen er i årenes løb også blevet fraveget i andre tilfælde. Det sker dog ikke længere, undtagen hvis man fx er kommet for skade at gøre et ord der kun findes i ét belæg til opslagsord i form af en henvisning i et tidligere bind (under *gud* har man fx henvist til *torg-gud* som kun findes i et enkelt belæg fra 1833). Et ord der kun er belagt i en anden ordbog eller ordliste medtages heller ikke.

SAOB's genstand er almene og etablerede ord og betydninger inden for skriftsproget og rigsmålet. Proprier medtages i reglen ikke, og dialekt og slang udelades også. Stof der har berøring med det folkløstiske og dialektale, som fx lege, ordsprog og talemåder, undgås også. Til gengæld udgør det fagsproglige stof et område hvor det kan være svært at trække grænsen for hvad der skal med. Dette er et velkendt fænomen inden for historisk leksikografi, og også SAOB har i tidens løb indoptaget en hoben dyre- og plantnavne samt tekniske udtryk og militæruddryk (sidstnævnte regnes vist for SAOB's “speciale”!). I alt regner man med at forskelligt fagstof fylder 15 % af ordbogen. Som eksempel gav Anki Mattisson ordet *rikt-skiva* (bind 22, spalte R 1953) der er medtaget i fem fagsproglige betydninger og fylder $\frac{3}{4}$ spalte. Artikler af denne type er nu forbudt, og antallet af fagudtryk som medtages (fx udenlandske dyrenavne) er langt færre, bl.a. ud fra den betragtning at de ofte allerede findes beskrevet i fagordbøger. Encyklopædiske oplysninger søges også begrænset.

Der er fokus på det historiske i SAOB. Derimod er det nyeste sprog, dvs. fra ca. 1950 og frem, ikke højt prioriteret. *Nyordboken* (2000) fra det svenske sprognævn (Språkrådet, tidl. Språknämnden) er excerperet til samlingerne, og man søger også i *Presstext*, en database over moderne avistekster (svarende til Infomedia). Men på spørgsmålet om man i realiteten beskriver det svenske sprog helt op til udgivelsesåret for hvert bind, er svaret tøvende. Det er og bliver den solide og detaljerede beskrivelse af det svenske skriftsprog i et længere historisk perspektiv der er i centrum, og fx er det ikke tilladt at medtage en sprogpåvisning der kun findes publiceret på internettet.

Siden 1997 har SAOB imidlertid selv været på nettet, og det gennemsnitlige antal af besøg på en dag ligger på mellem 4000 og 5000. At så mange slår op i ordbogen på nettet, er en yderligere grund til at overveje hvordan man kunne supplere ordbogen, hvis første bind jo som nævnt så dagens lys for over 100 år siden. Det ligger uden for den nuværende redaktions rammer at lave en supplering, men det er indlysende at det i givet fald skal ske som en netversion.

Birgit Eaker fortalte derefter om behandlingen af flerordsforbindelser i ordbogen. Disse er i princippet opdelt i tre hovedtyper: *uttryck*, *förbindelse* og *särskilt förbindelse*. *Uttryck* er et meget bredt begreb og omfatter både fraser og talemåder, dvs. alle forbindelser af betydningsbærende ord der på en eller anden måde har en “indre sammenhæng”. *Förbindelse* dækker over ordkombinationer med en bestemt grammatisk relation mellem de enkelte dele, mens *särskilt förbindelse* specifikt gælder partikelverber (fx *ställa in*).

Udtryk kan godt blive opslagsord, anført på alfabetisk plads efter bogstavprincippet. Det sker især hvis det første ord i udtrykket ikke findes som simpleks (fx *snick och snack*, *fata morgana*), eller hvis der findes sammensætninger til udtrykket (fx *dag och natt-frö*, *-gräs* til *dag och natt*). I disse tilfælde får udtrykket også en etymologi. Men det normale er at et udtryk behandles under et af de simpleksord der indgår i det. Af eksempler fra artiklen *tid* kan nævnes *under tiden* under betydning 1 g, *över tid* under betydning 1 h og *tidens melodi* under betydning 4 b β. I SAOB anbringes udtryk altså under den betydning der anses for den relevante, ikke samlet til sidst i artiklen, som fx i Den Danske Ordbog.

Efter foredragene blev vi delt op i hold. Vi fik en demonstration af de elektroniske ressourcer, først og fremmest det nye redigeringsystem – og så blev vi lukket ind i ordbogens hjerte: arkivet! En tung dør lukkede sig, og så stod vi midt i en samling af omhyggeligt ordnede og sirligt skrevne sedler på 8 mio. stk. Hvilket er ganske meget – seddelsamlingen til ODS er fx på 2,5 mio. sedler. Eftersom SAOB begyndte at udkomme helt tilbage i 1893, består arkivet også af en ganske omfattende accessions-samling.

Redaktionsprocessen er delt op i tre faser. Først en forredaktion hvor sedlerne til et ord findes frem, sorteres og behandles med hensyn til især de formelle oplysninger om udtale, bøjning osv. i artiklens hoved. Antallet af belæg, ældste og yngste belæg og citategnede eksempler udvælges også. Herefter går seddelkassen videre til den egentlige redaktør, som gennemgår oplægget og giver artiklen sin endelige udformning og struktur. Til slut skrives de udvalgte eksempler ind af en særlig medarbejder, og der læses korrektur osv.

Selve publiceringsprocessen er indtil videre foregået i editeringsprogrammet FrameMaker, men SAOB har i nogle år arbejdet på at udvikle sit eget database-baserede programmel, der kan sikre en umiddelbar og samtidig publicering, både på nettet og på tryk. Dette udviklingsarbejde er i sin afgørende fase, og vi fik demonstreret hvordan systemet fungerer og bl.a. arbejder sammen med databasen for kildeforkortelser. SAOB er med andre ord godt på vej ind i den digitale tidsalder, men vil dog også til det sidste blive udgivet på tryk – og tak for det!

Man kan i øvrigt læse mere om SAOB på www.saob.se.

Ny dansk disputats med forslag til en generel metaleksikografisk teori

Af Henning Bergenholtz

Sven Tarp, Center for Leksikografi, Handelshøjskolen i Århus, har fredag den 6. oktober 2006 forsvaret sin doktorafhandling med titlen: *Leksikografien i grænselandet mellem viden og ikke-viden - Generel leksikografisk teori med særlig henblik på lørnerleksikografi.*

Udvalget til bedømmelse af afhandlingen bestod af:

Universitätsdozent, dr. phil. et habil. Ulrich Heid, Universität Stuttgart

Professor, dr. phil. John Humbley, Université Paris 7

Professor, dr. phil. et habil. Henning Bergenholtz, Handelshøjskolen i Århus

Fra dette udvalg var John Humbley og Ulrich Heid de to officielle opponenter. En kort beskrivelse af indholdet af doktorafhandlingen, udarbejdet af Handelshøjskolens kommunikationsafdeling, følger her:

Ny doktorafhandling om ordbogen i videnssamfundet

Behovet for hurtig og effektiv adgang til viden er større end nogensinde før. Ny doktorafhandling fremlægger teori for ordbogen som redskab til vidensstilegnelse. Afhandlingen forsvares fredag den 6. oktober.

Doktorafhandlingen *Leksikografi i grænselandet mellem viden og ikke-viden* skrevet af lektor Sven Tarp (Institut for Sprog og Erhvervs kommunikation) præsenterer en sammenhængende teori for ordbøger, der kan være udgangspunkt for udvikling af helt nye leksikografiske redskaber.

Ordbøger har eksisteret i flere tusind år, men deres karakter har forandret sig i takt med samfundsudviklingen. Da de første ordbøger blev skrevet på lertavler for 4000 år siden, var det som led i fortolkningen af religiøse skrifter. I dag er ordbøger i stigende grad elektroniske og internetbaserede og spiller en vigtig rolle som redskaber til vidensstilegnelse. Samtidig betyder internationaliseringen, at der i fremtiden vil være et endnu større behov for ordbøger, som kan lette kommunikationen på tværs af grænserne.

Sammenhængende teori

Det centrale i afhandlingen har været at udvikle en generel leksikografisk teori. Tidligere forskningsprojekter inden for området har fokuseret på selve ordbogen og brugen af den. Sven Tarp flytter i stedet fokus uden for ordbogen og undersøger de forskellige typer potentielle brugere og de situationer, hvor de kan have behov, som kan dækkes ved ordbogsopslag. På det grundlag formulerer han en teori for de forskellige kommunikative og kognitive funktioner, som ordbøger kan have.

– Hvis man ser bort fra alle underordnede træk, er kernen i ordbøger deres evne til at dække brugernes behov for informationer om stort set hvad som helst. Netop behovet for hurtig adgang til informationer er i dag helt centralt. Det gælder både, når der er problemer i kommunikationsprocessen, og når der findes et ønske om at få større viden om et bestemt emne. Den teori, som er udviklet i afhandlingen, kan bruges som vejledning til udvikling af helt nye typer af leksikografiske redskaber, ikke blot trykte ordbøger, men også elektroniske opslagsværker og specialiserede søgeredskaber, forklarer Sven Tarp.

Fokus på brugeren

Anden del af afhandlingen fokuserer på lørnerordbøger, der er beregnet på brugere, som er i gang med at lære et fremmedsprog. Lørner-begrebet stammer fra det engelske learner. Lørnerne kan f.eks. være danskere, der er ved at lære engelsk eller spansk, eller udlændinge, som ønsker at lære dansk.

– Man har forskellige behov alt efter, hvor i læringsprocessen man befinder sig. Du har behov for én slags læringsredskaber, når du starter med at lære for eksempel engelsk på aftenskolen, hvor det primært handler om at forstå en tekst. Du vil derimod have andre behov, hvis du skal skrive en juridisk kontrakt på engelsk. Det er derfor også inden for leksikografien uhyre vigtigt at tage højde for målgruppen, når man starter et ordbogsprojekt, forklarer Sven Tarp.

Som et led i arbejdet med afhandlingen har Sven Tarp været på et halvt års forskerophold i Sydafrika. Ud over det videnskabelige arbejde fik han her afprøvet nogle af teorierne på egen krop:

– Jeg har selv befundet mig i rollen som lørner, da jeg besluttede mig for at lære afrikaans. Det var særdeles lærerigt at få både be- og afkræftet nogle af mine teorier om lørnerrollen.

Fremtidens ordbøger

Der er ingen tvivl om, at brugen af ordbøger har ændret sig med tiden. Sven Tarp forklarer, at ideen til doktorafhandlingen udspringer af en mangel på nye vidensredskaber, der i højere grad tager højde for de ændrede behov i samfundet:

– Det handler i bund og grund om produktudvikling. Mere end halvdelen af alle ordbogsprojekter bliver enten ikke færdiggjorte eller bliver forsinkede. Hvis man lavede en række guidelines på baggrund af denne teori, ville man ikke alene kunne højne kvaliteten og effektivisere arbejdsprocessen, man ville også kunne reducere udgifterne, fortæller Sven Tarp.

Der har allerede været stor international interesse for doktorafhandlingen. Sven Tarp er blevet inviteret til at afholde kurser om sine teorier i både Brasilien og Sydafrika.

Festskrift til Pia Jarvad

Af Jørgen Nørby Jensen

Den 20. juli 2006 fyldte seniorforsker, cand.mag. *Pia Jarvad* 60 år. Hun har været ansat i Dansk Sprognævn siden 1974.

Pia er medgrundlægger af LEDA. Hun var i flere år medlem af foreningens bestyrelse og desuden formand i perioden 1994-1998. Leksikografien har i det hele taget fyldt meget i Pias mangeårige virke: Hun var involveret i arbejdet med *Retskrivningsordbogen* fra 1986, hun står bag den populære *Politikens Lille Danske Ordbog* (3. udgave, 3. oplag, 2006), og for nylig har hun lagt sidste hånd på en helt nyskreven udgave af ordbogsklassikeren *Bevingede Ord*, der udkom første gang i 1940.

Bedst kendt er Pia Jarvad for gigantværket *Nye ord. Ordbog over nye ord i dansk 1955-1998* (Gyldendal, 1999).



Bogen indeholder over 10.000 ord, vendinger og forkortelser som er blevet almindelige i sidste halvdel af det 20. århundrede: fra *kernefamilie*, *sølvbryllups kvarter* og *rock and roll* over *bistandshjælp*, *weekendfar* og *aerobic* til *EU-borger*, *downloade* og *smiley*. I *Nye Ord* er der oplysninger om ordenes betydning, brug, oprindelse og tidligst registrerede forekomst, og man tager næppe munden for fuld ved at hævde at bogen har haft en ikke ringe betydning for ordudvalget i nyere ordbøger som fx *Den Danske Ordbog* (2003-2005), *Politikens Nudansk Ordbog med etymologi* (3. udg., 2005) og *Retskrivningsordbogen* (3. udg., 2001).

I sit arbejde med at indsamle og kategorisere de mange nye ord har Pia også forsket i principperne for orddannelse. Om dette kan man læse i bogen *Nye ord – Hvorfor og hvordan?* (Gyldendal, 1995) og i en lang række artikler.

Inden for en overskuelig fremtid vil en elektronisk version af *Ordbog over nye ord i dansk* blive lagt på Dansk Sprognævns hjemmeside (www.dsn.dk), og herefter vil den løbende blive opdateret med nye ord fra perioden efter 1998 – til glæde for leksikografer og alle andre med interesse for ordtilvæksten i det danske sprog.

Som en hyldest til Pia Jarvad har Dansk Sprognævn udsendt et festskrift med titlen *Ordet fanger. Festskrift til Pia Jarvad i anledning af 60-års-dagen*. Det indeholder følgende artikler:

Niels Davidsen-Nielsen: Hvad betyder *henholdsvis*?

Henrik Galberg Jacobsen: Gamle ord i nye klæder. Lille ortografisk ændringsordbog 1775-2001

Asgerd Gudiksen: Om *halv-* og *hel-* og halvaffikser

Arne Hamburger: Stumt *h* foran *v* i starten af et ord eller en stavelse

Erik Hansen: Kylling *sb.* (ny brug)

Tore Kristiansen og Jacob Thøgersen: Rigtige mænd er ikke bange for engelsk – om konstruktion af sprogholdninger og kønsidentitet

Jørn Lund: Nye ord. Gennemtræk og tilvækst
Ida Elisabeth Mørch: Bestemthed og nøgen form
Jørgen Nørby Jensen: 60 ord på 60-års-dagen
Marianne Rathje: Hvordan udtales *weekend* i dag?
Ole Ravnholt: Passivens subjekter
Jørgen Schack: Pseudohomonymer

Skriftet er redigeret af Jørgen Nørby Jensen, Ole Ravnholt og Jørgen Schack. Det er på 203 sider og kan købes hos Dansk Sprognævn (tlf. 35 32 89 82 eller e-post: hanne.broch@dsn.dk). Prisen er 95 kr. + porto.

Årsberetning for LEDA for året 2005-06

Af Henrik Lorentzen

De hidtidige beretninger følger kalenderåret, men holdes ved generalforsamlingen sidst i april. Den praksis giver en del forsinkelse, så derfor har formanden i samråd med bestyrelsen besluttet at beretningen bør føres op til det tidspunkt hvor den gives. Det følgende er altså en beretning for 2005 til -06.

Medlemsmøder

Der er blevet holdt 4 medlemsmøder i perioden.

Det første blev holdt den 18. april 2005 på Peter Skautrup Centret for Jysk Dialektforskning ved Aarhus Universitet. Det overordnede tema for mødet var Leksikografiske miljøer og projekter vest for Storebælt. Der var en række foredrag, det første ved Ruth Feil, Termplus ApS, og Lotte Weilgaard fra Institut for Fagsprog, Kommunikation og Informationsvidenskab ved Syddansk Universitet i Kolding. Det handlede om NorNa-projektet, dets begrebsindeks og søgegrænseflade. Birthe Gawinski fra Syddansk Universitet i Esbjerg fortalte om sin illustrerede online-ordbog om spansk mad, og Sandro Nielsen fra Center for Leksikografi ved Handelshøjskolen i Århus talte om udarbejdelse af regnskabsordbøger til internettet.

Efter kaffepausen var der foredrag om Jysk Ordbog og rundvisning ved Viggo Sørensen og Inger Schoonderbeek Hansen. Efter foredragene blev der holdt generalforsamling efterfulgt af en hyggelig spising, der i dagens anledning blev betalt af foreningen.

Det andet medlemsmøde blev holdt den 10. november 2005 på Center for Tegnsprog og Tegnstøttet Kommunikation i København. Her fik vi af Jette H. Kristoffersen, Janne Boye Niemelä og Thomas Troelsgård præsenteret projektet Ordbog over Dansk Tegnsprog, og det var vist i meget bogstavelig forstand en øjenåbner for mange af os, idet det ud over at være en introduktion til denne særlige type ordbøger også blev en mini-indføring i tegnsprog.

Det tredje medlemsmøde fandt sted den 27. februar 2006 på Københavns Universitet, Amager, og handlede om den store tysk-danske ordbog fra Gyldendal, der udkom ved årsskiftet. Bestyrelsen havde inviteret to af redaktørerne, nemlig Henrik Bergstrøm-Nielsen fra forfattersiden og Torben Christiansen fra forlagssiden, som begge fortalte om særlige aspekter ved tilblivelsen af ordbogen. Vibeke Winge leverede det kritiske modspil, som også kan læses i hendes anmeldelse i LEDA-Nyt nr. 41. Det gælder i øvrigt også for de andre foredrag jeg har nævnt, at de findes i skriftlig form i LEDA-Nyt.

Det fjerde medlemsmøde blev holdt i forbindelse med generalforsamlingen den 26. april 2006 på Københavns Universitet, Amager. Bolette Sandford Pedersen fra Center for Sprogteknologi og Jørg Asmussen fra Det Danske Sprog- og Litteraturselskab berettede om projektet DanNet, et leksikalsk-semantisk ordnet for dansk, hvor man udnytter resurser og viden fra Den Sprogteknologiske Ordbase, SIMPLE og Den Danske Ordbog.

Ved alle møderne var der pænt fremmøde og gode diskussioner. Som noget nyt forsøgte bestyrelsen sig med mødeindkaldelse pr. e-mail, og det så ud til at fungere så godt at vi vil gøre det igen i de tilfælde hvor vi ikke kan nå at annoncere i LEDA-Nyt. Hvis nogen medlemmer har modtaget mødeindkaldelse med posten og hellere vil have den pr. mail, bedes de henvende sig til kassereren.

Bestyrelsen har planlagt to møder efter sommerferien som vi allerede har sat dato på. Det første er et besøg på Svenska Akademiens ordbok i Lund, og det bliver den 20. september. Vi regner med at ankomme ved middagstid, spise frokost sammen med nogle fra redaktionen og derefter få en introduktion til projektet, blive vist rundt, se hvordan de arbejder osv. Sidst på eftermiddagen vil vi så vende hjem til Danmark.

Det andet møde er planlagt til den 22. november, hvor Henrik Køhler Simonsen vil fortælle om projektet ZooLex, et leksikografisk værktøj der er udviklet til Zoologisk Have i København. Man kan altså godt sætte kryds i kalenderen ved disse to datoer.

Bestyrelsesmøder

Bestyrelsen har i perioden holdt 7 bestyrelsesmøder og derudover haft regelmæssig mailkontakt. Det har været gode og konstruktive møder, hvor foreningens aktiviteter er blevet diskuteret og planlagt.

LEDA-Nyt

LEDA-Nyt er i perioden udkommet tre gange, nr. 39, 40 og 41, alle tre ret omfangsrige numre, bl.a. fordi de indeholder skriftlige versioner af de foredrag der er blevet holdt ved medlemsmøderne. Derudover indeholder de anmeldelser, indkaldelser til medlemsmøder, orientering om konferencer og senest også underholdning. Det redaktionelle arbejde varetages især af Inger Schoonderbeek Hansen og Liisa Theilgaard. Jeg vil gerne sige tak til dem og til bidragerne.

Hjemmeside

I år har arbejdet med hjemmesiden været koncentreret om at få lagt ældre numre af LEDA-Nyt ud (de nyere lå der allerede), og den første fase, som går ud på at scanne dem ind, er afsluttet. De ligger nu på hjemmesiden som pdf-filer helt tilbage til det første nummer fra februar 1989, skrevet af Ole Norling-Christensen med assistance af Henrik Holmboe. En af nyhederne fra dette nummer inden for dataleksikografi er at Dantes La Divina Commedia er udkommet – på diskette! Vi håber på at få etableret en søgefunktion på et senere tidspunkt.

Hjemmesiden ligger fortsat hos CST på www.cst.dk/leda og bestyres af Sanni Nimb.

Generalforsamling

Der blev afholdt generalforsamling den 18. april 2005. Referatet kan læses i LEDA-Nyt nr. 40. Det blev her nævnt at det var en god idé med møder uden for København, og det er da også planen at tage til Lund i september.

Medlemmer

Antallet af medlemmer ligger på 106 (hvoraf 93 har betalt kontingent for 2005). Det er nogenlunde stabilt i forhold til sidste år, men udgør en tilbagegang i forhold til årene før. Dette skyldes nok især at bestyrelsen har besluttet sig for at slette gamle restanter af medlemslisten.

Videnssamfundet

LEDA er blevet kontaktet af gruppen bag hjemmesiden videnssamfundet.dk, som vil registrere op mod 600 danske videnskabelige selskaber og foreninger. Formanden har derfor svaret på to skemaer med oplysninger om LEDA's formål, aktiviteter m.v. Hvad der kommer ud af det, er for tidligt at sige endnu.

Konferencer

Nordisk Forening for Leksikografi, NFL, holdt sin 8. konference, som fandt sted i Sønderborg i maj 2005 med mange deltagere fra LEDA's medlemskreds. Der kan læses beretninger fra konferencen i LEDA-Nyt nr. 40 og i LexicoNordica nr. 12. Den næste nordiske konference finder sted i Island i 2007 (se annonce andetsteds i dette nummer).

Referat af generalforsamling i foreningen Leksikografer i Danmark

Af Liisa Theilgaard

I alt 17 personer deltog i generalforsamlingen, der fandt sted 26. april 2006 på Københavns Universitet Amager.

Dagsorden:

1. Valg af dirigent og referent.
Lars Trap-Jensen (Det Danske Sprog- og Litteraturselskab) blev valgt til dirigent.
Liisa Theilgaard (Gyldendal) blev valgt til referent.
2. Formanden aflægger beretning.
Formand Henrik Lorentzen (Det Danske Sprog- og Litteraturselskab) aflagde beretning (se s. 34).
3. Kassereren fremlægger det reviderede regnskab.
Kasserer Christian Becker-Christensen (Politikens Forlag) fremlagde det reviderede regnskab, som blev godkendt.
4. Fastsættelse af kontingent.
Kontingentet forbliver uændret 125 kr. om året.
5. Valg af formand.
På valg var Henrik Lorentzen, som blev genvalgt.
6. Valg af bestyrelsesmedlemmer og suppleant.
På valg var Inger Schoonderbeek Hansen (Jysk Ordbog), Sanni Nimb (Center for Sprogteknologi) og Liisa Theilgaard, som alle tre blev genvalgt. Christian Becker-

Christensen, Henrik Hovmark (Ømålsordbogen) og Jørgen Nørby Jensen (Dansk Sprognævn) var ikke på valg og sidder stadig i bestyrelsen.

7. Valg af revisor og revisorsuppleant.

På valg var revisor Ebba Hjorth (Det Danske Sprog- og Litteraturselskab) og revisorsuppleant Pernille Neergaard-Dibbern (Gyldendal). Begge blev genvalgt.

8. Forslag.

Der var ikke indkommet nogen forslag.

9. Eventuelt.

Sanni Nimb foreslog at forlænge fristen for indsendelse af forslag til generalforsamlingen fra 15. januar.

Det blev foreslået, at der gøres mere for at annoncere LEDA's arrangementer tidligere og til en bredere kreds, fx på lingvistik, i Humanist (deadline længe før udgivelse) eller Universitetsavisen (koster penge).

Det blev foreslået at spare portoudgifter til udsendelse af LEDA-Nyt ved i stedet at maile bladet til medlemmerne. Man kunne også samle bladene og kun sende til én kontaktperson på hver institution, som så fordelte bladene videre. Hertil indvendte kassereren, at det er forskelligt, om man modtager bladet på sin arbejdsplads eller derhjemme.

Carsten Hansen, Lexitext, tilbød derefter generøst, at Lexitext betaler portoen for udsendelse af LEDA-Nyt det næste år, maks. 1.500 kr. Det blev modtaget med stor tak.

Ud over kreditering foreslog Anna Braasch, at sponsorer kunne vedlægge reklame.

Bestyrelsen arbejder videre med disse forslag.

Endelig blev vinderen af Lexi-krydsen i LEDA-Nyt nr. 41 udtrukket. Vinder blev Simon Skovgaard (Det Danske Sprog- og Litteraturselskab), og præmien var et eksemplar af Simon Winchesters *The Meaning of Everything* om OED's historie. Tak til Lars Trap-Jensen for krydsogtværsen. Bestyrelsen modtager gerne ideer til sjove opgaver til bladet.

9. Konference om Leksikografi i Norden

Akureyri, Island, 22.-26. maj 2007

NORDISK FORENING FOR LEKSIKOGRAFI

Stofnun Árna Magnússonar í íslenskum fræðum, Sprákrádet og organisationskomitéen har den store glæde at invitere til den 9. Konference om Leksikografi i Norden. Det er anden gang Island er vært for konferencen, og arrangørerne har derfor – inspireret af det danske eksempel fra den seneste konference – besluttet at placere konferencen i den nordlige del af landet. Valget er faldet på den gamle handelsplads Akureyri ved fjorden Eyjafjörður, nu en ung universitetsby med en god lufthavn. Konferencestedet bliver Hótel KEA, der ligger i byens hjerte ved foden af byens kendte kirke og med mulighed for at deltagerne kan bo på stedet. Akureyri er en by med knap 16.000 indbyggere, tæt på konferencelokalene ligger det dejlige udendørs termalbad, og der vil således være mulighed for friluftsbadning inden programstart om morgenen.

Fagligt program

I overensstemmelse med traditionen for de tidligere nordiske konferencer i leksikografi vil vi ikke pege på bestemte temaer, men opfordrer deltagerne til at tilmelde foredrag inden for de områder de interesserer sig for og arbejder med til daglig, således at programmet kommer til at give et bredt billede af leksikografien i Norden lige nu. Programkomitéen vil forbeholde sig ret til at udvælge enkelte foredrag til plenumforelæsninger, men regner i øvrigt med at foredragene skal vare 30 minutter inklusive spørgsmål og diskussion. Afhængigt af tilmeldingen vil der blive to parallelle sessioner.

Socialt program

En af dagene vil der blive en udflugt fra ca. kl. 15 til kl. ca. 22, en tur der vil vise deltagerne nogle af fjordens naturmæssige og kulturelle seværdigheder. Konference-middagen holdes fredag aften.

Tid og sted

Deltagerne bør ankomme om eftermiddagen/aftenen den 22. maj, hvorefter der er registrering og velkomstarrangement. Generalforsamling i Nordisk Forening for Leksikografi afholdes under konferencen. Vi gør opmærksom på at der ikke vil være program lørdag morgen, men at der er mulighed for at skaffe lokaler til interne workshops og andre faglige møder, hvis deltagerne ønsker dette.

Nærmere oplysninger på NFL's hjemmeside <http://www.nordisk-sprakrad.no/nfl.htm> og halldora@lexis.hi.is.

Tidsfrister

Frist for tilmelding med foredrag, workshop eller softwaredemonstration: 1. november 2006.

Tilbage melding om accepteret foredrag: 1. december 2006.

Sjáumst á Akureyri í maí!
Organisationskomitéen