

LEDA-Nyt nr. 39 – Marts 2005

ISSN 0908-355

Redigeret af Inger Schoonderbeek Hansen og Liisa Theilgaard

Indhold

Meddelelser fra bestyrelsen: Indkaldelse til generalforsamling	s. 2
Indkaldelse til medlemsmøde	s. 3
Handbok i Lexikografi	s. 5
Fraseologi i Danmark: En status	s. 8
Den SprogTeknologiske Ordbase	s. 13
Den Danske Ordbog – nok engang	s. 20
Dansk-Hollandsk Ordbog og Hollandsk-Dansk Ordbog	s. 23
Dansk-Hollandsk Ordbog: Udgangspunkter og målsætninger	s. 31
En polyfon ordbogsrecension	s. 34
Kommentar fra forlaget: Lidt om produktionen	s. 39
8. Konference om Leksikografi i Norden	s. 41

Medlemsmøde og generalforsamling **mandag den 18. april 2005 kl. 13.00**

Leksikografiske miljøer og projekter **vest for Storebælt**

Præsentation af NorNa-projektet (Ruth Feil & Lotte Weilgaard, SDU Kolding),
Spansk mad – en illustreret onlineordbog (Birthe Gawinski, IT-Vest),
Dansk-Engelsk Regnskabsordbog (Sandro Nielsen, HHÅ) og
Jysk Ordbog (Inger S. Hansen & Viggo Sørensen, AU)

Sted: Aarhus Universitet, Niels Juels Gade 84, "Trøjborgkomplekset", lokale 344
Efter generalforsamlingen byder LEDA på middag!

se side 3

Meddelelser fra bestyrelsen

Indkaldelse til generalforsamling

Mandag den 18. april 2005 kl. 16.45

Aarhus Universitet, Niels Juels Gade 84, "Trøjborgkomplekset", lokale 344.

Dagsorden ifølge vedtægterne:

1. Valg af dirigent og referent
2. Formanden aflægger beretning
3. Kassereren fremlægger det reviderede regnskab
4. Fastsættelse af kontingent
5. Valg af bestyrelsesmedlemmer.
På valg er:
Christian Becker-Christensen
Henrik Hovmark
Jørgen Nørby Jensen
6. Valg af revisor og revisorsuppleant.
På valg er:
Ebba Hjorth
7. Forslag
8. Eventuelt

Bestyrelsen opfordrer medlemmer fra alle dele af landet til at stille op til bestyrelsen.

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 2005. 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 (Pol. Forlag) (kasserer), Christian.Becker-Christensen@pol.dk
Liisa Theilgaard (Gyldendal), Liisa.Theilgaard@gyldendal.dk
Sanni Nimb (KUA, CST), sanni@cst.dk
Jørgen Nørby Jensen, jnj@firkant.net
Inger Schoonderbeek Hansen (AU, Jysk Ordbog), jysis@hum.au.dk

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

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

Henvendelse vedrørende indmeldelse og adresseændring sker til foreningens kasserer.

Øvrige henvendelser rettes til formanden.

Medlemsmøde

Leksikografiske miljøer og projekter vest for Storebælt

Mandag den 18. april 2005 kl. 13.00

**Aarhus Universitet, Niels Juels Gade 84, "Trøjborgkomplekset", lokale 344
Medlemsmødet efterfølges af generalforsamling og middag**

13.00-13.05 Velkomst

13.05-14.00 NorNa – Begrebsindeks og søgegrænseflade

v/ Ruth Feil og Lotte Weilgaard (Institut for Fagsprog, Kommunikation og Informationsvidenskab, Syddansk Universitet)

Vi vil gerne præsentere resultaterne fra det nordiske netværkssamarbejde NorNa (Nordic Navigator) om udvikling og testning af en prototype for et flersproget, nordisk søgeinterface, der i samspil med eksisterende søgemaskiner giver monolingval adgang til informationer på flere nordiske sprog samtidig. De behandlede sprog er norsk, finsk, svensk og dansk samt engelsk; det sidste åbner mulighed for senere udbygning med ikke-nordiske sprog. Projektet tager udgangspunkt i kendte terminologiske principper. Netværksprojektet har modtaget økonomisk støtte fra Nordisk Ministerråds Sprogteknologiske forskningsprogram. Desuden vil vi kort fortælle om projektet "Studerende som brugere af TERMbank SYD".

Spansk mad – en illustreret onlineordbog

v/ Birthe Gawinski (specialestuderende på masterstudiet i it, sprog og læring, IT-Vest)

Formålet med ordbogen er at hjælpe ordbogsbrugeren med at få en bedre fornemmelse af, hvad et givet opslagsord dækker over, især i de tilfælde, hvor der ikke findes en ækvivalent. Der findes bl.a. definitioner og/eller forklaringer på både kilde- og målsproget samt illustrationer og links til relevante websteder, informationer som man sjældent vil finde i papirudgaver af gængse bilingvale ordbøger.

Jeg opfatter ikke min målgruppe som en entydig størrelse, men ser den mere som målgrupper på forskellige niveauer eller med forskellige behov. Hypertekststrukturen i onlineordbogen giver gode muligheder for at lade brugeren selv beslutte, hvilke informationer og hvor mange han/hun har brug for i den givne situation.

14.00-15.00 Dansk-Engelsk Regnskabsordbog

v/Sandro Nielsen (Center for Leksikografi, Handelshøjskolen i Århus)

I de senere år har Center for Leksikografi arbejdet med og udgivet to dansk-engelske regnskabsordbøger, en internetordbog og en trykt ordbog. Ordbøgerne er udarbejdet efter de nyeste principper inden for leksikografien, hvilket bl.a. betyder, at udgangspunktet er rent leksikografisk og ikke lingvistisk. Den nye leksikografiske definition af en ordbog fokuserer på tre elementer: leksikografiske funktioner, data og strukturer. Disse skal relateres til den påtænkte brugergruppe, brugernes lingvistiske og faglige kompetence samt deres behov i bestemte brugersituationer. Ordbogens funktioner er oversættelse, produktion, reception og revision af tekster samt videns tilegnelse, og brugerne kan inddeles i eksperter, semi-eksperter og lægfolk. De mange muligheder, det elektroniske medium giver, er udnyttet, og forskellige datatyper kombineres med brugernes behov i funktionsrelaterede situationer, fx: definitioner, grammatiske oplysninger, kollokationer, fraser, henvisninger, synonymer og antonymer. Både de faglige og lingvistiske data hjælper brugeren i kommunikationsorienterede situationer inden for en registerspecifik kontekst som regnskabsvæsen.

15.00-15.20 Pause hvor der serveres kaffe/te og kage

15.20-16.30 Jysk Ordbog (www.jyskordbog.dk)

v/Inger Schoonderbeek Hansen & Viggo Sørensen (Peter Skautrup Centret for Jysk Dialektforskning, Aarhus Universitet)

Vores oplæg til LEDA's medlemsmøde vil falde i to dele. Først besøger vi Peter Skautrup Centret, hvor Viggo Sørensen kort fortæller om historien bag Jysk Ordbog samt om dialektsamlingerne der ligger til grund for ordbogen. Derefter fremviser Inger Schoonderbeek Hansen Jysk Ordbog på internettet og kommer ind på bl.a. fordelene ved netpublicering af en omfattende ordbog som Jysk Ordbog.

Alle oplæg varer 45 minutter inkl. efterfølgende diskussion.

Gæster til foredragene er meget velkomne. Vel mødt!

16.45 Generalforsamling (se side 2)

17.30 Middag

LEDA inviterer deltagerne til middag efter generalforsamlingen ca. kl. 17.30! Middagen forventes at foregå i samme bygning. Drikkevarer er for egen regning. **Tilmelding er påkrævet** og skal ske til Inger Schoonderbeek Hansen (e-post jysis@hum.au.dk eller tlf. 89 42 19 15) inden 4. april.

Sted: Aarhus Universitet, Niels Juels Gade 84, "Trøjborgkomplekset", lokale 344.

Transport

Tog

Lyntog (uden ekstra omkostninger) kører hver time .56 fra Københavns Hovedbanegård. Rejsen varer 2.48 time (fx Kbh. 08.56 – Odense 10.13 – Århus 11.44). IC3-tog kører hver hele time, og her varer rejsen 3.08 time (fx Kbh. 09.00 – Odense 10.30 – Århus 12.08). Fra kl. 20.00 kører der tog fra Århus mod København hver time .02 (fx Århus 20.02 – Kbh. 23.18).

Se: www.dsb.dk

Busforbindelser i Århus

Bus nummer 7 holder på Banegårdspladsen (samme side som banegårdens hovedindgang) og kører 5 gange i timen (22 34 46 58 10). Stoppestedet er på Aldersrovej og hedder Trøjborg-centret eller Aarhus Universitet; turen varer 11 minutter. Efter kl. 18.26 kører der kun 2 busser i timen mod banegården (42 12).

Se: www.sporveje.dk

Bil

Motorvej E45 mod Århus. Afkørsel Århus S. Tag Ydre Ringvej til venstre (der på det stykke hedder Viby Ringvej, O2) til Randersvej. Kør her ikke mod Randers, men til højre ned mod byen (Centrum), og fortsæt til ring 1 (der hedder Ringgade, O1). Drej til venstre (Nordre Ringgade). Drej ved 2. trafiklys til højre (Aldersrovej). Niels Juels Gade er 2. sidevej til venstre; nummer 84 er hjørnebygningen.

Se: www.krak.dk

Anmeldelse af Bo Svensén: *Handbok i lexikografi*

Ordböcker och ordboksarbete i teori och praktik. Andra, omarbetade och utökade upplagan. Stockholm: Norstedts Akademiska Förlag 2004.

Af Sven Tarp

Indledning

Da jeg i 1989 begyndte at arbejde på min ph.d. i fagleksikografi, var den første bog om leksikografi, som jeg overhovedet læste, Bo Svenséns *Handbok i lexikografi*, der var udkommet to år forinden. Den gav mig en første og nødvendig introduktion til en verden, som jeg på daværende tidspunkt endnu nærmede mig med en vis portion berøringsangst. Jeg mindes, at jeg var glad for denne læsning, selv om jeg senere i stigende grad blev kritisk over for bogen og dens grundlag. Nu hvor jeg netop er blevet frikøbt til at skrive en doktorafhandling om lærerleksikografi, vil skæbnens ironi, at den første bog, jeg kaster mig over, igen er *Handbok i Lexikografi*, der i 2004 kom i en helt ny udgave.

I de 18 år, der er gået, siden Bo Svensén udsendte sin første *Handbok*, er der løbet meget vand i den leksikografiske strand. Det har derfor været spændende at undersøge, i hvor høj grad Bo Svensén har fanget hele denne udvikling i sit stort anlagte værk. Jeg må klart melde ud, at jeg er både imponeret og skuffet. Imponeret, fordi Bo Svensén er en flittig og omhyggelig forfatter, der kommer vidt omkring på bogens godt 600 sider. Skuffet, fordi jeg stadig savner den teoretiske helhedsopfattelse og sammenbindingskraft, som efterhånden fik mig til at være temmelig kritisk over for første udgave.

Indhold

Handbok i Lexikografi er på i alt 624 sider. Den indeholder 28 kapitler samt en bibliografi og et register. Bogen starter med en oversigt over leksikografiens grundlag og vigtigste begreber og fortsætter på den baggrund med en gennemgang af forskellige ordbogstyper. Herefter følger kapitlerne slag i slag i en logisk og velstruktureret rækkefølge, hvor Bo Svensén begynder med dataindsamling og selektion og en kort introduktion til ordbogens komponenter og strukturer, hvorefter han gennemgår en lang række enkeltproblemer såsom lemmaer, ortografi, udtale, morfologi, ordklasse, syntaks, kollokationer og idiommer, definitioner, ækvivalenter, eksempler, encyklopædi, illustrationer, markeringer og etymologi. Derefter følger en række kapitler, hvor ordbogsstrukturer som mikro-, makro-, mega- og henvisningsstrukturer behandles nærmere. Og i den sidste del af håndbogen berører Bo Svensén en række relaterede og centrale spørgsmål såsom planlægning af ordbogsprojekter, juridiske og etiske aspekter, elektroniske ordbøger, ordbogsanvendelse og ordbogskritik.

Som det fremgår, kommer Bo Svensén ind på de centrale og vigtigste spørgsmål i forbindelse med leksikografien. Dertil kommer omfattende henvisninger til relevant litteratur. Ud over de mange konkrete henvisninger i selve teksten afsluttes hvert kapitel med en kort omtale af den vigtigste litteratur inden for det berørte emne. De mange litteraturhenvisninger viser Bo Svensén som en flittig og alsidig forsker, om end man i enkelte tilfælde kan indvende, at han ikke peger på de mest relevante forfattere og problemstillinger. Det gælder f.eks. i kapitlet om encyklopædiske oplysninger i ordbøger, hvor man savner en henvisning til de mange artikler af Henning Bergenholtz, som betvivler relevansen for leksikografien af den lingvistiske skelnen mellem sprog og encyklopædi. Men alt i alt afspejler Bo Svenséns håndbog på trods af dette og andre mindre huller langt hen ad vejen den eksisterende speciallitteratur.

Kritik

Spørgsmålet om relationen mellem leksikografi og lingvistik har i stigende omfang trængt sig på i den leksikografiske debat gennem de seneste 18 år. På den ene side står de teoretikere, som opfatter leksikografi som en deldisciplin af lingvistikken, som “anvendt lingvistik”. På den anden side de teoretikere, som betragter leksikografien som en helt selvstændig disciplin. På side 4 i *Handbok i lexikografi* tager Bo Svensén formelt stilling i denne diskussion: “Somliga betraktar lexikografen som ett slags tillämpad lingvistik, medan andra (däribland författaren till denna bok) anser att den står på egna ben och inte bör ses som en lingvistisk deldisciplin”.

Denne formelle stillingtagen er en forandring og – efter min mening – et frem skridt i forhold til håndbogens første udgave. Problemet er imidlertid, at den ikke slår igennem i håndbogen. De data, der selekteres til ordbøger, behandles i høj grad ud fra en lingvistisk synsvinkel. På side 3 defineres praktisk leksikografi således: “Lexikografi är en verksamhet som går ut på att observera, samla, utvälja, analysera och i en ordbok beskriva ett antal lexikaliska enheter (ord, orddelar och ordförbindelser) tillhörande ett eller flera språk.”

Det, der gør leksikografien til en selvstændig disciplin i forhold til lingvistikken, er, at den som teori har ordbøger og ikke sproget som genstandsområde, og at den som praksis består af opslagsværker, der er udarbejdet med henblik på at bistå ved løsningen af bestemte problemer eller ved besvarelsen af bestemte spørgsmål hos bestemte typer af brugere. Det er således ikke “lexikaliska enheter”, som er leksikografiens omdrejningspunkt, men de spørgsmål og problemer, som de potentielle ordbogsbrugere har eller måtte have.

Bo Svenséns tilbagefald til den “lingvistiske leksikografi” kommer også til udtryk, når han vælger at bruge en “traditionellt inriktad typologi” i modsætning til en typologi, der bygger på ordbøgernes formål og funktion. Han skriver på side 27:

“Olika slag av referensverk och deras förhållanda till språket resp. världen har redan diskuterats i föregående avsnitt. Man kan därvid skilja mellan sådana referensverk som enbart ger information om språket (språkordböcker), sådana som är avsedda att enbart ge upplysningar om världen utanför språket (encyklopedier) och sådana som är avsedda för båda syftena (encyklopediska ordböcker).”

Denne opdeling bygger fuldt ud på en meget omdiskuteret lingvistisk skelnen mellem sproget og verden og tager alene udgangspunkt i karakteren af de data, som er til stede i opslagsværkerne. Sammenholder man dette med kapitel 1.1, hvor forfatteren i overensstemmelse med Hausmann opdeler den praktiske leksikografi i brugsleksikografi og dokumentationsleksikografi, hvor førstnævnte har til formål “att tillgodose individens behov av information”, dvs. til at “understödja kommunikation” og “inläring av språk”, betyder det i praksis, at Bo Svensén hermed udelukker et stort antal fagordbøger, som falder ind under de opslagsværker, som han kalder encyklopædier, fra leksikografiens verden.

Dette er en ærgerlig og helt unødvendig indskrænkning af leksikografiens genstandsområde. Endnu mere ærgerligt er det imidlertid, at denne indfaldsvinkel til leksikografien går igen i hele håndbogen. Dermed bliver de mange interessante data og andre aspekter ved ordbøger, som Bo Svensén behandler, alene anskuet på deres egne præmisser og ikke i lyset af den funktion, de har i ordbøgerne.

Dette gør endvidere, at *Handbok i lexikografi* ikke er videre egnet som hjælp i forbindelse med produktion af ordbøger. Hvornår de enkelte data overhovedet er relevante, og hvordan de i så fald skal præsenteres, sættes kun sjældent i relation til de for-

skellige ordbogsfunktioner. Og det er jo netop dette, som en ordbogsforfatter har brug for at vide.

Måske af samme årsag indeholder håndbogen et kapitel om ordbogsprojekter, hvor Bo Svensén falder tilbage til de værste leksikografiske traditioner, der blandt andet har ført til et stort antal skibbrudne og yderst langtrukne ordbogsprojekter. Forfatteren skelner her mellem en *projektplan* og en *ordbogsplan*: "Projektplanen är i huvudsak en angelägenhet för administratörer, marknadsfolk, ekonomer och jurister, men dessa måste givetvis också samråda med lexikografisk expertis. Ordboksplanen däremot kan i stort sett utarbetas av lexikograferna själva utan medverkan av andra specialistgrupper" (s. 472).

I projektfasen skal man blandt andet "undersöka marknaden och söka identifiera de eventuella behov som finns hos olika målgrupper" (s. 472). Ydermere skal man lave en tidsplan, der indeholder såvel planlægningsfase som implementeringsfase, hvor Bo Svensén med reference til Landau lægger op til, at planlægningsfasen kan udgøre hele 30 procent af den samlede tidsplan. Det interessante er her, at denne fase skal foretages af andre typer eksperter end lige netop leksikografer, som allerhøjst kan tages med på råd, og at det her også forventes, at man har klargjort, "vilken eller vilka ordbogsfunktioner det är fråga om: reception, produktion, översättning L1 ? L2 resp. L2 ? L1" (s. 479). Det vil sige, at det centrale og meget vanskelige spørgsmål om en ordbogsfunktioner ifølge Bo Svensén hovedsagelig skal afgøres af folk, som ikke har nogen uddannelse i leksikografi. Dette kan dårligt betragtes som andet end en invitation til ordbogsprojekter, som ikke er leksikografisk gennemtænkt, og som risikerer at munde ud i produkter af tvivlsom kvalitet, der inden da vil have sprængt såvel budgetter som tidsramme.

Håndbog

Handbok i lexikografi udgiver sig, som titlen siger, for at være en håndbog. Men hvad betyder dette? Og hvem kan have glæde af den? Og hvornår?

Handbok i lexikografi er forsynet med en detaljeret indholdsfortegnelse og et 15 sider langt register. Herigennem kan man finde de steder i bogen, hvor de forskellige problemer behandles. Håndbogen kan på denne måde bruges som opslagsværk, når man søger oplysninger om bestemte problemer.

Men som allerede antydnet, kan man stille spørgsmålstegn ved, om for eksempel kommende ordbogsforfattere kan have nogen gavn af ordbogen. Har de problemer med konkrete data eller strukturer, kan de via indholdsfortegnelse og register gå ind på det sted, hvor disse er behandlet. Men eftersom der ikke er en gennemgående diskussion af, hvordan de forskellige data, strukturer og andre problemstillinger forholder sig til en ordbogs formål, kan man risikere, at en ordbogsforfatter vil kopiere den måde, som håndbogen fremstiller de forskellige ting på i dens mange og i øvrigt meget nyttige eksempler. Det kan nogle gange være udmærket, men i andre tilfælde kan det føre til data, som er medtaget og præsenteret uden for tid og sted. Man må derfor advare mod at bruge håndbogen som hjælp til praktiske ordbogsprojekter.

Selv om en håndbog næppe er lavet til at læse fra ende til anden, kan man (som undertegnede anmelder) vælge at læse den på denne måde. I den forstand kan man helt sikkert opnå en vis indsigt i leksikografiens forunderlige verden. Problemet her er imidlertid, at den leksikografiske teori som allerede nævnt stort set er fortrængt til nogle enkelte specialkapitler, mens hovedparten af håndbogens emner diskuteres løsrevet fra leksikografisk teori og i høj grad relateret til lingvistisk teori trods forfatterens erklæring om, at leksikografien er en selvstændig disciplin. Som introduktion til leksikografisk teori er håndbogen derfor decideret uegnet.

Efter nøje overvejelse er det min opfattelse, at *Handbok i lexikografi* først og fremmest kan benyttes af folk, der allerede har et vist kendskab til leksikografi og kan relativere de mange oplysninger. Den kan i så fald bruges som opslagsbog, hvor man kan søge inspiration til løsning af konkrete problemer eller blive henvist til relevant litteratur, hvis man ønsker at trænge dybere ned i et bestemt spørgsmål.

Fraseologi i Danmark: En status

Af Jan Katlev

Den 23.-24. november 2004 afholdt **Dansk Selskab for Fraseologi (DANFRAS)** på initiativ af sin formand **Ken Farø** et symposium med ovenstående titel på Københavns Universitet med foredragsholdere og deltagere fra Københavns Universitet, Handelshøjskolen i København, Aarhus Universitet, Det Danske Sprog- og Litteraturselskab, Gyldendals ordbogsredaktion, Islands Universitet og Universitetet i Osaka (Japan).

Der blev holdt foredrag med udgangspunkt i (a) dansk fraseologi om (1) fraseologi og tekstlingvistik, (2) indlæring af fraseologi, (3) fraseologi og dialekt, (4) fraseologi og kollokationer, (5) fraseologi og etymologi, (6) fraseologisk terminologi samt med udgangspunkt i (b) kontrastiv fraseologi i dansk og fremmedsprog som (7) islandsk, (8) tysk, (9) engelsk, (10) fransk & italiensk, (11) nygræsk og (12) japansk.

a) Foredrag med udgangspunkt i dansk fraseologi

1) I foredraget *Frasemer set tekstlingvistisk* drøftede **Lisbeth Falster Jakobsen** selve størrelsen et *frasem* defineret som en idiomatisk størrelse

- fx af typen *en torn i øjet* med en sekundær semantik der ikke er simpelt sammensat af semantikken i de enkelte ord
- med semantisk transparens, også ved lidt søgte anvendelser som i eksemplet *en forandringsivrig mand der kører med langt lys og måske ikke har empati og indlevelsessevne som sine stærke sider*
- med en mulig paradigmatiske substitution som når et idiom som *ikke den skarpeste kniv i skuffen* dvs. 'en knap så klog person' får den fantasifulde variant *ikke det klareste lys på juletræet*
- med reproducerbarhed som kan danne en skabelon – som når udtrykket *tage benene på nakken* 'stikke af, flygte' brugt om en kat får formen *tage poterne på nakken*
- der kan alluderes indirekte til i en tekst, som når overskriften *Et plakatfuldt konvent* handler om et partikonvent hvor der er mange valgplakater, og ikke om et partikonvent hvor alle er *plakatfulde* dvs. 'meget berusede'.

2) I foredraget *Om indlæring af fraseologi* drøftede **Erla Hallsteinsdóttir** Kühns model for indlæring af fraser:

- *identifikation* (= *Erkennen*) af fraser i tekster
- *afkodning* (= *Entschlüsseln*) af frasernes betydning
- *indlæring / leksikalisering* (= *Erwerben*) af fraser i øvelser
- [*oversættelse* af fraser]

med særligt henblik på afkodningsstrategier over for ukendte eller mindre kendte fraser i fremmedsprog, og afgrænsningen af de egentlige fraser vis-a-vis ikke-frasema-

tiske kollokationer, på grundlag af *frasematisk analogi*, *semantik*, *fonetisk analogi*, *visualisering*, *logisk følgeslutning*, *etymologisk viden & kontekst*.

Dette blev illustreret med en lang række af fraser der indeholder substantivet *ansigt*, såsom *i sit ansigts sved – blive lang i ansigtet – lægge ansigtet i de rette folder*.

3) I foredraget ***Fraseologi og dialekt*** behandlede **Torben Arboe** bl.a. forskellige problemer ved ordbogsregistrering af fraser på dialekt, som

- dialektrim der er fremmede for rigsmål: *snø-brø* ‘sne-brød’, *døww-løww* ‘døje-leve’, *æri-græri* ‘arving-græden’; *rörren ær snå:r brörren* ‘rådden er snart bruden’
- dialektrim parallelle med rigsmålsrim: *to:r – go:r* ‘tåre – gårde’, *le:? – te:?* ‘ligge – tigge’
- tvillingformler & bogstavrim: ‘*Degne og duer de gør skidne stuer*’
- dialektidiomer med rigsmålsfremmede betydningsnuancer eller underbetydninger af ord: *han æ mæ i fajeskav*? ‘han er med i fadderskabet’ dvs. ‘han hører lidt til familien’, *fåld et får ..* ‘fald ikke, far ..’ dvs. ‘fald ikke, min ven (jovialt til dreng eller mand)’
- dialektale særbetydninger af enkeltord: *go:j* egl. ‘gode’, nu ‘ejendom, formue, gods’
- dialektale særbetydninger af frasematiske tryktabsforbindelser: *give 'af* ‘afgive, afstå’, *give 'af sig* ‘udstøde en lyd’, *give 'op* ‘samle [fløde] på overfladen; kaste op, brække sig’.

4) I foredraget ***Kollokationer set fra forskellige synsvinkler*** gennemgik **Anna Braasch** kollokationer som en undertype af faste ordforbindelser set fra forskellige synsvinkler, nemlig bl.a.:

- en empirisk-kvantitativ-distributionel synsvinkel: hvilke ord optræder hyppigt sammen og i hvilken kontekst?
- diverse lingvistiske synsvinkler: en formsynsvinkel: kollokationens grundform og variation, en syntakssynsvinkel: den indre frasestruktur samt div. andre synsvinkler på kollokationer, semantisk, leksikalsk, typologisk, funktionelt, didaktisk, stilistisk
- foruden i egentligste forstand leksikografiske og datalingvistiske synsvinkler såsom bl.a. selektion, repræsentation.

5) I foredraget ***Fraseologi og etymologi*** fremlagde **Jan Katlev** eksempler på den sproghistoriske baggrund for forskellige danske idiomer og kollokationer, anskuet som komplekse resultater af lydudviklinger, lydforvanskninger, betydningsforskydninger og øget tilsløring af en oprindelig lydlig og betydningsmæssig transparens i flerordsforbindelser:

- *Gå fløjten* af nedertysk *gân floiten* / højtysk *gehn flöten*, enten opr. af ældre nedertysk *floiten* ‘flyde bort’ eller opr. af vestjiddisch **flejtn* / **pflejtn* / **b(i)flejtn* ‘gå fallit / i ruin’.
- *Høg over høg* som oprindelig ikke har noget med rovfuglen en *høg* at gøre; udtrykket er en fonetisk baglæns hyperkorrektion fra formen *høj over høj* i betydningen ‘mægtig over mægtig’ – fremmet af idiomer som det oprindelige *høg over hane* altså ‘rovfugl over tamfugl’.

- *Nogen* (pronomen) leksikondefineret i en flerordsfrase

	'person	ikke kendt	af-den-talende'
= (let parafraseret)	'hvilken-person	ikke vidst	af-mig'
= (parafraseret)	'hvem	ikke ved	jeg'
= (med ordstillingsændring)	'ikke ved	jeg	hvem'
= (oversat til urnordisk)	<i>ne wait</i>	<i>ek</i>	<i>hwariR</i>
= (via lydudviklinger til gammeldansk)	<i>N O</i>	<i>GH</i>	<i>ÆR</i>

altså *noghær*, med akkusativen *noghæn* >> moderne dansk *nogen*.

6) I foredraget **Fraseologisk terminologi i Danmark** drøftede **Lars Brink** fraser / idiom / fraser bestemt som 'etablerede flerordsudtryk med en uforudsigelig og uigen-nemsigtig faktisk betydning der afviger fra udtrykkets pålydendebetydning'; han plæde-rede for at undgå termen fraser og i stedet at foretrække udtryk som fraser med uforud-sigelig betydning eller idiom. Han opstillede som centrale fraseologiske størrelser bl.a.

- *frase* 'etableret flerordsudtryk'
- *idiom* 'frase med uforudsigelig betydning'
- *metaforisk idiom* 'idiom med implicit sammenligning'
- *kollokation* 'frase med forudsigelig betydning'

foruden nærmere definerede særtermer som *rutineytring*, *rituel ytring*, *høflighedsfrase*, *sproglig kliché*, *konstruktion* m.m.

b) Foredrag med udgangspunkt i kontrastiv fraseologi i dansk og fremmedsprog

7) I foredraget ***Að vera ekki á hverju strái / Det hænger ikke på træerne. Dansk-islandsk fraseologi*** behandlede **Auður Hauksdóttir** ligheder og forskelle mellem dan-ske og islandske flerordsforbindelser, især idiom / idiom og kommunikative formler.

I visse tilfælde er der en påfaldende lighed mellem idiom / idiom i de to sprog. I an-dre tilfælde ligner deres flerordsforbindelser tilsyneladende hinanden, men adskiller sig alligevel på en uforudsigelig måde. Det giver vanskeligheder ved tilegnelsen af denne del af ordforrådet. Man kan både komme ud for at der er ækvivalens mellem de to sprog selv om man ikke regner med det – og så omvendt for tilfælde hvor der ikke foreligger ækvivalens selv om man regner med det.

- Ækvivalente idiom / idiom:
bryde isen – i den syvende himmel =
brjóta ísinn – í sjöunda himni
- Næsten ækvivalente idiom / idiom:
gå i hundene – være på det rene med noget =
fara í hundana – vera með eitthvað á þurru
- Ækvivalente idiom / idiom med synonymvariation:
puste nyt liv i noget – redde ansigt =
blása nýju lífi í eitthvað – bjarga andlitinu
- Idiom / idiom med partiel ækvivalens:
ikke have salt til et æg – hænger ikke på træerne =
eiga ekki salt í grautinn – eru ekki á hverju strái
- Danske idiom / idiom uden islandsk ækvivalent:
være ude at svømme – skære noget ud i pap
- Islandske idiom / idiom uden dansk ækvivalent:
draga ýsur – komið annað hljóð í strokkinn

- Høflighedsvendinger:
tak for hjælpen – god dag – god rejse =
takk fyrir hjálpin – góðan dag – góða ferð.

8) I foredraget **Alles in einen Topf? Problemer i dansk-tysk fraseologi – leksikologisk, translatorisk og leksikografisk** diskuterede **Ken Farø** en teoretisk skelnen mellem de tre distinkte discipliner *leksikologi*, *translatologi*, *leksikografi* i forhold til beskrivelsen af idiomer. Ved kontrasteringen af to sprogs idiomer i ordbogssammenhænge og ved tekstoversættelser opererer man ofte med en overdrevet absolut distinktion mellem klart ækvivalente idiomer, klart ikke ækvivalente idiomer og delvis ækvivalente idiomer. Denne overdrevne absolutte distinktion skyldes en overdreven baseren sig på ydre sproglig *form* med for ringe hensyntagen til sproglig *funktion* og sproglig *kontekst*.

De tre discipliner *leksikologi*, *translatologi*, *leksikografi* defineres ideelt som

- *kontrastiv leksikologi*: ‘leksikalsk sammenligning mellem forskellige sprogs ordforråd som sådan uden hensyn til formål og anvendelse’
- *translatologi*: ‘undersøgelse af transmission af tekstindhold / tekstbetydninger mellem forskellige sprog’
- *(flersprogs)leksikografi*: ‘undersøgelse af flersprogsordbøger ud fra de funktioner de skal varetage’.

De tre discipliner indebærer hver for sig sin karakteristiske *fejlslutning*.

- *Den leksikologiske fejlslutning* består typisk i at oversætte standardidiomer i et kildesprog med standardækvivalenter i et målsprog uden hensyn til oversættelses-naturligheden i den lokale kontekst, som fx når man oversætter tysk .. *blamiere mich bis auf die Knochen* til dansk .. *blamerer mig selv helt ind til skindet* hvor det tyske udtryk i sammenhængen er idiomatisk naturligt, mens den foreslåede danske ækvivalent ikke er naturlig i den konkrete lokale kontekst ‘blamerer mig’.
- *Den translatoriske fejlslutning* består i uden forbehold at bruge fulde tekstoversættelser som kilde til gode idiomækvivalenter i ordbøger – denne fejlslutning kan være mildere som når man i en ordbog gengiver et nutidigt gangbart idiom, fx tysk *ihre Brötchen verdienen*, med et tilsyneladende tilsvarende forældet dansk idiom som *tjene deres brød* – men den bliver selvfølgelig grovere hvis den udspringer af lodret manglende kendskab til kildesprogets idiomatik, som når det overførte danske udtryk *lige ud af landevejen* bliver litteralt oversat til et ikke overført tysk udtryk *gleich draussen an der Landstrasse*.
- *Den leksikografiske fejlslutning* består i at anvende tosprogede ordbøger som kilder til ækvivalentforslag i tekstoversættelser som hvis det danske udtryk *sommerfugle i maven* ‘nervøs kildren i mellemgulvet’ uden videre i en tekst oversættes med det tilsyneladende ækvivalente tyske udtryk *Schmetterlinge im Bauch* som næsten kun betyder ‘forelskelse’.

9) I foredraget **Bliver falske venner mere ægte med tiden? Om engelsk indflydelse på danske flerordsenheder** gennemgik **Henrik Gottlieb** en række nyere danske fraselån fra engelsk, både i form af engros-lån af nyt stof og betydningsændringer af udvendigt set ækvivalente danske idiomer, med eksempler som fx *cold turkey* >> *kold tyrker*; *bring up* >> *bringe op* [dvs. ‘bringe på bane’]; *make ends meet* >> *få enderne til at mødes* [dvs. ‘få pengene til at slå til’]; *rotten apples* >> *rådne æbler* [dvs. ‘uheldige elementer’], i stedet for det etablerede danske *brodne kar*; ortografisk afsmitning som fx dansk **adresse* << engelsk *address*; indsnævring af betydningsomfang ved lån som fx dansk *mail* ‘elektronisk post’ vs. engelsk *mail* ‘post (i almindelighed)’.

10) I foredraget *Jeg har det lige på tungen! Fraseologiske somatismer i italiensk, fransk og dansk* drøftede **Mirella Cristofoli & Lilian Stage** det kendte fænomen at kroppen er udgangspunkt for vores erfaringer og at dette også gengives sprogligt, idet betegnelser for legemsdele i alle sprog fungerer som 'byggeklodser' i fraser, sikkert fordi de samme legemsdele har relativt samme symbolske betydning i mange sprog. Men netop kun *relativt* samme betydning, idet der kan være kulturelt betingede asymmetrier mellem forskellige sprogs idiommer med betegnelser for legemsdele, såkaldte *somatismes*.

- Et eksempel på samme somatisme i italiensk, fransk og dansk har man i *jeg har det lige på tungen* 'jeg er lige ved at komme på et ord' = *je l'ai sur le bout de la langue* = *l'ho sulla punta della lingua*.
- En somatisme 'der ligner' har man på fransk i *l'Europe démocratique, mon oeil!* – dvs. '.. nej den tror jeg ikke på, den går ikke..', som bedst oversættes til dansk som *det demokratiske Europa, i øjet (på Karl den tolvte)!*
- Med en lille forskydning i betegnelsen for legemsdel, 'tå' >> 'fod', får man *træde over tæerne* = *marcher sur les pieds* = *pestare i piedi*
- Med større forskydninger i bl.a. den anvendte legemsdel, men med samme idiom-betydning, har man fx i *ikke lægge fingrene imellem* = *ne pas avoir la langue dans la poche* egl. 'ikke have tungen i lommen' = *non avere peli sulla lingua* egl. 'ikke have hår på tungen'.

11) I foredraget *Fraseologi i almene tosprogede ordbøger og i det virkelige liv. Forskelle på kontekstafhængige og kontekstuaafhængige oversættelser* fremlagde **Rolf Hesse** eksempler på tolkeoversættelser ved EU-Domstolen fra nygræsk til en række andre EU-sprog af en passus i en tekst vedrørende tilladelse til at eksportere 'tvivlsomt' oksekød fra Irland til de øvrige EU-lande. I denne tekst står, ordret oversat til dansk, ca.: '.. jeg mener at en accept af en elastisk løsning, ... – *ville åbne Aiolos*' sække i retning af at eventuelle uregelmæssigheder blev godkendt'.

Dette udtryk *ville åbne Aiolos*' sække refererer græsk-internt til en scene fra Odysseen hvor det at åbne vindguden Aiolos' sække med alle vindene fra forskellige verdenshjørner får Odysseus' skib til at sejle forvirret hid og did i stærk storm uden retning. Dette billedsproglige idiom blev her konkret oversat af de forskellige tolke:

- dansk: *ville åbne for sluserne*
- svensk: *skulle riskera att störa ordningen* 'ville risikere at forstyrre (ro og) orden'
- engelsk: *would tend to disturb the waters* 'ville kunne bringe vandene [havet] i oprør'

men med erstatning med et andet billedsprogligt idiom [*åbne Pandoras æske*] på

- tysk: *eine Büchse der Pandora .. öffnen würde*
- fransk: *ouvrir une boîte de Pandore*.

12) I foredraget *Dansk-japansk fraseologi* præsenterede **Masako Suzuki** en dansk idiomordbog for japanere indeholdende ca. 1000 danske idiommer.

Med udgangspunkt i den danske faglitteratur om idiommer og talemåder er der blevet udvalgt et materiale af centrale danske idiommer, som er blevet ordnet efter deres semantiske kerner, fx substantivet *agurk* i idiomet *at gå agurk*, og derefter forsynet både med (1) en bogstavelig oversættelse til japansk, (2) en japansk betydningsækvivalent til

det danske idiom, og (3) et dansk eksempel indeholdende det danske idiom, fx *hendes forældre gik totalt agurk da hun kom for sent hjem*.

Eksempler på danske idiomer som de japanske leksikografer havde formmæssige og betydningsmæssige problemer med var fx *sætte på den anden ende – ikke stikke dybt – bore i et problem – slå nogen for penge – have skudt papegøjen – komme på benene igen – gå op i en spids*.

Eksempler på danske idiomer som danske leksikografer har været konsulteret om er fx *have fidus til – hvile på laurbærrene – der er krummer i ham – tabe ansigt*.

Eksempler på idiomatiske inkongruenser mellem dansk og japansk er fx *vaske sine hænder* (overført), på japansk = *vaske sine fødder*, *leve som hund og kat*, på japansk = *leve som hund og abe*, *han er chefens højre hånd*, på japansk = *han er chefens højre arm*.

Symposiet sluttede med en opsamlingsdiskussion med to hovedemner: (a) en nærmere fastlæggelse af en fælles dansk fraseologisk terminologi og (b) retningslinjer for det fremtidige arbejde inden for Dansk Selskab for Fraseologi. Under (a) var beklageligvis foredragsholderen om **Fraseologisk terminologi i Danmark**, **Lars Brink**, ikke til stede. Man drøftede bl.a. ud fra hans oplæg hvor heldigt det var at anvende termen *frasem* i sammenligning med det tysk inspirerede tungere alternativ *fraseologisme* eller mere tilvante termer som *idiom*, *flerordsforbindelse* m.m.

Et andet centralt problem var spørgsmålet om 'puristiske' eller internationale termer. **Ken Farø** påpegede her det i dansk tradition meget forskelligt denoterende *tale-måde*, der snart bruges om ordsprog og andre 'mikrotekster', snart om idiomer (i snæver forstand), snart almentsprogligt om fraser, klicheer og uforpligtende formuleringer generelt. – Man nåede ikke frem til fuld enighed om disse spørgsmål, men det blev foreslået at lade dem viderebehandle i et arbejdsudvalg.

Denne terminologidiskussion kan genoptages ved de foredragsarrangementer i DANFRAS-regi som planlægges senere i løbet af 2005.

DANFRAS har omkring 80 medlemmer i Danmark, foruden fra Polen, Japan, Island, England, Tyskland, Belgien, Sverige og Schweiz.

DANFRAS' formand **Ken Farø** redigerer bind 35 af tidsskriftet HERMES som udkommer efteråret 2005 med fraseologi som tyngdepunkt, med danske, nordiske og udenlandske bidrag.

Den SprogTeknologiske Ordbase (STO) for dansk – Leksikografiske aspekter

Af Anna Braasch og Sussi Olsen (Center for Sprogteknologi, KUA)

Indledende bemærkninger

Dette indlæg fokuserer på enkelte aspekter af STO-projektet. Derved er der mange emner som vi ikke kan komme ind på her, eksempelvis detaljerede oplysninger om sammensætningen af ordforrådet (fordelt på ordklasser, domæner mv.), beskrivelsen af sprogteknologiske anvendelser af materialet, lingvistiske specifikationer for STO samt leveringsbetingelser mv. – alt dette findes på adressen www.cst.dk, og her er flere muligheder: 'STO-ordbase' giver adgang til brugergrænsefladen, 'STO licensvilkår' indeholder oplysninger om materialet, links til artikler, levering mv. Generelle oplysninger om projektet findes på www.cst.dk/sto.

Hvad er STO?

En sprogteknologisk ordbase består af en samling data der beskriver ords lingvistiske egenskaber, der er udarbejdet fortrinsvis til anvendelse i sprogteknologiske værktøjer.

Formålet med at udarbejde STO var at afhjælpe flaskehalsproblemet i dansk datalingvistik og sprogteknologi, idet der på disse områder var et udpræget behov for en stor ordbase, men udviklingen kunne ikke varetages af en enkelt institution eller i forbindelse med udviklingen af et enkelt produkt fordi arbejdet er ressourcekrævende og dermed dyrt at udføre.

Løsningen var at udvikle en ordbase hvis omfang, indhold og beskrivelsesmetode er velegnet til at dække forskelligartede behov. På den tekniske side var det et grundlæggende krav at ordbasens indhold skal kunne konverteres, udnyttes og eventuelt udbygges både til forskningsformål og i kommercielle anvendelser.

Hvad indeholder STO?

Ordforrådet stammer primært fra almensprog (68.000 ord) og er fortrinsvis baseret på et avis-korpus. Det fagsproglige ordforråd (ca. 13.500 ord) stammer fra seks udvalgte emneområder, nemlig edb, miljø, sundhed/helse, finans, forvaltning samt handel/erhverv.

De lingvistiske oplysninger omfatter ordenes bøjning, stavning og deres syntaktiske konstruktionsmuligheder samt for en del af ordforrådet også betydningsrelaterede semantiske oplysninger, og hele materialet er udarbejdet på basis af korpusbelæg. Leksikografer, lingvister og interesserede sprogbrugere kan få indblik i en del af STO-databasens indhold over internettet vha. brugergrænsefladen, jf. afsnittet "Brug af grænsefladen". Der er dog en række oplysningstyper som er kodet i databasen, men ikke vises på skærmen, da udviklingen af brugergrænsefladen blev afsluttet i 2002, og på det tidspunkt var ordbasens struktur endnu ikke fuldt implementeret. Vi håber at få mulighed for at videreudvikle grænsefladen, således at man vil kunne se alle oplysninger der er relevante for sprogbrugere.

STO i anvendelser

Informationsteknologi bruges i dag i stadig større udstrækning, og da ingen danske sproglige værktøjer eller programmer kan arbejde uden adgang til en stor dansk ordbase (også kaldet leksikon), er STO et oplagt basisprodukt til anvendelse som *leksikonkomponent* i sådanne applikationer.

De primære anvendelsesområder er online informationssøgning, automatisk og maskinstøttet oversættelse, spørgsmål/svar-systemer, sproglige hjælpemidler for handicappede, programmer til sprogundervisning mv. STO benyttes i en række sprogteknologiske værktøjer, eksempelvis i en *parser* (til automatisk sætningsanalyse), en *tagger* (morfosyntaktisk opmærkning af tekster), et *lemmatiseringsprogram*, samt i en prototype til et indholds-baseret *søgesystem*.

Ordbasens materiale er også velegnet til brug i datalingvistisk forskning, eksempelvis i testning af grammatikker. I sprogforskning og leksikografi kan STO danne grundlag for systematiske undersøgelser af ords egenskaber, eksempelvis med hensyn til deres konstruktionspotentiale eller typer af komplemente.

Enkeltopslag i STO over internettet kan desuden anvendes i undervisning/indlæring af dansk som fremmedsprog på højt niveau fordi ordbasen indeholder de oplysninger der er relevante mht. korrekt bøjning, stavning og syntaktiske konstruktionsmuligheder i en overskuelig form.

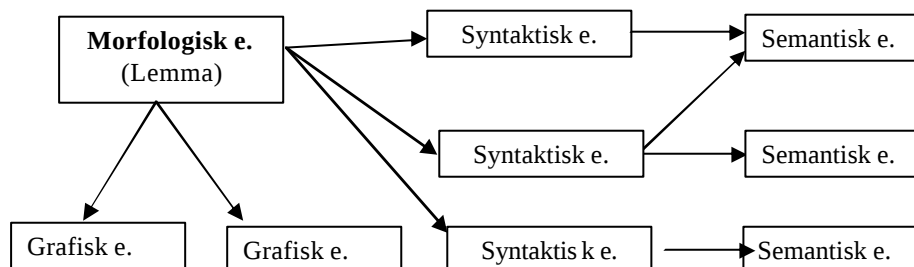
Leksikografiske principper og arbejdsmetode

I udarbejdelsen af STO blev der lagt vægt på at dække et bredt sammensat ordforråd, at beskrive ords egenskaber entydigt, præcist og detaljeret og at formulere oplysningerne på en sådan måde at materialet kan konverteres til forskellige formalismer. Den leksikografiske arbejdsmetode byggede på relevante resultater fra datalingvistiske forsknings- og udviklingsprojekter samt veldokumenterede erfaringer fra leksikografiske projekter. I det danske leksikografiske miljø havde STO især gavn af den faglige kontakt med Den Danske Ordbog.

Arbejdsgangen

Udarbejdelsen af materialet (vi kaldte processen 'kodning') var fordelt på medarbejdere fra fire institutioner med forskellige fysiske adresser og med forskellig faglig profil. Arbejdsgangen omfattede fire hovedtrin: *kodning* af lingvistiske oplysninger, *krydstjek* hos en anden institution end den der stod for kodningen, *validering* samt *import* af de validerede data i STO-basen. De sidste to trin blev altid udført af CST. Arbejdet var tilrettelagt af CST og blev udført på grundlag af en fælles *kodningsmanual* der indeholdt både de lingvistiske retningslinjer og de nødvendige tekniske vejledninger. Kodningsmanualen blev trinvist videreudviklet og revideret i løbet af projektet på grundlag af de praktiske erfaringer fra ordbogsarbejdet. I denne proces deltog alle samarbejdspartnere. Der blev udviklet særlige *kodningsværktøjer* i ACCESS som gav god støtte til koderen mht. konsistens og samtidig effektiviserede arbejdet.

Kodningsarbejdet blev udført i moduler, fordelt på almensprog og fagsprog, på ordklasser og hhv. på morfologi, syntaks og semantik. Denne opdeling gav sig naturligt fra den lagdelte lingvistiske beskrivelsesmodel som danner grundlaget for STO. Figur 1 er en skitse af denne model.



Figur 1. Den lagdelte beskrivelsesmodel: Oplysningerne er fordelt på 'enheder' (e.).

Forud for hvert nyt modul i kodningsprocessen blev der afholdt et fysisk møde med deltagelse af alle projektmedarbejdere hvor de relevante lingvistiske afsnit fra kodningsmanualen blev drøftet og kommenteret. Derefter fulgte en kodningstest hvor retningslinjerne blev afprøvet i praksis af alle på en fælles lemmaliste, sådan at vi kunne se om der var forskelle i implementeringen af retningslinjerne – og hvad det skyldtes. Derefter justeredes retningslinjerne hvis det viste sig at der var uafklarede eller uhensigtsmæssige ting i selve vejledningen. Formålet med denne proces var at sikre så ensartet og konsekvent kodning som overhovedet muligt.

Som det fremgår af figur 1, er STOs beskrivelsesmodel opdelt i morfologiske, syntaktiske og semantiske enheder. I det følgende vil oplysningstyperne for hver type enhed blive gennemgået.

Den morfologiske enhed

De morfologiske oplysninger til et lemma er i praksis fordelt på flere blokke af oplysninger og omfatter ordklasseangivelse, stavemåder (ofte flere, fx *hæfte/hefte*), bøjningsmønstre (ofte mere end et), dekomponering (for sammensatte substantiver), fugeelement (for substantiver), autonomi ('NO' angiver at ordet kun kan forekomme i kollokationer eller andre faste udtryk, fx *a conto*) samt om hver stavemåde er godkendt af Retskrivningsordbogen. Desuden omfatter den morfologiske enhed oplysninger om status (om ordet er kodet med både morfologi, syntaks og semantik).

Id: <i>MEDLEMSBONUS</i>	
Lexcat: <i>noun</i>	Sublexcat: <i>Common</i>
Autonomi: <i>YES</i>	
Origin: <i>H og E-korpus</i>	
Status for morfologi: <i>EFTER SAP</i>	<i>SAP=samarbejdspartnere</i>
Status for syntaks: <i>KODET AF SAP</i>	
Gmu_id: <i>GMU_MEDLEMSBONUS_1</i>	
Spelling: <i>medlemsbonus</i>	RO-approved: <i>YES</i>
Dekomposition: <i>medlem+s+bonus</i>	
Fugeelement: <i>Removed: Added: 0</i>	
GINP_ID: <i>MFG0225 (+en,+er,+erne)</i>	RO-approved: <i>NO</i>
<i>bøjningsmønstre</i>	
GINP_ID: <i>MFG0246 (+sen,+ser,+serne)</i>	RO-approved: <i>YES</i>

Hvis et ord har flere stavemåder, vil der være flere gmu_id'er (graphical morphological unit) med hver deres spelling (stavemåde), dekomponering, fugeelement og bøjningsmønstre tilknyttet.

STO har valgt at medtage visse stavemåder der ikke er anerkendt af Retskrivningsordbogen 2001 (RO2001), jf. RO-approved YES/NO ovenfor. Det drejer sig først og fremmest om stavemåder der har ændret status fra godkendt til ikke-godkendt eller omvendt i de seneste udgaver af RO. Da STO skal kunne bruges til automatisk genkendelse, og da dette nødvendigvis også må omfatte tekster ældre end den sidste udgave af RO, er der i STO medtaget former i overensstemmelse med RO 86 og frem. Disse former mærkes som ikke-godkendte. Andre ikke-godkendte stavemåder i STO er fx *canarisk* og *sclerose*, som begge er uhyre hyppigt forekommende.

Også bøjningsmønstre kan være ikke-godkendte i tilknytning til visse ord. Følgende lemmaer med bøjningsmønstre er ikke godkendte if. RO2001

tilbud (+et,+,+ene)
kursus (+et,+,+ene)

og et ord som *test* har if. RO aldrig måttet bøjes som (+en,+s,+ene), men s-pluralisformen forekommer ikke desto mindre i mange tekster og er derfor medtaget i STO.¹

Der kan tilknyttes flere fugeelementer til hver ord (gmu_id), fx *papir*, så såvel *papirspose* som *papirflyver* kan håndteres. I fugeelement angives hvad der fjernes fra spelling før fugeelementet tilføjes, fx

arbejde removed: e added: s ? *arbejds...*

¹ Det skal bemærkes at arbejdet med STO og dets forløber PAROLE-projektet har strakt sig over tre udgaver af Retskrivningsordbogen, og nogle stavemåder og bøjningsmønstre der ikke længere er godkendte, kan stadig optræde som godkendte. Dette gælder for ord der blev behandlet tidligt i projektforløbet, og vi håber at vi inden længe får tid til at rette op på det.

Den syntaktiske enhed

De syntaktiske oplysninger omfatter ordets valensmønster hvorved vi forstår både dets aritet – hvor mange led det knytter til sig – leddenes syntaktiske funktion (subjekt eller objekt fx) samt deres syntaktiske kategorier, fx om et tilknyttet led er et nominalsyntagma, et præpositionssyntagma eller en sætning. Desuden oplyses om reflektivitet, partikelverbum, hjælpeverbum og om et led er optionelt eller obligatorisk, og der gives for de fleste ord et eksempel (til brug for mennesker).

Nedenfor følger to eksempler på syntaktiske enheder:

Lemma: **indsprøje**

Syntaktisk beskrivelse: **Dv3NP_{a0}-i** ? Trivalent verbum, styrede led: subjekt (ikke udtrykt), objekt=NP, adverbialt PP-objekt (optionelt) indledt af præp. 'i' hvor styrelsen kun kan være et NP.

Eksempel: Lægen **indsprøjter** stoffet i patienten

Lemma: **præference**

Syntaktisk beskrivelse: **Dn2GP_{ni}-for** ? Divalent substantiv, styrede led: valensbunden genitiv, optionelt PP-objekt indledt af præp. 'for', hvor styrelsen kan være et NP eller en infinitiv.

Eksempel: Frischs **præference** for arbejderklassen

En af de vanskeligste problemstillinger i en valensbaseret ordbog er at afgøre hvorvidt et led er valensbundet og dermed skal medtages i den syntaktiske beskrivelse. Klart valensbundne komplementer, obligatoriske såvel som optionelle, medtages altid i valensbeskrivelsen i STO, hvorimod frie modifikatorer som tid og sted ikke medtages. Den problematiske gruppe der ligger midt imellem, de såkaldte 'middles' if. Harold Somers, behandles individuelt for de forskellige ordklasser, og frekvensen af korpusbelæg spiller i den forbindelse en vigtig rolle. Især for substantiver medtages led som ikke klart er valensbundne, men som er højfrekvente og samtidig centrale for kernebetydningen af substantivet. Dette bunder i sprogteknologiske hensyn.

Som det fremgår af ovenstående, angives i valensbeskrivelsen materialet for styrelsen i præpositionssyntagmaer. Det kan være hhv. NP, at-sætning, infinitiv og spørgebisætning. Ved infinitiv angives også styrelsesforhold, dog kun hvor dette har relevans, fx ikke ved substantiver. Ved infinitiv angives også om denne indledes med *at*.

Af de øvrige syntaktiske træk som STO indeholder, kan nævnes oplysningen om hvorvidt et verbum er ægte reflektivt eller evt. blot kan konstrueres med reflektivt objekt samt oplysning om hvorvidt et verbum kan indgå i ekspletivkonstruktioner (både *det* og *der*).

Den semantiske enhed

STO arbejder med flere niveauer af semantik. Projektet blev i opstartsfasen beskåret mht. semantikdelen, og derfor blev målet for semantikken todelt: dels at integrere semantikdataene fra SIMPLE-projektet², dels at forsyne ordene fra ét af domænerne med semantisk information, men i noget mindre omfang end SIMPLE-semantikken. Ordforrådet er derfor beskrevet med semantik i tre forskellige detaljeringsgrader. Niveau 1 består i at alle ord er forsynet med en angivelse af hvilket domæne de tilhører. På niveau 2 består den semantiske information ud over domæneangivelse af betydningsopdeling, ontologisk tilhørsforhold baseret på SIMPLE-ontologien, argumentstruktur og selektions-

²SIMPLE var et EU-projekt (1999-2000) der drejede sig om at udvide de eksisterende PAROLE-ordbøger med kompleks semantisk information, se www.cst.dk/simple.

restriktioner. Dette niveau er kodet for STOs 2500 sundhedsord. Nedenfor ses et eksempel på en semantisk enhed på niveau 2:

Korpuseksempel: *Læger indsprøjter stoffet N. i narkoofre.*

Ontologisk tilhørsforhold: Event Act **Relational act** Purpose act

Syntaktisk konstruktion: **Dv3NP_a0-i**

Verbum trivalent, styrede led: subjekt (default), NP objekt, optionel PP-objekt med præposition 'i'

Argumentstruktur: **arg12_ADJi**

subjekt=arg1, objekt=arg2, præpositionssyntagma='adjunct' (modifikator)

Selektionsrestriktioner: **ARG12ADJhum_sub_bpahum**

subjekt: **human**, objekt: **substance**, modifikator: **body part/hum**

hvoraf det fremgår at *indsprøjte* i SIMPLE-ontologien er en *purpose act* som er en type *relational act* som har *act* som overbegreb. Derpå beskrives argumentstrukturen og selektionsrestriktionerne som viser at trækket *human* restringerer subjektet, *substance* objektet og modifikatoren trækket *hum* eller *body part*. Da ordet her stammer fra sundhedsdomænet, som ikke inkluderer det veterinære domæne, er *animate* ikke et muligt træk for denne modifikator.

Niveau 3 indeholder informationerne fra niveau 2 + qualiaroller, polysemi- og synonymirelationer som defineret i SIMPLE-modellen. Dette gælder for de 7000 almensproglige ord som blev integreret fra SIMPLE. For eksemplet ovenfor ville der ud over polysemirelationer og synonymer også kunne tilskrives en semantisk relation kaldet *concerns* som ofte udfyldes netop ved semantiske enheder af *purpose act*-typen.

Arbejdsgrundlag

I *kodningsprocessen* havde alle foruden kodningsværktøjet adgang til samme arbejdsgrundlag, nemlig korpusser, lemmalister, kodningsmanual og lingvistiske specifikationer. Der er benyttet følgende korpusser i kodningen af almensproglige ord:

- Berlingske Tidende og Weekendavisen (1999) ca. 30 mio. ord (primært korpus)
- Berlingske Tidende (1990-92) ca. 20 mio. ord
- DK87-90 (1987-90) ca. 4 mio. ord
- Korpus 2000 (1998-2002) og Korpus 90 (1988-92), begge på ca. 28 mio. ord.

Inden adgangen til de to sidstnævnte DSL-korpusser fik projektet lov til at benytte Den Danske Ordbogs korpus og søgeværktøjet Semaskop.

Lemmaselektion

Forholdet mellem korpus og ordforråd afspejler sig i *lemmaselektionen*, som forløb på to måder. Det almensproglige ordforråd tog sit udgangspunkt i DDO's frekvensbaserede ordlister, hvorfra der blev udvalgt en lemmakandidatliste for STO på grundlag af frekvensoplysningerne. Lemmakandidaterne blev derefter verificeret vha. aviskorpusserne og senere ligeledes vha. Korpus 2000.

Det fagsproglige ordforråd fra de tidligere nævnte domæner er udvalgt og kodet på grundlag af tekster der er fundet på internettet inden for de valgte domæner. For hvert domæne er der selekteret en tekstmængde på mellem 1 og 2 mio. løbende ord; den valgte teksttype repræsenterede kommunikation fra fagmand til lægmand, dvs. det er ikke eksperttekster med høj fagsproglighedsgrad. Formålet med at vælge netop denne teksttype var at det fagsproglige ordforråd skulle supplere det almensproglige ordforråd med domæneord der forstås og anvendes i almindelig kommunikation om fagområdets emner, eksempelvis i bekendtgørelser til borgerne om miljø eller sundhed. Vi kalder disse ord for gråzoneord da de ikke er deciderede termer, men heller ikke er generelle ord.

STO er i stor udstrækning korpusbaseret, idet både lemmaselektion, kodning af morfologisk, syntaktisk og semantisk information baseret på systematiske undersøgelser af ords forekomster i teksterne udgør korpusgrundlaget.

Brug af STOs grænseflade

Der er, som nævnt, mulighed for opslag i STO over internettet fra CST's hjemmeside, som også har nogle tips til interessante søgninger. Grænsefladen har fire forskellige søgemuligheder:

- *Ordsøgning* med lemma eller ordform hvis man har brug for opslagsordets bøjningsformer (et eller flere paradigmer), eller den søgte ordforms lemmaform og dets bøjningsformer (et eller flere paradigmer). Lemmaets/lemmaernes syntaktiske konstruktionsmuligheder bliver også vist.
- *Substantiviske sammensætninger*. Søgeresultatet er en liste med alle substantiver som søgestrengen indgår i, dvs. både sådanne ord som er registrerede i STO som sammensætninger, og alle ord der ikke er registreret som relevant sammensætning i basen, men eventuelt kan være det (markeret med kursiv). Desuden kan man få vist sammensætningens struktur.
- *Korpussøgning* med opslagsordet eller en enkelt ordforms forekomster i et korpus bestående af artikler fra Berlingske Tidende, årgang 90-92.
- *Parameterbaseret søgning* med kombination af forskellige lingvistiske kriterier. Søgeresultatet er en liste af tilfældigt udvalgte ord (op til 30) der opfylder den valgte kombination.

Den sidstnævnte søgemulighed er den leksikografisk og lingvistisk set mest interessante, fordi man kan få vist grupper af ord der har samme lingvistiske egenskaber ud fra den valgte parameterkombination, eksempelvis en gruppe af divalente verber med partiklen og/eller præpositionen *for*. Figur 2 viser et udsnit af resultatet for denne søgning. Den forkortede beskrivelse af konstruktionen kan være lidt svær at gennemskue, derfor er det muligt at klikke på lemmaet og få vist alle ordets konstruktioner i detaljer og med eksempler. I tilfældet *ængste* er nedenstående konstruktion en ud af de fem registrerede for dette verbum. Ved samme opslag vises også ordets morfologi for at skabe sammenhæng mellem det morfologiske og det syntaktiske beskrivelseslag.

Lemma	Ordklasse	Forkortet beskrivelse af konstruktion (klik på lemmaet for uddybning)
1 stå,1	verbum	divalent: NP, PP(NP/infinite subject equi control) prep=for
2 ængste	verbum	divalent reflexive: NP, PP(NP/that-c/infc w. subject control) prep=for
3 udgive	verbum	divalent reflexive: NP, obligatory PP(NP/infc) prep=for
4 grue	verbum	divalent: NP, PP(NP/that-clause) prep=for

Figur 2: Søgeresultat for en parameterbaseret søgning (udsnit)

Links fra STOs brugergrænseflade

For at kunne sammenligne STOs kodninger med andre ordbøger og se ordets forekomster i tekster er der etableret links til følgende sproglige ressourcer på nettet:

- Retskrivningsordbogen (3. udgave 1996-2002, Dansk Sprognævn)
- Den Danske Netordbog (2002-2004, Center for Leksikografi, HHÅ)
- Korpus 2000 (Det Dansk Sprog- og Litteraturselskab)
- GOOGLE (søgning på danske internetsider)

Om STO-projektet

Projektet modtog bevilling fra Ministeriet for Videnskab, Teknologi og Udvikling for perioden 1. marts 2001 – 29. februar 2004. Ordbasen er produceret i et samarbejde mellem Center for Sprogteknologi, KUA; Institut for Datalogistik, HHK; Institut for Almen og Anvendt Lingvistik, KUA; Institut for Fagsprog, Kommunikation og Informationsvidenskab, SDU. Projektledelse og koordinering blev varetaget af CST.

Brugergrænsefladens udvikling er finansieret af Danmarks Elektroniske Forskningsbibliotek (DEF) i 2002.

Den Danske Ordbog – nok engang

Af Henrik Lorentzen

Den Danske Ordbog (DDO) blev præsenteret og diskuteret ved et medlemsmøde i LEDA i april 2004, og Christian Becker-Christensen (CBC) har samlet sine kommentarer fra dette møde i en artikel i LEDA-Nyt nr. 38 (nov. 2004). I det følgende vil jeg knytte nogle bemærkninger til CBC's artikel, idet jeg selvfølgelig hverken kan eller vil kommentere alt, men i øvrigt er jeg enig med CBC i en del af hans betragtninger og iagttagelser. Mine bemærkninger er grupperet efter overskrifterne i CBC's artikel.

Layout og typografi

CBC er generelt tilfreds med den overskuelige og luftige måde artiklerne i ordbogen præsenterer sig på; især noterer han sig at artiklens tre hoveddele (hoved, krop og fod) markeres ved at begynde på ny linje. Desuden er der et linjeskift mere – inden for artiklens hoved – idet bøjningsoplysninger altid begynder på ny linje. CBC anfører at det i visse tilfælde medfører artikler der "brækker fra hinanden" (eksempler er **gudsdyrkelse**, **afglans**, **a la** og **boldøje**). Hertil vil jeg bemærke at problemet opstår ret sjældent, nemlig ved artikler hvor oplysningstyperne bøjning og udtale fylder meget lidt, eller hvor én af dem er helt fraværende. Ser man fx på den ordbogsside hvor **afglans** optræder, har ingen af de andre 17 artikler dette, vel især æstetiske, problem, idet de alle har både bøjning og udtale, således at artiklerne ikke "brækker over" på det pågældende sted.

Dette luftige udseende, mener CBC, står så i modsætning til det mere kompakte indtryk som sektionen med faste udtryk kan give, fx får de to faste udtryk til verbet **brede** lov til at "brede sig" over ganske mange linjer med kun ét linjebrud. Det er rigtigt at det kan virke kompakt når der som i tilfældet *brede sig* er tale om tre betydninger, der hver især har underbetydninger. Det redaktionen har villet vise, er at de faste udtryk er underordnet selve opslagsordet (sub-lemma har netop været den redaktionsinterne betegnelse), og derfor har vi syntes at linjebrud ved hver betydning af de faste udtryk ville signalere et for stort brud. De enkelte betydninger er markeret med bogstaverne **a**), **b**) osv. i halvfed, og det har da også været overvejet om en kraftigere fremhævelse end halvfed ville være hensigtsmæssig, men vi endte med at konkludere at den valgte mellemting mellem det faste udtryks fede stil og den øvrige teksts ordinære sats ville give det vi ønskede. I artiklen **ben** virker det fx umiddelbart enkelt at orientere sig i de 31 faste udtryk selvom nogle af dem har flere betydninger.

Markering af informationstyper

CBC nævner et par tilfælde hvor han mener det er vanskeligt at lokalisere de enkelte oplysningstyper. Et af disse tilfælde har nu også været diskuteret og overvejet grundigt inden det endelige layout blev fastlagt. Hvis en artikel kun indeholder én betydning, evt. med tilhørende underbetydninger, er overgangen fra hoved til krop kun markeret med

linjeskift, og den overgang, må jeg indrømme, er ikke tydeligt nok markeret idet den foregående linje (som CBC ganske rigtigt bemærker) ofte er fyldt ud (med bøjning og udtale). Desværre vil man faktisk ofte lægge mere mærke til en efterfølgende underbetydning, som jo er markeret med en udtrykket bullet. Som nævnt har vi overvejet flere løsninger på dette problem, bl.a. en udtrykket bullet som ved unummererede underbetydninger, og en eller anden form for specialtegn har også været på tale. Vi har ikke været tiltrukket af den noget ulogiske løsning som Nudansk Ordbog har valgt, nemlig at markere betydningen med et 1-tal uanset om der fortsættes med betydning 2, 3 osv. eller ej.

Et andet problem CBC nævner, er at de faste udtryk typografisk kan opfattes som om de hører til den sidste betydning. I og med de faste udtryk ikke er rykket ud, men følger spaltens venstre kant, er det principielt muligt at tolke det på denne måde, men vi mener at den fede skrift, som er lidt mere komprimeret end opslagsordenes også fede skrift, er et tydeligt signal om overgangen til ny oplysningstype.

Endelig bemærker CBC at der kan være et problem ved etiketter (sigler hedder de hos os) af typen AFL, SMS, HIST, som bruges i artiklens fod (orddannelse og etymologi). Det er korrekt at der er sammenfald i typografien mellem de netop nævnte etiketter og dem der bruges i betydningsdelen, bl.a. SYN, JF, ANT, ENC. Problemet skulle så være at AFL, SMS, HIST på grund af sammenfaldende typografi ikke tydeligt viser at der er tale om en anden del af artiklen. Dertil vil jeg bemærke at vi bevidst har ønsket den samme typografi til disse stempellignende etiketter, og når de nævnte typer netop ikke bruges i samme del af artiklen, ser vi ikke det som et problem. Derimod kunne man måske have ønsket sig at AFL, HIST o.l. var rykket til venstre, men det ville givetvis forstyrre i forhold til det næste opslagsord, som jo også er rykket til venstre.

Fordeling af informationstyper

CBC har bidt mærke i den sektionsopdelte fordeling af oplysningstyperne og påpeger nogle uheldige konsekvenser, fx i tilfælde hvor opslagsordet har flere varianter med forskellig bøjning og udtale. Jeg må medgive at vi i artikler som **gudsforgåen** får en meget kompliceret fremstilling af variantforholdene, men når der som her er tale om ikke bare en dobbeltform, men en tripelform, bliver det svært at undgå en høj grad af kompleksitet. De tilsvarende artikler i andre ordbøger som fx Retskrivningsordbogen og Nudansk Ordbog er også indviklede og svære at afkode for en ubefæstet bruger. For redaktionen har det været afgørende at følge det princip at opslagsformer, bøjning og udtale behandles færdigt hver for sig, altså en konsekvent strukturel tilgang (ingen sammenblanding af oplysningstyper) der sikrer en ensartet præsentation, men indimellem kan udløse nogle resultater der kan synes svært tilgængelige og præget af redundans. Der er altså tale om et bevidst og velbegrundet valg, hvilket der også er ved udfoldning af bøjningsparadigmer. Det har været diskuteret om man skulle gøre som CBC foreslår, nemlig føje supplerende bøjningsformer til ved de relevante betydninger (fx *fodbolde* ved den tællelige betydning af dette ord) i stedet for at indskrænke paradigmet ved de betydninger hvor det er nødvendigt. Igen er der tale om et redaktionelt valg der bygger på at bøjning i princippet afvikles i artiklens hoved, men kan kommenteres senere i artiklen, fx i form af markeringer som “især i præs. ptc. brugt som adj.” i artiklen **efterfølge** (*den efterfølgende debat, de efterfølgende dage* osv.).

Faste udtryk og kollokationer hører til de evigt tilbagevendende problemer for ordbogsmagere og -brugere. CBC gør opmærksom på den overskuelige fremstilling af de (sub)lemmatiserede flerordsudtryk i DDO, nemlig efter betydningerne og før eventuelle orddannelser, men han har jo ret i at man derved kapper forbindelsen mellem hovedbetydningerne og de dertil relaterede udtryk. Jeg kan da her oplyse at planen oprindeligt var at placere udtrykkene under de relevante hovedbetydninger, men ulemperne herved var så store at vi besluttede os for den valgte løsning. Ulemperne er bl.a. at det

kan være urimelig svært at finde et givet udtryk i en lang artikel, for slet ikke at tale om det rent praktiske forhold for leksikografen, og tilsvarende for brugeren, at det kan være vanskeligt, måske endda umuligt at placere et billedligt udtryk under en bestemt betydning. Men selvom leksikografen har skaffet sig af med det sidste problem, resterer der stadigvæk en anden vanskelig afgørelse: Hvornår er der tale om en gennemskuelig kollokation – der skal stå under en hovedbetydning – og hvornår er der tale om et ikke-gennemskueligt fast udtryk – der skal forsynes med definition m.m. og stå i den dertil indrettede del af artiklen? Det der er gennemskueligt for den ene (leksikograf eller bruger), er det ikke nødvendigvis for den anden, hvilket i øvrigt også fremgår af CBC's artikel og de eksempler han nævner.

Jeg er enig med CBC i at der er flere formål med at anføre afledninger og sammensætninger; ud over at vise orddannelsesmuligheder og supplere det behandlede ordforråd kan orddannelserne eksemplificere betydningerne på linje med kollokationer og citater. Placeringen er vi derimod ikke enige om, idet CBC ville have foretrukket at orddannelserne var placeret under de relevante betydninger og ikke i artiklens fod. Det er klart at det havde været en mulighed – som også har været overvejet – at placere orddannelserne i artiklens krop ligesom det er gjort i fx Nudansk Ordbog og i Svensk ordbok, men her har vi gerne villet undgå den redundans der ville opstå i mange artikler hvor en given orddannelse knytter sig til flere, måske alle betydninger af opslagsordet. Et par eksempler er **elendig**, hvor *elendighed* skulle nævnes under alle tre betydninger, og **grill**, hvor *grillmad* skulle nævnes under tre af de i alt fire betydninger. Redaktionen har altså vægtet to hensyn højest: dels det pladsbesparende, dels ønsket om at give en samlet fremstilling af opslagsordets orddannelsesmuligheder. Det pladsbesparende har også været afgørende for princippet om kun at anføre et bestemt orddannelseseksempel enten i kroppen eller i foden; til gengæld er det nok en forsyndelse mod redaktionsreglerne at **-bar** i betydningen 'mindre butik' illustreres ved *smykkebar* når *hælebar* er langt mere oplagt. I artiklerne **-bank** og **-blind** er der da også som "kropeksampler" valgt oplagte og frekvente ord som *blodbank*, *databank* og *ordblind*.

Betydningsforklaringer

CBC opholder sig ved brugen af synonymdefinitioner (i stedet for de traditionelle leksikografiske definitioner), og det er da heller ikke uden problemer at definere ved hjælp af synonymy idet man så at sige "sender aben videre" (for at genoptage CBC's eksempel) til et andet opslagsord i håb om at brugeren så der finder en reel definition. Det er også rigtigt bemærket at redaktionen forholdsvis sjældent betjener sig af denne definitions-måde da den indimellem også fører til cirkularitet i betydningsforklaringerne. Cirkularitet er nærmest uundgåelig, men vi har haft som princip at direkte cirkularitet af typen gal = vred = gal ikke måtte forekomme. Synonymdefinitionerne er, som CBC antyder, også netop mest blevet brugt til sekundære betydninger, og i tilfældet **abe**, betydning 2 er der nok tale om at de to synonymy "skyld; ansvar" er udtryk for både en semantisk og en stilistisk variation.

Til sidst et par ord om antallet af betydninger. Det er helt rigtigt at DDO i sammenligning med Nudansk Ordbog har mange betydninger, hvilket virker rimeligt i betragtning af de to ordbøgers størrelse og anlæg. Præcis hvor meget man skal dele et ords betydninger op, er afhængigt af den enkelte redaktørs temperament (i litteraturen tales her om "lumpers" og "splitters", altså "de der slår sammen" og "de der deler op", se fx Jackson 2002, p. 88 ff.), men en afgørende faktor – som CBC nævner – har været synonymy og antonymy, altså det forhold at et ord kan have ét synonym i én betydning og et andet i en anden betydning.

Det fremgår af CBC's artikel og af Henning Bergenholtz' oplæg ved LEDA-mødet at de opfatter DDO som mere rettet mod reception end mod produktion. Jeg

håber at fremtiden vil vise at produktionshensynet også er blevet tilgodeset; fx bør de grammatiske oplysninger om valens og om konstruktion i øvrigt komme brugere til gode der skal producere sprog. Måske kunne dette aspekt blive genstand for en debat i LEDA-Nyt eller andetsteds?

Andre omtaler

Skulle man have lyst til at læse andre omtaler og anmeldelser af DDO, vil jeg sluttelig nævne et par der er udkommet inden for det sidste halve år. For de to førstes vedkommende har medlemmer af redaktionen haft lejlighed til at kommentere anmeldelserne. På DSL's hjemmeside (www.dsl.dk/ddo_litteratur.html) findes en længere oversigt over presseomtale og litteratur.

Henning Bergenholtz og Vibeke Vrang: Den Danske Ordbog imponerer og skuffer. I: *Hermes, Journal of Linguistics*, nr. 33. Århus 2004.

Henning Bergenholtz og Vibeke Vrang: Ny dansk ordbog i seks bind for sekretærer og forskere. I: *LexicoNordica*, nr. 11. Oslo 2004.

Karen Margrethe Pedersen: Den Danske Ordbog. Udgivet af Det Danske Sprog- og Litteraturselskab og Gyldendal. Bd. 1: A-D, 2003. Bd. 2: E-H, 2004. I: *Danske Studier 2004*. København 2004.

Referencer

DDO = Den Danske Ordbog. DSL og Gyldendal. 2003-.

Jackson, Howard 2002. *Lexicography. An introduction*. Routledge.

Nudansk Ordbog = Nudansk Ordbog med etymologi. 2. udg. Politikens Forlag. 2001.

Retskrivningsordbogen. 3. udg. Alinea – Aschehoug Dansk Forlag. 2001.

Svensk ordbok. Esselte. 1986.

Dansk-Hollandsk Ordbog og Hollandsk-Dansk Ordbog

Gyldendals Røde Ordbøger. 1. udgave, 1. oplag, København 2004.

Af Inger Schoonderbeek Hansen

Sidste år udkom de længe ventede bilingvale ordbøger, *Dansk-Hollandsk Ordbog* (herafter DHO) og *Hollandsk-Dansk Ordbog* (herafter HDO) i rækken af *Gyldendals Røde Ordbøger*.¹ Ordbøgerne har været under udarbejdelse siden 1997, et år efter det fælles sprognævn for Nederland og Flandern, 'Nederlandse Taalunie', havde udliciteret projektet. DHO er udarbejdet ved Universiteit van Amsterdam (Leerstoelgroep Scandinavische Taal- en Letterkunde) under supervision af redaktørerne Annelies van Hees og Harry Perridon. HDO er udarbejdet ved Universiteit Gent (Vakgroep Scandinavistiek en Noord-Europakunde) under supervision af hovedredaktøren Godelieve Laureys. På begge ordbøger har en række forfattere, redaktører, partnere m.m. ydet deres større eller mindre bidrag, hvilket der omhyggeligt redegøres for i begge ordbøger. Ordbøgerne er hverken revision eller genoptryk, men erstatter helt eksisterende ældre og meget mindre omfattende ordbøger. Her foreligger altså nogle meget gedigne og kærkomne værker!

¹ Ordbøgerne udkom parallelt i Nederland og Belgien på forlaget Het Spectrum med titlerne: *Prisma groot woordenboek Deens-Nederlands* og *Prisma groot woordenboek Nederlands-Deens* (Utrecht 2004; pris: €49.50 pr. del; Gyldendals koster 320 DKK pr. del).

Ordbøgernes layout er i store træk ensrettet efter *Gyldendals Røde Ordbøger's* sædvanlige praksis ved bilingvale ordbøger. På indersiden af ordbøgernes omslag er der et kort over det hollandske sprogområde, der omfatter Nederland og den nordlige halvdel af Belgien, Flandern. Titelbladene med kolofon følges af et forord forfattet af ovennævnte respektive redaktører, en redegørelse for ordbøgernes brug af betegnelserne hollandsk, flamsk og nederlandsk, en vejledning i ordbogens brug (om særlige tegn, opslagsordenes rækkefølge, de forskellige ordklassers morfologi og oversættelserne), en redegørelse for den anvendte udtale (med en oversigt over udtalen af alle vokaler og konsonanter samt over de forskellige prosodiske tegn) og endelig en oversigt over redaktionelle tegn og forkortelser ('Forkortelser og betegnelser' i DHO hhv. 'Tegn og forkortelser' i HDO). De indledende afsnit optræder som paralleltekst på både dansk og hollandsk.² Herefter starter ordbogsdelen i tre spalter med opslagsordene i alfabetisk rækkefølge, A-Å (DHO) hhv. A-Z (HDO), fordelt på 843 (DHO) hhv. 1095 sider (HDO). Sidst i begge ordbøger er der optaget en liste over uregelmæssige verber i dansk (DHO) hhv. hollandsk (HDO). Bagerst på indersiden af ordbøgernes omslag ses endnu engang oversigterne 'Forkortelser og betegnelser' (DHO) hhv. 'Tegn og forkortelser' (HDO).

Ordbøgernes brugere

Ordbøgerne retter sig både mod dansk- og hollandsktalende brugere, akademiske og ikke-akademiske. I HDO beskrives brugergruppen mest udførligt, og her lægges først og fremmest vægt på ordbogens dobbelte formål. På den ene side kan HDO anvendes som produktionsordbog for hollandsktalende, fx i forbindelse med oversættelsesarbejde, og på den anden side som receptionsordbog for dansktalende, fx i forbindelse med at forstå betydningen af hollandske ord og udtryk. Begge ordbøger henvender sig bl.a. også til danskstuderende ved universiteterne i Nederland og Belgien og må kunne bruges på samme måde af hollandskstuderende i Danmark. – Ordbøgerne ville da helt klart have været til stor gavn for ca. 10-15 år siden da jeg sad og svedte over adskillige oversættelseseksaminer i Amsterdam!

DHO indeholder "lidt over" 45.000 opslagsord og "et alsidigt udvalg af kollokationer og faste udtryk" (s. 5).³ Henning Bergenholtz (Center for Leksikografi, Handelshøjskolen i Århus) har som dansk partner stået for lemmaselektionen; desuden er der anvendt danske og hollandske sprogdata-baser og internettet. Denne ultrakorte beskrivelse af ordbogens grundlag må brugeren så nøjes med. Betydelig mere information får man i HDO. Den indeholder ca. 45.500 opslagsord og knap 55.000 kollokationer, idiommer og eksempler der er hentet fra 'Referentiebestand Nederlands' (en leksikalsk database udarbejdet på initiativ af ordbogskommissionen under 'Nederlandse Taalunie', 'Commissie Lexicografische Vertaalvoorzieningen'), suppleret med både danske og hollandske sprogdata-baser og internettet. I HDO tilføjes det at der især lægges vægt på de mange kollokationer der illustrerer de hollandske ords brug i forskellige kontekster samt faste vendinger. Ifølge redaktøren er ordbogens mange brugseksempler vigtigere end antallet af opslagsord idet "de produktive orddannelsesmekanismer i [to nærtbeslægtede sprog som dansk og hollandsk] ... er ret ensartede og gennemskuelige" (s. 8).⁴

² Til sammenligning er indledningen i fx *Tysk-Dansk Ordbog* i rækken af *Gyldendals Røde Ordbøger* (14. udgave, 4. oplag, København 2003) kun forfattet på dansk – naturligvis med tyske eksempler fra ordbogsdelen.

³ Til sammenligning har fx *Tysk-Dansk Ordbog* 80.000 opslagsord og 14.000 udtryk og vendinger fordelt på 812 sider (jf. note 2 ovenfor).

⁴ I øvrigt kan man hverken i DHO eller HDO finde ordet kollokation ('collocatie', dvs. idiomatisk forbindelse mellem to eller flere ord), der både optræder i *Retskrivningsordbogen* (3. udgave, København 2001) og den hollandske ækvivalent *Woordenlijst Nederlandse taal*, også kaldt 'Groene boekje', Den lille grønne bog (3. oplag, Nederlandse Taalunie, Den Haag 1995).

Ordbøgerne omfatter “det moderne standardsprog” samt “et stort antal udtryk der forekommer i talesprog” (HDO s. 7 – i DHO beskrives ordstoffet ikke eksplicit). Mere specielle områder som fx ord fra grænseområdet mellem almensprog og fagsprog, formelle udtryk og vendinger (markeret med F for formelt sprog), navnestof, som fx navne på europæiske lande og hovedstæder, eventuelle hollandske oversættelser af danske stednavne mv., er ligeledes optaget i begge ordbøger, ord der afspejler nationale forskelle i det binationale sprogområde (her særmærkes opslagsord og udtryk med *Belg.* der er karakteristiske for Flandern, ligesom ord der kun forekommer i Nederland særmærkes med *NL*).

Hollandsk, flamsk og nederlandsk

I denne anmeldelse vil jeg ikke drøfte betegnelserne hollandsk, flamsk og nederlandsk. Ordbøgernes betegnelse ‘Hollandsk’ i titlen på Gyldendals udgave har adskillige steder været genstand for (begrundet og ubegrundet) kritik, hvilket uden tvivl må have vakt en vis undren hos danskere. Jeg vil derimod henvise til redaktøernes udmærkede, noget diplomatiske, fælles gennemgang af (og stillingtagen til) betegnelserne i indledningens kapitel ‘Hollandsk, flamsk og nederlandsk’.

I vid udstrækning følges i denne anmeldelse ordbøgernes betegnelse hollandsk for det sprog der tales i Nederland og i den nordlige del af Belgien (hvilket danskere fejlagtigt kalder ‘flamsk’). Hollandsk tales af godt 16 millioner hollændere (!) og 6 millioner flamlændere (!), desuden har det officiel status i Surinam (Nederlands tidligere koloni i Sydamerika). Landene nævnes dog ved deres rigtige navne, Nederland og Belgien (jf. ovenfor).

Vejledning til ordbøgernes brug

Afsnittet ‘Særlige tegn’ indeholder en oversigt over redaktionelle tegn der til dels er standardiseret efter praksis i rækken af *Gyldendals Røde Ordbøger*, fx ~ (DHO) hhv. - (HDO) som erstatning for opslagsordet, homonymnummer (^{1 2 3} “flere opslagsord, der staves ens”) etc. Det kan være et problem at få de specielle grafiske tegn til at slå til uden at læsningen eller brugen af ordbogen besværes. Således anvendes [] traditionelt til udtaleangivelser; i begge ordbøger bruges firkantede parenteser desuden til tre (DHO) hhv. to (HDO) andre formål, fx til at angive en definition i de tilfælde hvor der ikke kan gives en oversættelse (i begge), til at angive en supplerende forklaring til en oversættelse (DHO), til at angive en dansk fuldform af en forkortelse (HDO). Runde parenteser bruges til hele fem (DHO) hhv. tre (HDO) formål, fx til at vise et eller flere eksempler på ordets brug (i begge), til at angive at noget kan udelades (DHO). Umiddelbart kan for mange tegn, klammer, parenteser mv. forstyrre læsningen. Således skrives (*pl.*) i HDO til at angive at foregående danske ord er pluralis, mens DHO nøjes med et *pl.* uden parenteser. I HDO adskiller (*pl.*) sig imidlertid fra *pl.* ved at stå efter den danske oversættelse hvis denne afviger fra det hollandske opslagsord hvad angår numerus, fx **broek**² ... bukser (*pl.*), benklæder (*pl.*); *pl.* står derimod kun efter det hollandske opslagsord når det udelukkende findes i pluralis: **mensenrechten** *sb. pl.* (HDO). I DHO nævnes de danske substantivers talbøjning ikke eksplicit (kun ved angivelse af bøjningsendelserne, jf. nedenfor), og *pl.* optræder derfor kun i forbindelse med den hollandske oversættelse, fx **adoptivforældre** ... (-ne) adoptieouders *pl.* Tegnene ? og ? (sidstnævnte kun i DHO) omtales i afsnittet ‘Ordbogsdelen’ nedenfor.

Opslagsordenes rækkefølge er naturligvis strikte alfabetisk; der er blot nogle små forskelle ordbøgerne imellem, fx placeres stort enkeltstående bogstav foran lille i DHO, men omvendt i HDO. Nogle af forkortelserne synes at være lidt tilfældigt udvalgt, fx har DHO **A** *fork. for: ampere* A (ampère), og **E** (*fork.*) (*europavej*) E (fx de E2) eller **K** ... *kunst med stort ~ ...; Københavner med stort ~ (!)*, der ikke optræder i HDO,

men findes i hollandsk; HDO har fx forkortelsen **G-7**, der ikke optræder i DHO. I DHO er **HA** forkortelse for *handelsvidenskabelig afgangseksamen*, mens forkortelsen også for en dansker meget oplagt står for Hells Angels. Ordbøgerne indeholder en del forkortelser for titler, dog især dem der tilhører den akademiske verden, fx **cand.**, **phil.**, **dr.** (DHO), **drs.**, **ing.** (HDO) mv. Det kan dog undre at et **mag.scient.** (efter endt naturvidenskabeligt studium!) er kommet med i denne sammenhæng, men ikke fx et mag.art. eller cand.mag.

Afsnittet om ordklassernes morfologi er i begge ordbøger meget kortfattet, men meget informativt med masser af eksempler. I DHO behandles således bøjningen af substantiver, adjektiver og verber, mens HDO også tager adverbier og talord (ganske kort) med. I DHO savner man en kort gennemgang af adverbiers bøjning. For fx en danskstuderende med hollandsk som modersmål kan det nemlig være svært gennemskueligt hvornår man tilføjer *-t* når adverbiet er dannet af et adjektiv, fx *hun synger smukt*. Desuden tillader dansk retskrivning en vis valgfrihed, fx *det er dejlig(t) nemt*, og for nogle adjektiver tilføjes ikke *-t*, fx *hun synger fantastisk smukt* (jf. *Retskrivningsordbogen* § 36-39).

Det virker noget ulogisk i HDO at fordobling af konsonant efter kort vokal ved adjektiver ikke er gengivet i bøjningsendelsen, fx **rot** ... *adj. (-er/-st)* med betydningen rådden eller elendig, der bøjes 'rotte - rotter - rotst'; eksemplerne viser naturligvis bøjningen, men der er ingen eksempler ved fx ordet **eervol** ... *(-er/-st)* ærefuld, der bøjes 'eervolle - eervoller - eervolst'. Fordobling af konsonant ved fx substantiver gengives derimod, fx **kat** ... *sb. (de; -ten)* ...

DHO nævner ikke eksplicit opslagsordenes ordklasse i ordbogsdelen (jf. ovenfor); denne må brugeren slutte sig til ved at studere ordets bøjningsformer, hvilket formentlig ikke ville være et problem for især den lidt øvede danskstuderende. I HDO nævnes ikke kun opslagsordenes ordklasse, men for visse adjektivers vedkommende tilføjes desuden betegnelserne attributiv eller prædikativ (hvis adjektivets brug er begrænset til en af disse muligheder), for verbernes vedkommende angives om verbet er transitivt eller intransitivt, om verbet er hjælpeverbum, om verbet bøjes med hjælpeverbum 'hebben' (have) eller 'zijn' (være) etc. Strengt taget er det imidlertid overflødigt for HDO's oplagte brugere, danskstuderende med hollandsk som modersmål, om et verbum bøjes med 'hebben' eller 'zijn'. Det ville snarere være mere relevant at få oplyst hvilket hjælpeverbum pågældende verbs danske ækvivalent bøjes med. Denne information bør naturligvis fremgå af de anvendte eksempler, kollokationer mv. I indledningens afsnit om verbers bøjning bruger HDO eksemplet **kopen** *vb. tr. (h.)* (s. 23), men af de anvendte eksempler og kollokationer i ordbogsdelen fremgår ikke hvilket hjælpeverbum der anvendes i dansk: (fx *een huis kopen* købe et hus, *een boek kopen voor veertig euro* købe en bog for fyrre euro, *de vrijheid werd met bloed gekocht* friheden blev betalt med blod).

I afsnittet om verbers morfologi henvises desuden til listen med de uregelmæssige verber sidst i begge ordbøger. Disse kommenteres kort her. De er meget praktiske – især for hollandske danskstuderende (DHO) hhv. danske hollandskstuderende (HDO). I DHO forklares ganske kort bøjningssystemet bag præteritum og præteritum participiumsformer af svage, regelmæssige verber der optræder hyppigst og er mest produktive. I listen med uregelmæssige verber der består af to spalter, fremgår infinitiv, præsens (dog kun når den ikke bøjes regelmæssigt ved tilføjelse af *-r* efter infinitiv), oversættelsen (kursiv i parentes), præteritum og præteritum participium. Dobbeltformer som fx *fnise* – *fnes/fnisede/fniste* – *fniset/fnist*, *fare* – *for/farede* – *faret* indgår også (jf. *Retskrivningsordbogen*).

I HDO er listen med uregelmæssige verber noget længere. Ligesom i DHO omfatter den kun usammensatte og ikke-afledte verber. Af den grund er sammensatte og afledte verber i ordbogsdelen markeret med hævet prik (·), fx **aan·doen**, **toe·spreken** mv.

Listen er opdelt i tre spalter der gengiver infinitivformen med oversættelse, præteritumsformerne i singularis og pluralis og præteritum participiumsformen. Præsensformerne (singularis og pluralis) gengives ikke da de har et ret regelmæssigt mønster dannet efter infinitiv.

Af afsnittet om oversættelserne fremgår det at et polysemisk opslagsords forskellige betydninger i ordbogsdelen inddeles ved hjælp af tal. Den pågældende betydning angives i parentes på samme sprog som opslagsordet, fx **jævn**sides ... **1.** (*side om side*) zij aan zij, naast elkaar; **2.** (*parallel*) tegelijk, tegelijkertijd (DHO), **centimeter** ... **1.** (*lengte-eenheid*) centimeter ...; **2.** (*meetlint*) målebånd ... (HDO). I DHO inddeles oversættelsernes eventuelle betydningsforskelle yderligere ved hjælp af bogstaver, fx **dem** ... **1.** (*som objekt*) **a.** (*om ting*) ze, die (fx ik heb ze, die heb ik); **b.** (*om personer*) ze, F hen (fx ik ken ze/hen, die ken ik, hen ken ik niet) Efter oversættelsen tilføjes i parentes ofte et eller flere eksempler til yderligere illustration; i DHO tilføjes et eksempel på hollandsk, fx **jævne** ... **1.** (*legere*) binden (fx de saus) ..., i HDO tilføjes et hollandsk eksempel med dansk oversættelse, fx **centraal** ... central (fx *een centrale figuur/rol* en central skikkelse/rolle; *een centrale ligging* en central beliggenhed ...). Semikolon og komma adskiller mindre hhv. ringe eller slet ingen betydningsforskelle oversættelserne imellem. Som eksemplerne illustrerer bruges semikolon imidlertid også til at adskille de nummeropdelte oversættelser.

I hollandsk inddeles substantiver, ligesom i dansk, i fælleskøn (*de*-ord) og intetkøn (*het*-ord). I DHO nævnes oversættelsernes grammatiske køn normalt ikke eksplicit, dog tilføjes et *o.* ved substantiver i intetkøn, fx **pige** ... meisje *o.* Fornemmelsen for grammatisk køn er imidlertid geografisk betinget: I Flandern og i den sydlige del af Nederland er *de*-ord stadig enten maskulinum eller femininum, mens de i resten af Nederland næsten udelukkende er maskulinum. Denne opdeling har ingen betydning for valg af artikel, men den er afgørende for valg af pronomen. Af den grund tilføjer HDO *f.* bag det substantiv der kun kan bøjes i femininum, fx **muziek** ... *sb. f. uden plur. (de)* musik ... eller *f./n.* for *het*-ord, som der henvises til med 'zij' eller 'haar' (hende/deres), fx **meisje** ... *sb. f./n. (het; -s)*

Udtale / Uitspraak

Udtaleangivelserne i DHO er udarbejdet af Peter Molbæk Hansen (Institut for Almen og Anvendt Lingvistik, Københavns Universitet) og er til dels baseret på Molbæk Hansens udtaleordbog *Dansk Udtale* (København: Gyldendal 1990). I HDO er der ingen oplysninger om udtaleangivelsernes ophavssted. I begge ordbøger er alle opslagsord – bortset fra bøjningsformer og forkortelser der ikke udtales som et ord – forsynet med en udtaleform. I HDO angives geografiske udtaleforskelle mellem Nederland og Belgien ved hjælp af et *Belg.* bag den pågældende variant. Begge ordbøgers udtaleangivelser bygger på IPA-lydskriften.

I afsnittene 'Vokaler/Klinkers' og 'Konsonanter/Medeklinkers' opstilles alle lyde for dansk hhv. hollandsk belyst ved et eller flere eksempler samt ofte med en kort forklaring om hvordan pågældende lyds udtale kan sammenlignes med hollandsk hhv. dansk. Fx **u kulde** ['kul?] (als de *oe* in 'koe') (DHO), **u** som i **koe** ['ku] eller **a** som i **maat** [mat] (udtales som et langt /α/ omtrent som i 'marts') (HDO). Det er lidt besynderligt at udtaleangivelser af lange vokaler ikke markerer længde med et længdetegn, som [:] i IPA, fx [ma:t], men kun ved vokalkvaliteten, [a] over for [a]. Oversigterne giver naturligvis ikke en detaljeret beskrivelse af udtalen; en omsætning af udtaleformerne til en korrekt udtale kræver for begge ordbøgers vedkommende en bruger der er velorienteret i dansk hhv. hollandsk fonetik. I DHO savner man en særlig omtale af de karakteristiske diftonger; i HDO nævnes de blandt udtalen af vokalerne. Ofte svarer diftongens 1. komponent til en eksisterende vokal, mens diftongens 2. komponent svarer til [j] eller [v]. I

nogle tilfælde kan skriftsproget dog afvige markant fra talesproget, fx udtales talordet *seksten* ['sajsd?n], et ord som *vegne* ['vajn?]. Et par linjer om diftonger ville have været tilstrækkeligt.

I DHO anføres både hoved- og bitryk, mens der i HDO kun anføres hovedtryk. I DHO gælder dette alle ord, mens det i HDO gælder alle flerstavellesord; ved enstavellesord anføres ikke trykstreg, fx **doop** [dop] ... dåb. Trykgengivelse på andre stavelser end første led er i danske øjne noget mærkværdigt placeret, nemlig ikke foran den trykbærende stavelse, men umiddelbart foran den trykbærende vokal, fx efter trykssvag forstavelse i **bewonderaar** [b?w'cnd?rar] ... beundrer, i sammensætninger som fx **autovrij** [otovr'ei] ... bilfri mv.

Forkortelser og betegnelser (DHO) / Tegn og forkortelser (HDO)

Redaktionelle forkortelser, betegnelser, tegn mv. er overvejende dannet med udgangspunkt i ordbogens danske udgave og er således nogenlunde ensrettet efter praksis i rækken af *Gyldendals Røde Ordbøger*. De fremgår af en liste med en dansk beskrivelse og en hollandsk oversættelse. Listerne er som nævnt ovenfor også trykt bagerst i ordbøgerne på indersiden af omslaget, hvilket er meget brugervenligt.

Langt overvejende svarer forkortelserne altså til danske betegnelser; der er dog nogle få undtagelser, fx svarer *o.* i DHO til intetkøn (hollandsk: 'onzijdig'), men dette står udelukkende efter en hollandsk oversættelse; *bep.* i HDO svarer til bestemt (hollandsk: 'bepaald'); det internationale *NL* optræder i begge ordbøger med den danske beskrivelse 'Holland', i DHO desuden 'hollandsk', og den hollandske oversættelse 'Nederland', i DHO desuden 'Nederlands'. Især de groteske enkeltbogstaver *F*, *T*, *S*, *V* (sidstnævnte to kun i HDO), der betegner de vigtigste stilistiske og pragmatiske områder formelt sprog, talesprog/uformelt sprog, slang hhv. vulgært sprog, træder tydeligt frem i de enkelte artikler.

HDO indeholder mere end dobbelt så mange redaktionelle tegn og forkortelser som DHO (114 hhv. 52), heriblandt en del forkortelser og betegnelser for specielle semantiske områder, fx *agr.* 'landbrug' (agratiske betegnelser), *billedkunst*, *fodb.* 'fodbold', *håndværk*, *pengv.* 'pengevæsen', *med.* 'lægevidenskab', *sold.* 'soldatersprog' mv. I øvrigt adskiller fællesbetegnelserne i DHO og HDO sig ikke meget fra hinanden, fx forkortes i DHO betegnelser der vedrører de store trosretninger (*kat.*, *prot.*), mens de skrives fuldt ud i HDO (*katolsk*, *protestantisk*, desuden *jødisk*).

Ordbogsdelen

Informationerne i de enkelte artikler i ordbogsdelen er ordnet i rækkefølgen:

Opslagsord (i fed og lettere udtrykket), evt. homonymnummer (hævet), udtale (i firkantet parentes), ordklassebetegnelse (kun i HDO), bøjning (i kursiv i rund parentes), evt. et tal der angiver betydningsinddelingen (i fed), evt. nærmere betydningsforklaring af opslagsordet (i kursiv i rund parentes), oversættelse (i antikva), evt. bøjningsoplysningen *o.* (kun i DHO), evt. videre betydningsoplysning (i kursiv i rund parentes), evt. eksempel på målsproget (i antikva i rund parentes), evt. kollokationer, udtryk mv. (markeret ?) der indeholder kildesprogets eksempler (i kursiv) og oversættelser deraf (i antikva), evt. ordsprog (markeret ?, kun i DHO). I HDO fremtræder kollokationer, udtryk mv. mindre tydeligt, idet de oftest bliver bragt som eksempler i direkte forbindelse med opslagsordets danske oversættelse (i kursiv i rund parentes med efterfølgende dansk oversættelse i antikva).

top¹ ['tɒp] (-pen, -pe) 1. top (fx van golven, de top bereiken); 2. (snur-retop) tol; 3. (lille paryk) toupetje o., haarstukje o.;
 □ *det er ~pen* (het is) geweldig!, (het is) te gek!; *~pen af isbjerget* het topje van de ijsberg; *fra ~ til* tå van top tot teen; *være i ~* op en top zijn; *i ~pen af* in de top van (fx de boom); *sætte håret op i en ~* je haar opsteken; *med ~ på* met een kop (fx een lepel met een kop); *en skefuld med ~* een volle lepel, een schepje met een kop; *på ~pen af a.* op de top van (fx de berg); *b. (overført)* op het hoogtepunt van (fx je carrière); *holde sig på ~pen* aan de top blijven; *gå til ~s a.* in top gaan (fx de vlag ging in top); *b. (om status, karriere o.l.)* aan de top komen; *o Der er koldt på toppen.* Het is eenzaam aan de top.

ring [rɪŋ] sb. (de; -en) 1. (sieraad) ring (fx een ring om/aan zijn vinger hebben have en ring på fingeren; *een ringetje in het oor* en ring i øret); 2. (cirkelvormig ding) ring (fx de ringen van een boom et træs ringe; *de ringen van Saturnus* Saturns ringe; *oefeningen aan de ringen* øvelser i ringene; *ringen om de ogen* hebben have (sorte) rande under øjnene); 3. (arena) ring (fx de bokseksamen i ringen; *de handdoek in de ring* werpen kaste håndklædet i ringen); 4. (rondweg) (Belg.) ringvej (fx *de ring rond Brussel* ringvejen rundt om Bruxelles);
 □ *er uitzien om door een ringetje te halen* være nydeligt/pænt klædt på.

Artiklerne **top**¹ fra DHO og **ring** fra HDO

Opslagsordet og tallene mellem væsentlige betydningsforskelle skrives med fed skrift, hvilket giver den enkelte artikel et brugervenligt, letlæseligt fremtrædende – det er også de informationer der umiddelbart er vigtigst for brugeren. En del artikler indeholder udelukkende information om opslagsordets udtale, ordklasse (kun i HDO), bøjningsform(er) og oversættelse, mens hovedparten indeholder fra ét eksempel til rigtig mange eksempler på ordets brug, kollokationer, ordsprog mv. Her ligger helt klart ordbøgernes force.

De to ordbøgers præsentation af eksempler, kollokationer, faste udtryk, ordsprog mv. er temmelig forskellig, i hvert fald for hovedparten af artiklernes vedkommende. I DHO er artiklen nærmest todelt: Den første del indeholder 'basisoplysninger' (opslagsform, bøjning, evt. betydningsdifferentiering, oversættelse med et enkelt eller to eksempler), den anden del indeholder udvidede informationer om opslagsordets anvendelse i form af alskens kollokationer, udtryk mv. og evt. ordsprog; denne del indledes med de særlige tegn ? hhv. ? (jf. artiklen **top**¹ ovenfor). Denne inddeling bevirker en forholdsvis let orientering i artiklens informationer, og det giver samtidig brugeren mulighed for selv at vælge fra, fx ved straks at lade øjnene glide fra opslagsordet til ordets brug i forskellige kontekster.

I HDO udgør artiklens basisoplysninger og tilhørende kollokationer, udtryk, ordsprog mv. snarere en enhed. I de lidt længere artikler (jf. ovenfor) optræder – i rund parentes – eksempler, kollokationer mv. i direkte forlængelse af opslagsordets nummerinddelte oversættelser. Opbygningen har store ligheder med den i de hollandske *van Dale*-ordbøger. Denne type præsentation kan ved nogle af de 'store' ord (som fx præpositionerne) give lange, lidt overskuelige artikler, hvor det kun er de fede tal der særmarkerer opslagsordets væsentlige betydningsforskelle, og som er brugerens eneste mulighed for at 'zappe' igennem artiklen. Sjældnere adskilles faste forbindelser mv. fra artiklen ved tegnet ?, der ifølge indledningen "markerer indledning på afsnit med faste udtryk og vendinger" (s. 19, eller ifølge afsnittet 'Tegn og forkortelser': "markerer begyndelsen af afsnit med udtryk og idiomer", s. 34). I nogle tilfælde er det svært at se hvornår nogle faste udtryk og vendinger er bragt efter ? og hvornår ikke. Efter ? kan stå faste forbindelser som fx *ridder in de orde van Oranje-Nassau* ridder af Oranje-Nassau-Ordenen; *de orde van advocaten* advokatsamfundet (**orde**²), talemåder som *paal en perk stellen aan iets* sætte en stopper for noget, sætte en grænse for noget (**perk**); og i artiklen **recht** optræder eksemplerne *het recht van de sterkste* jungleloven; ... *zijn rechten laten gelden* kræve sin ret, mens der i samme artikel efter ? optræder *tot zijn recht* komen komme til sin ret.

Det er meget vanskeligt at inddele faste ordforbindelser som kollokationer, udtryk, talemåder, ordsprog, idiommer mv. problemfrit, og området er langt fra færdigdebat- teret (jf. rapporten fra DANFRAS på side 8 ff. i dette nummer af LEDA-Nyt). Også i DHO afspejler denne problematiske inddeling sig i eksemplerne efter ? og ?. Således bringes efter ? fx *jeg synes jeg hørte en loppe gø* (**loppe**), *det koster ham kun et greb i lommen* (**lomme**), *kæmpe med ryggen mod muren* (**mur**), mens der efter ? optræder fx *livet er ikke lutter lagkage* (**liv**²), *kundskab er magt* (**magt**), der efter min mening adskiller sig fra ordsprog som *Som man råber i skoven, får man svar* (**svar**), *Brændt barn skyer ilden* (**barn**). Desuden er placeringen af ordsprogene i DHO ikke i alle tilfælde gennemskuelig. Så vidt muligt må det første substantiv i et ordsprog være afgørende for ordsprogets placering, men her ses fx at *Den, der ager med stude, kommer også med* kun står under **age**, mens *Brændt barn skyer ilden* imidlertid både står under **barn**, **sky**⁴ og **ild**.

Lige bortset fra det besværlige omkring disse faste forbindelsers afgrænsning og rette placering, er det et meget stort plus for begge ordbøger at der er bragt så mange af dem! Både i DHO og HDO præsenteres således ordenes mange kombinationsmulig- heder og dertil ækvivalente oversættelser. I DHO er desuden præpositionen i de mange præpositionsforbindelser markeret med fed skrift, hvilket igen bidrager til overskuelig- heden (jf. artiklen **top**¹ ovenfor).

Ovenfor nævnte eksempler og påpegede småfejl er (indrømmet) tilfældige nedslag i ordbøgerne, for nok har jeg studeret ordbøgerne nøje, men jeg har ikke læst hver enkelt artikel i dem (endnu). Enkelte oversættelser gav mig anledning til undren. Således over- sættes **kommafejl** i DHO med fejl i interpunktionen ("fout in de interpunctie"). En kom- mastrid som den danske er jo også særegen, hvorfor ordet ikke altid kan forventes at ha- ve en præcis ækvivalent på andre sprog. Ordet *ijdtluit* svarer til en forfængelig person der tilbringer mange timer foran spejlet, men i HDO oversættes det med 'pralhals'. Ordet *hjemmeside* er bogstaveligt oversat med 'thuispagina'; 'website' er ikke nævnt, men bruges fx på 'Nederlandse Taalunie's hjemmeside. Hverken *thuispagina* eller *website* optræder i HDO. Det ifølge danskere meget typisk danske ord (!) *hyggelig* oversættes i DHO med den gode hollandske ækvivalent 'gezellig', der også optræder i HDO. Til det virkelig typisk danske ord eller rettere begreb *janteloven* gives i DHO en god beskrivelse og en passende hollandsk gengivelse.

Ordbøger (såvel mono- som bilingvale) bliver ofte kritiseret for deres medta- gelse og udeladelse af opslagsord. Det er også her DHO og HDO skal stå deres prøve! Redaktionen af DHO og HDO blev påbegyndt i 1997, og lemmaudvælgelsen er sand- synligvis baseret på databaser der stammer fra dette tidspunkt og/eller tidligere. Brug- en kan således ikke forvente at finde ord som *dvd*, *dvd-afspiller*, *mp3*, *website* m.fl., mens eksemplet *fås på cd og mc* (mon musikkassetter stadig fås?) under opslagsordet **cd** nok ville blive erstattet af fx *fås på cd og dvd*. DHO har også **cd-plade**, altså selve skiven, og **cd-skifter**, men kun de hollandske ord 'cd-rom' og 'cd-rom-brander' som eksempler. Dette problem gælder ikke specielt for DHO og HDO, men naturligvis for alle trykte ordbøger hvis redaktion strækker sig over en længere årrække. Egentlig kræver de en hurtig genoptrykning og opdatering med nye og helst endnu flere opslagsord og kollo- kationer. Det kræver naturligvis både penge og tid, men ordbøger kan med stor fordel også gøres tilgængelige via internettet og således uden større problemer løbende opdateres.

Dansk-Hollandsk Ordbog: Udgangspunkter og målsætninger

Af Harry Perridon

I anledning af Inger Schoonderbeek Hansens anmeldelse af ordbøgerne skal jeg her gå ind på de udgangspunkter og målsætninger vi havde, da vi begyndte arbejdet med den dansk-hollandske del. Målene vi ville opnå, var først og fremmest følgende: (a) ordbogen skulle være bidirektionel, dvs. den skulle kunne bruges af både dansk- og hollandsksprogede; og (b) den skulle være så brugervenlig som muligt.

Fordi den person som bruger en ordbog til at producere tekster på det fremmede sprog, har behov for meget mere detaljerede oplysninger om hvordan et givet ord eller udtryk fungerer i målsproget, end den som bare vil vide hvad et ord eller udtryk betyder i den tekst man læser, har vi valgt de danske brugere som den målgruppe vi skulle tage mest hensyn til. En vigtig følge af dette valg var at vi besluttede at den semantiske inddeling af et (dansk) lemma skulle styres af de mulige oversættelser til målsproget (hollandsk), og ikke af nogen kvasi-universel semantik der deler det ind i betydningselementer uden at der bliver taget hensyn til om de betydninger man på denne vis finder frem, leder til forskellige oversættelser. Lad mig give et eksempel. Det danske ord *bånd* har ifølge Politikens Store Nye Nudansk Ordbog (1996) tre hovedbetydninger: 1. et langt, smalt stykke stof el. andet materiale som kan bruges til at binde med; <i sammen-sætn.> en lang genstand der er sammenstykket af enkelte led, mest i metal; 2. langt, smalt, fladt, magnetiserbart materiale til lagring af oplysninger; fx billeder og lyd; 3. stærk følelse af fællesskab. Denne inddeling kan muligvis forsvares i en ensproget ordbog, skønt det ikke er selvindlysende hvorfor fx *hårbånd* og *samlebånd* tilhører samme gruppe, og *lydbånd* og *videobånd* en anden gruppe. Den Danske Ordbog (2003) for eksempel henfører Nudansk Ordbogs betydning 3 'stærk følelse af fællesskab' til betydning 1, har samme gruppe 2, men udskiller en ny betydning 'vandret, bred rem beregnet til at transportere genstande eller varer fra et sted til et andet idet den kan bevæges fremad.' I en tosproget ordbog derimod er en sådan inddeling ikke på sin plads idet det ikke giver noget fingerpeg om hvordan ordet skal oversættes. Det danske *bånd* svarer til *band*, *lint* eller *touw(tje)* i hollandsk, men ikke på den måde at de betydninger Nudansk Ordbog eller Den Danske Ordbog udskiller, nøje svarer til de tre hollandske ord. I langt de fleste tilfælde kan *bånd* oversættes med *band*, men ikke når det betyder 'meget smalt og langt stykke stof e.l. der bruges til at binde noget med el. til pynt', som Den Danske Ordbog betragter som ordets oprindelige betydning. I dette tilfælde er *touw(tje)* ('til at binde noget med') eller *lint* ('til pynt') det rigtige ord, men kun hvis *bånd* ikke er brugt i overført betydning. I Dansk-Hollandsk Ordbog (herefter DHO) bliver det så til:

bånd ['b? n'] (-et, -) 1. band (fx iets op een band opnemen, de banden verbreken, aan de lopende band); 2. (bændel) lint o.; 3. (snor) touw o.;

At det hollandske ord *band* kan bruges i mange af de samme kontekster som det danske *bånd*, bliver tydeliggjort gennem de eksempler der står i parentes.

Vi mener at det også for udforskningen af det danske sprogs semantik er vigtigt at se hvordan det reagerer når det bliver konfronteret, eller måske sågar 'beskudt', med et andet sprog. Enhver sådan konfrontation vil give os oplysninger om hvordan danske ord er opbygget, og hvordan de forholder sig til ord og udtryk i andre sprog. Det er måske ikke helt som Hjelmslev siger i *Omkring sprogteoriens grundlæggelse* (1966, s. 48), hvor han sammenligner de forskellige måder sprogene beskriver virkeligheden på: "Det er som en og samme haandfuld sandskorn, der formes i vidt forskellige mønstre, eller som den sky paa himlen, der for Hamlets blik skifter skikkelse fra minut til minut. Ligesom de samme sandskorn kan lægges i forskellige mønstre, og den samme sky antage stadig nye skikkelser, saaledes er det ogsaa den samme mening, der formes eller

struktureres forskelligt i forskellige sprog”. Det er jo ikke på forhånd givet at det faktisk er den samme virkelighed der bliver beskrevet af de forskellige sprog. Men det er i hvert fald klart at ethvert sprog fremhæver, eller snarere lader os fremhæve, andre detaljer i den sociale virkelighed vi lever i. Tosprogede ordbøger, og især de der i den semantiske analyse af lemmaerne *ikke* udgår fra at der findes et sæt universelle betydningspartikler (“de samme sandskorn”), har en stor værdi for studiet af de forskellige sprogs betydningside idet de på en håndgribelig måde tydeliggør at også betydninger er konventionelle og derfor må læres.

Hvad er et lemma?

Et spørgsmål man må tage stilling til når man laver en ordbog, er ‘hvad er et lemma?’. Det synes at være selvindlysende at et lemma må være et ord i en ordbog, men så enkelt er det alligevel ikke. Man plejer for eksempel at betragte participierne som verbalformer, og de optages derfor almindeligvis ikke som opslagsord i ordbøger; men hvis de er afledt af en ordgruppe, fx af et løst sammensat verbum som *hænge sammen*, har vi det problem at det participiet skulle være en form af, ikke er et verbum, men netop en ordgruppe: *sammenhængende* er ikke en sammensætning af *sammen* og *hængende*, men snarest en samdannelse af *sammen* + *hæng* + *ende*. Det såkaldt løst sammensatte verbum *hænge sammen* får i den danske tradition, i modsætning til den svenske, ikke et eget lemma, men behandles som en kollokation. I hollandske og tyske ordbøger er alle løst sammensatte verber lemmaer, hvilket utvivlsomt har sin årsag i at opslagsformen, infinitiven, skrives som ét ord, fx holl. *samenhangen*, ty. *zusammenhängen*.⁵ Men i mange af de kontekster hvor ordet/ordgruppen forekommer, står de to dele *samen* og (en form af) *hangen* ikke i denne rækkefølge og er adskilt fra hinanden af et eller flere ord, fx:

alles <i>hangt</i> met alles <i>samen</i>	
alt hænger med alt sammen	‘alt hænger sammen med alt’

Det er ikke bare biord/partikler som *in* ‘ind’, *uit* ‘ud’, *samen* osv. der i den hollandske retskrivning betragtes som en del af infinitiven, også tillægsord og navneord kan forekomme som første led af en infinitiv, fx: *schoonmaken* ‘gøre rent’ (vs. *hij maakt de stal schoon* ‘han gør rent i stalden’), *koffiedrinken* ‘drikke kaffe’ (vs. *ze drinken ’s ochtends altijd koffie* ‘de drikker altid kaffe om morgenen’). Hvorfor visse kombinationer skrives som et ord og andre som to, er ofte fuldstændig uklart, man skriver fx *koffiedrinken* og *theedrinken* som et ord, men *bier drinken* og *water drinken* som to. For den danske bruger af Hollandsk-Dansk Ordbog kan det derfor være svært at vide hvilket ord han/hun skal slå op, særlig når afstanden mellem delene af den ‘løse sammensætning’ er stor.

Som det fremgår af ovenstående er grænsen mellem morfologi og syntaks ikke skarp, hvilket betyder at man som ordbogsforfatter egentlig kunne drage den selv. Man kan imidlertid ikke se bort fra at der findes en vis tradition i både den danske og den hollandske leksikografiske praksis, der generelt set går ud fra at kun det som skrives som ét ord, er et lemma, og at det som skrives som to eller flere ord, er en ordforbindelse. Om end jeg personligt synes at den svenske praksis med at opføre løst sammensatte verber som selvstændige lemmaer er mere brugervenlig (ordartiklerne bliver betydelig mindre, hvilket gør at man nemmere finder det udtryk man søger efter) har vi alligevel valgt at følge den tradition de fleste danskere er vant til. Som kompromis er disse ‘løse sammensætninger’ – i det mindste i de længere ordartikler – markeret med fed skrift så de ser ud som en slags sublemmaer:

⁵ Mærkeligt nok skrives infinitiven med infinitivmærket *te* aldrig som ét ord, fx: *het schijnt ermee samen te hangen (dat)* ‘det ser ud til at hænge sammen med (at)’

hale 2 [ha:l?] (-ede, -et) halen (fx een drenkeling aan boord halen);

? ~ i land **a.** binnenhalen (fx een order binnenhalen); **b.** (gøre indrømmelser) terugkrabbelen; få et forlig ~t i land een compromis bereiken; ~ torsk i land zagen, hard snurken; ~ en line ind een lijn binnenhalen; ~ ind på nogen iemand inhalen; ~ ned neerhalen (fx een zeil, een vlag); ~ op i bukserne je broek optrekken; ~ ud af trekken/halen uit; jeg kunne ikke ~ et ord ud af hende ik kon geen woord uit haar krijgen.

Mange af de andre ordforbindelser der står opført efter tegnet ?, kunne også betragtes som 'løse sammensætninger', men dette strider endnu mere mod gængs praksis. Derimod var den forskel i status der på denne måde etableres mellem ægte og uægte sammensætninger, ikke udslagsgivende for beslutningen om et vist ord eller en vis ordforbindelse skulle optages i ordbogen. Her brugte vi de samme kriterier: Hvis betydningen af den ægte eller uægte sammensætning er fuldstændig determineret af de led der indgår i sammensætningen eller ordforbindelsen, bliver ordet/frasen i princippet ikke optaget.

Grammatiske angivelser

I og for sig er det ikke tilrådeligt at bruge alt for meget lingvistisk metasprog i ordbøger. Det virker fremmedgørende for de fleste brugere, der enten har glemt eller fortrængt den grammatik de lærte i folkeskolen. Desuden giver det næppe nogen brugbar information, det er for eksempel ikke let at se hvilken glæde man har af at ordbogen oplyser at det og det ord tilhører den og den ordklasse. Den information er ikke tilstrækkelig når man vil danne ordets øvrige former, og irrelevant hvis ordet kun forekommer i én form: For brugeren er det lige meget om fx ordet *stå* i udtrykket *gå i stå* kaldes et navneord eller et udsagnsord. I praksis kan man derfor godt nøjes med at angive de former som lemmaerne forekommer i. Hvordan lemmaet forbindes med andre ord, fremgår så af de eksempler der gives.

I DHO har vi så vidt muligt forsøgt at spare på brugen af et lingvistisk metasprog, men i visse tilfælde viste traditionen sig at være stærkere end læren, som fx i det ovenfor nævnte eksempel (*gå i*) *stå*, der er forsynet med det ret meningsløse (*sb.*).

Slutbemærkninger

Enhver der har lavet eller er i færd med at lave en ordbog, ved at arbejdet aldrig er færdigt. Sjældent er man helt tilfreds med resultatet når det foreligger i trykt form: "Hvis vi bare havde haft lidt mere tid (og penge) var det hele blevet (endnu) bedre."

Man opdager at man ikke har været helt så konsekvent som man troede man var, at der stadigvæk forekommer fejl som man troede man havde rettet, men formodentlig i den forkerte fil, og at visse udtryk og ord er blevet umoderne og er afløst af andre, som man desværre ikke har fået med. Sådan er det også med denne ordbog, hvilket til dels skyldes den ret lange tid det tog at producere den. Vi påbegyndte arbejdet i 1997, i slutningen af 2001 blev redaktionen afsluttet, og filerne sendt til *De Taalunie*, hvor de blev liggende til 2003 da Gyldendal tog hånd om produktionen af ordbøgerne. I mellemtiden var Holland og Belgien gået over fra gylden og frank til euro, båndoptagere og kassettebånd, som Inger Schoonderbeek Hansen påpeger, var blevet museumsgenstande, samfundet i de tre lande havde ændret sig på mange måder osv. osv.

Det ser derfor ud til at den eneste mulighed for at holde en ordbog rigtigt up to date er at udgive den som e-bog. Men en sådan e-udgave kan efter min mening aldrig helt erstatte en 'rigtig', trykt ordbog.

En polyfon ordbogsrecension

Af Godelieve Laureys

Da Inger Schoonderbeek Hansen kontaktede mig i forbindelse med en anmeldelse af ordbøgerne dansk-hollandsk/hollandsk-dansk, blev jeg glad. Det er vigtigt for enhver forfatter at få feedback på det man har skrevet og det gælder især et så komplekst og samtidig så brugerrettet produkt som en ordbog. At min amsterdamske kollega og jeg desuden som forfattere fik lejlighed til at gennemlæse recensionen og komme med kommentarer, synes jeg er storartet, idet man på den måde undgår direkte misforståelser eller fejltolkninger, som det ville være synd at sende ud i verden. At vi derudover blev bedt om selv at skrive et indlæg for at redegøre for vore egne erfaringer med projektet, er enestående. Værket bliver således belyst fra forskellige synsvinkler og læseren kan forhåbentlig danne sig et nuanceret billede på baggrund af denne polyfone tekststart.

Baggrund

En omfattende og professionel ordbog for sprogparret dansk/hollandsk har længe været savnet. En moderne professionel bilingval ordbog er korpusstøttet og kendetegnes af en opdateret makrostruktur og en rig og kontrastivt relevant mikrostruktur. Indtil for nylig var man for det pågældende sprogpar imidlertid henvist til lommeordbøger¹, eller man var nødt til at gøre brug af et dobbelt sæt af ordbøger for at slå noget op: fx først hollandsk-engelsk og så engelsk-dansk. Geerte De Vries' mellemstore ordbog², der udkom i 1971 og blev genoptrykt i 1986, dækkede ganske vist behovet i nogen tid, men denne ordbog har en meget begrænset mikrostruktur og er desuden forældet og for længst udgået. "Paludans hollandsk-dansk, dansk-hollandsk ordbog" (1 bind, 12.000 opslagsord), der udkom 1997, er et uændret genoptryk af Van Goor's lommeordbog.

Da CLVV (= ordbogskommissionen, der sorterer under det fælles hollandsk-flamske sprognævn, Nederlandse Taalunie) i 1996 udliciterede en hollandsk-dansk/dansk-hollandsk ordbog, ansøgte både det Nordiske Institut ved Universitetet i Gent (Belgien) og Institut for Skandinavistik ved Amsterdams Universitet (Holland) om hele projektet. Opgaven blev imidlertid fordelt, således at Gent fik den hollandsk-danske del og Amsterdam den dansk-hollandske del. Umiddelbart syntes det for os i Gent at være en meget udfordrende og svær opgave at stå for den del der oversætter fra kildesproget hollandsk til målsproget dansk. Jeg begyndte da også med at sammensætte en redaktionsgruppe, der bestod hovedsagelig af hollandskkyndige dansksprogede redaktører suppleret med nogle danskkyndige hollandsksprogede medarbejdere. Nogle af redaktørerne er tosprogede, alle har læst dansk og hollandsk på akademisk niveau.

HDO blev udarbejdet i Gent i perioden oktober 1997 til sommeren 2002.

En kontrastiv hollandsk-dansk leksikalsk database og en bilingval ordbog

Udvælgelsen af hollandske opslagsord og eksempler er foretaget ud fra en relationel leksikalsk database, Referentiebestand Nederlands (RBN), der er udarbejdet på initiativ af CLVV. RBN omfatter ca. 45.000 opslagsord med et stort antal eksempler og bygger på et korpus af både skrift- og talesprog (5 millioner ord-korpus udarbejdet af INL ("Instituut voor de Nederlandse Lexicografie" ved Leiden Universitet)) og på en eksisterende relationel database på 21.000 ord udarbejdet af Instituttet for Leksikologi ved Vrije Universiteit Amsterdam.

¹ Van Goor's Mini-pockets Deens woordenboek, 1995/13 (12.000 artikler); Wolter's Miniwoordenboek Deens-Nederlands, Nederlands-Deens, 1999/5 (10.000 artikler);

² De Vries, Deens-Nederlands Woordenboek

RBN omfatter forskellige oplysningstyper. Ethvert opslagsord er opmarkeret ved et antal kategorier: stavemåde, udtale, morfologi, syntaks, semantik, pragmatik og kombinatoriske egenskaber (kollokationer, valens m.m.). RBN bygger på to koncepter: "form unit" (FU), der nærmest svarer til opslagsordets overfladestruktur og "lexical unit" (LU), som henviser til en vis betydning af en FU. Ved monosemiske opslagsord sammenfalder FU og LU virtuelt, men ved polysemiske opslagsord er det de forskellige betydningsvarianter (LU), der gælder, i og med at hvert LU oversættes separat og linkes til en tilsvarende betydningsvariant i målsproget. Ethvert LU specificeres i databasen ved hjælp af en kort semantisk omskrivning (diskriminator) og en ordbogsdefinition. Disse data er til stor hjælp for ordbogsredaktørerne, når de skal finde den rette oversættelsesækvivalent til en bestemt LU. Ved polysemiske opslagsord repræsenteres den korte betydnings specificifikation i ordbogen, ved monosemiske ord ligger denne information i databasen, men den repræsenteres ikke i ordbogen.

Ordbogsdefinitionerne er mere udførlige og svarer til definitionerne i en mellemstor forklarende ordbog. Disse er til stor hjælp for redaktørerne under arbejdet, men kommer naturligvis ikke med i ordbogen.

Opdelingen af lemmaerne er baseret på rent formelle, dvs. kategoriale eller subkategoriale, kriterier. Den mindste forskel i det morfologiske paradigme giver anledning til et nyt lemma (se også Morten Winklers bidrag).

For hvert lemma angives ordklasse og morfologisk type, fx angives det ved substantiver, om der er tale om et simpleksord, en sammensætning eller en afledning, og der angives pluralisendelse; ved verber angives bøjningsklasse og typologi. Den syntaktiske beskrivelse af et opslagsord omfatter oplysninger om ordets kombinatoriske egenskaber, fx angives ved substantiver genus og artikelbrug; ved verber, visse adjektiver og substantiver angives valensstruktur og faste præpositioner; ved adjektiver angives, om de kan bruges prædikativt og/eller adverbialt, og om de kan kompareres. I det semantiske felt får opslagsordet en kode, der angiver semantiske tilhørsforhold, ved substantiver angives bl.a. semantisk type (fx abstrakt, konkret), leksikaliseret metonymi (fx dynamisk, ikke-dynamisk). Det pragmatisk felt omfatter syv forskellige etiketter, der indkredser ordets brugsområde: fx angives, om et ord er et fremmedord, et regionalt ord, om det har en vis konnotation, om det tilhører et socialt register, og om det er tidsdateret. Systemets finmaskede struktur illustreres ved at etiketten "konnotation" igen er repræsenteret af fem forskellige markører: "neutral (default), offensive, pejorative, euphemistic, jocular".

Den hollandske database indeholder mange informationstyper, der ikke alle er repræsenteret i den grafiske ordbog (se også Morten Winklers bidrag). Det har været vores hensigt fra starten at udarbejde en hollandsk-dansk kontrastiv leksikalsk database, hvorefter en hollandsk-dansk ordbog, men ikke kun den, kunne blive afledt. Den kontrastive database, der blev udarbejdet inden for projektets rammer i Gent, rummer mange muligheder for fremtidig forskning. De danske opslagsord er lagret i en relationel database af samme format og specificationsdybde som RBN, hvilket indebærer at vi har oprettet LU's for alle danske ord og tilføjet en kort semantisk omskrivning af enhver betydningsvariant og indført de nødvendige morfologiske, semantiske og pragmatisk opmarkeringer. Den danske del af databasen kan således også forsøges linket til en af de andre målsprogsdatabaser der er fremstillet i RBN-format, fx finsk, portugisisk eller græsk.

Denne fremgangsmåde er helt i overensstemmelse med de to målsætninger som Nederlandse Taalunie har med RBN: en praktisk-politisk og en videnskabelig. Den praktiske målsætning er at stille en hollandsk leksikalsk database til rådighed for et stort antal bilingvale ordbogsprojekter; den videnskabelige målsætning er udarbejdelsen af en multi-

lingval, flersproget database, hvor hollandsk fungerer som en hub. Ud af denne flersprogede parallelle database vil man så kunne aflede råmateriale for nye sprogpar (se også Morten Winklers bidrag).

I denne henseende skal også påpeges at HDO udgår fra helt andre redaktionsprincipper end DHO. HDO udgår klart fra en hollandsk database og dette indebærer at opdelingen i betydningsvarianter (LU) af de hollandske opslagsord ikke primært styres af oversættelserne, men genspejler den semantiske analyse af kildesprogsordet. Selvfølgelig er der også taget hensyn til den kontrastive relevans. Det har fx ingen mening at opremse alt for mange betydningsspecifikationer der alle oversættes med det samme opslagsord på dansk.

Ordbogens profil

Ordbogens redaktører har fra starten lagt stor vægt på et alsidigt udvalg af eksempler og kollokationer. Det skyldes, at værdien af en ordbog mellem to nært beslægtede sprog som dansk og hollandsk ikke så meget ligger i et stort antal opslagsord, idet de produktive orddannelsesmekanismer i begge sprog er ret ensartede og gennemskuelige. Derfor er der i højere grad givet plads til et stort antal brugseksempler, der viser, hvordan et ord bruges i en kontekst eller indgår i faste vendinger.

Ordbogen behandler det moderne standardsprog. Derudover indeholder den et stort antal udtryk, der forekommer i talesprog, ligesom den rummer mange ord og udtryk fra grænseområdet mellem almensprog og fagsprog samt en del vendinger, der hører ind under det, man vil betegne som formelt sprog. Sådanne opslagsord og eksempel-sætninger er forsynet med en stilmarkør. Desuden er der, især som hjælp for danske brugere, optaget et antal opslagsord og udtryk, som udelukkende eller fortrinsvis forekommer i Flandern; i så fald er de forsynet med en geografisk markør.

I og med at det nederlandske sprog er et binationalt sprog bliver dette princip også brugt i det nederlandske kildemateriale for at skelne mellem kulturspecifikke ord og termer, der udelukkende forekommer i enten Holland eller Belgien. Der bruges forskellige etiketter for de såkaldte belgicisme, dvs. ord, der kun bruges i den nederlandsktalende del af Belgien, men ikke i Holland, fx: 'panikeren' (= 'in paniek raken') eller 'microgolfoven' (= 'magnetron'), og for de såkaldte samfundsord, dvs. ord der betegner institutioner eller samfundsinstanser i resp. Holland og Belgien. I det senere tilfælde får også den hollandske ækvivalent en hyponymmarkering (fx 'wethouder' (i NL) vs. 'schepen' (i B), der begge skal oversættes med 'byrådsmedlem').

I parentes bemærket er det en skam, at betegnelsen 'nederlandsk' er så lidt brugt i Danmark og at der er så stor uvidenhed omkring det forhold at 'flamsk' og 'hollandsk' er et og samme sprog. Da vi diskuterede ordbogens titel med forlaget – og dét har vi gjort – har udgangspunktet dog klart været at den gængse danske sprogbrug skulle respekteres.

DHO og HDO retter sig mod både hollandsksprogede og dansksprogede brugere. For den hollandsksprogede bruger fungerer HDO fortrinsvis som en produktionsordbog og for den danske bruger fungerer den fortrinsvis som en receptionsordbog. Denne dobbelte målsætning har medført en række redaktionelle valg og haft betydning for, hvordan oplysningerne i ordbogen er angivet (fx angives forklaringer der indkredser en dansk oversættelses betydning eller brug i HDO på hollandsk, fordi de skal hjælpe den hollandsksprogede bruger til at vælge den rette oversættelse og/eller til at producere idiomatisk dansk).

Det er vigtigt at lægge mærke til at ovennævnte bifokale dimension ikke realiseres i hver enkelt del, men kun i det samlede sæt af de to ordbøger (DHO og HDO). Oplysningerne om stavemåde, morfologi, syntaks, semantik og pragmatik knytter sig

således til hvert opslagsord eller hvert LU i kildesproget, dvs. at udtaleangivelsen og morfologiske og grammatiske opmarkeringer for danske ord står i DHO og tilsvarende for hollandske ord i HDO, selv om de næppe er rettet mod den danske henholdsvis hollandske bruger. Man skal altså bruge begge ordbøger samtidig for at have rigtig gavn af det bifokale perspektiv med henblik på både at forstå og producere på de to sprog.

Arbejdsgangen

HDO er fremstillet ved hjælp af ordbogsredigeringsprogrammet OMBI. OMBI (et akronym for 'Omkeerbare Bilinguale Lexicale Databanken') er en editor med en vende-funktion, der er udarbejdet specielt til tosprogsordbøger, og som er skæddersyet til RBN, hvilket indebærer at alle data i RBN direkte kan indlæses i programmet, og at alle danske data der laves til enheder i RBN, automatisk falder på plads i den danske relationelle database.

Arbejdsprocessen for udarbejdelsen af HDO var delt op i fem faser:

- *Fase 1:* bearbejdelse af den hollandske database med henblik på udvælgelse af ordbogens makro- og mikrostruktur og import af det selekterede materiale i OMBI. Selv om RBN kan kendetegnes som en database der er skræddersyet til at producere bilinguale ordbøger, måtte redaktørerne mange gange tage stilling til om et givet opslagsord eller EU ('example unit') skulle optages eller ej. På tilsvarende måde har redaktørerne også bestemt at visse opslagsord og visse eksempler skulle tilføjes, således at det oprindelige RBN også blev udvidet.

Udvælgelsen af opslagsord var baseret på tre kriterier: frekvens, kulturel specificitet og kontrastiv relevans. Til grund for udvælgelsen af mikrostrukturens elementer lå følgende inddelingsprincipper: grammatiske kollokationer, leksikalske kollokationer, idiomer, pragmatiske udtryk og derudover i et antal tilfælde også eksempler der illustrerer opslagsordets betydning i en konkret kontekst. Da RBN indeholder et (for) stort antal kollokationer, idiomer og eksempler, skulle der vælges. Først og fremmest indeholder RBN ofte både eksempler i kanonisk form og autentiske korpus eksempler. I redaktionsgruppen valgte vi som oftest den kanoniske form, men i et antal tilfælde foretrak vi et (forkortet) korpuseksempel. Den kanoniske form er som regel at foretrække i en bilingval ordbog ikke bare af pladshensyn, men også af hensyn til de tiltænkte brugere. Den receptive bruger starter normalt fra en bestemt kontekst og skal kunne tolke det ukendte ord eller udtryk i den kontekst det optræder i. Eksemplerne i en bilingval ordbog skal ligeledes hjælpe den produktive bruger til at producere korrekte tekstprøver på målsproget. En bilingval ordbog fungerer derfor bedst, når den gengiver betydningen af et opslagsord i kildesproget ved en oversættelsesækvivalent i målsproget i en ret generisk kontekst. Et yderligere argument mod de lange og komplicerede autentiske eksempler, som man ofte møder i et korpus, er at selve eksemplerne ofte indeholder ord der er ukendte for den receptive bruger, der jo har begrænsede kundskaber i målsproget og altid må karakteriseres som lærer. For at forstå eksemplet bliver brugeren altså nødt til at slå yderligere ukendte ord op, hvilket er uønskeligt og psykologisk. Dette indebærer dog ikke at vi helt kan undvære korpuseksempler. Det gælder især for pragmatiske udtryk og for stilistisk markerede udtryk, hvis brug netop skal illustreres i en konkret kontekst.

- *Fase 2:* oversættelse. Denne fase omfatter selve oversættelsen af de hollandske opslagsord, kollokationer og eksempler til dansk. Hvert EU ('example unit') hænges op under et såkaldt 'target word', hvilket ikke altid er lig med oversættelsesækvivalenten til det hollandske opslagsord (fx under opslagsordet 'beer' (oversat med 'bjørn') står eksempelsætningen 'beren op de weg zien', der oversættes med 'se spøgelse'; i den

danske database hænges dette eksempel selvfølgelig op under opslagsordet 'spøgelse' og ikke under 'bjørn').

Som oftest oversættes et hollandsk opslagsord med én dansk ækvivalent, men det forekommer også at flere oversættelsesækvivalenter angives. For den danske bruger udgør dette ikke noget problem, tværtimod, da en dansk bruger umiddelbart vil forstå de semantiske og kontekstuelle begrænsninger, der knytter sig til enhver oversættelsesækvivalent; den hollandsksprogede bruger derimod kan ikke stille så meget op med en opremsning af forskellige oversættelsesækvivalenter, uden samtidig at få oplyst hvad forskellene er mellem de respektive oversættelsesækvivalenter. Med tanke på den bifokale indretning af ordbogen har vi i mange tilfælde angivet en nærmere betydningsspecifikation for hver enkelt oversættelsesækvivalent. Af disse specifikationer fremgår eventuelle selektionsrestriktioner (fx kun for personer), afvigelser i oversættelsens denotation i forhold til kildesproget (fx kun i bogstavelig betydning) og/eller stilistiske eller pragmatiske konnotationer, der knytter sig til en bestemt oversættelsesækvivalent.

- *Fase 3:* korrekturlæsning af manuskriptet og kontrol af oversættelserne. Denne del af arbejdet blev udført på et Word-print.
- *Fase 4:* I denne fase blev der taget stilling til alle ændringsforslag fra fase 3, som enten blev godkendt og indført i databasen eller endeligt forkastet. Denne del af arbejdet blev udført af hovedredaktøren i samarbejde med en dansksproget redaktør.
- *Fase 5:* endelig korrekturlæsning af manuskriptet.

Da manuskriptet var afleveret til Gyldendal, indtrådte for mig en spændende periode, da jeg blev inddraget i postediting-processen. Jeg arbejdede meget sammen med Morten Winkler, og vi sad mange gange på Gyldendal og diskuterede, hvilke informationstyper der skulle med i den grafiske ordbog og hvilke udvælgelsesprincipper vi bedst kunne håndtere. Der skulle, som også Morten Winkler skriver i sit indlæg, ofte "semantiske øjne" til at se på resultatet af et såkaldt automatisk indgreb.

Brugen af korpora

Redaktørerne havde adgang til alle forhåndenværende hollandske og danske (elektroniske) ordbøger. Derudover havde redaktørerne adgang til Bergenholtz' 4 millioner ords korpus og til Den Danske Ordbogs 6 millioner ords korpus, og lige så snart den forelå, gjorde de også flittigt brug af Korpus 2000. Selvfølgelig havde redaktørerne også permanent adgang til internet, og de brugte tit de forskellige søgemaskinens tekstbaser. Brugen af korpora opfyldte forskellige funktioner:

- at sikre sig en god forståelse af et ords eller udtryks kontekstuelle betydning og registerbegrænsninger. Tværsproglig konsultation af kildesprogs- og målsprogskorpus var ofte nødvendig for at finde den mest adækvate oversættelse og gav desuden ofte anledning til indføjelser af en hyponymbetegnelse.
- at få en indblik i et ords frekvens og at få en oversigt over de mest hyppige grammatiske og leksikalske kollokationer.
- at verificere om en sammensætning eller en afledning, der ikke er optaget i Retskrivningsordbogen, alligevel forekommer på dansk.
- at bestemme over rækkefølgen i tilfælde af at flere oversættelser bliver opført ved et opslagsord.

Et nyttigt instrument?

Dansk og hollandsk hører til de mindre udbredte sprog i Europa. I begge sprogområder er der gode kundskaber i engelsk, og man kunne derfor sætte et spørgsmålstejn ved den konkrete nytte af et ordbogsspar som DHO og HDO. Selvom meget af den transnationale

kommunikation mellem vore lande foregår på engelsk, vil jeg dog understrege vægten af en god direkte forståelse af de mindre sprog. Vi er os mere og mere bevidste om hvad sproglig identitet betyder i et multilingvalt og multiethnisk Europa. Instrumenter der fremmer indbyrdes forståelse af vore nationalsprog, er derfor eftertragtet. Bilingvale ordbøger åbner op for rigdommen i vore respektive sprog og viser samtidigt også dets begrænsninger ved "konfrontationen" med et andet sprog. Ordbøger danner immaterielle broer mellem vore sprogsamfund og bidrager til en bedre selvforståelse og forståelse af den anden. Dette er ikke bare ment som en politisk påstand, men som en konkret henvisning til de mange kontakter der er mellem borgere i vore lande, hvad enten der er tale om migration, handel, turisme, akademiske eller tværkulturelle udvekslinger. For så slet ikke at nævne oversættelse af skønlitteratur. Det er meget vigtigt at (litterære) budskaber, der er affattet på et mindre sprog, kan formidles direkte til et andet mindre sprog, uden at man hver gang behøver at tage omvejen via engelsk. Dette gælder i særdeleshed for sprogparret dansk og hollandsk, der har mange fælles rødder, og som ofte ligger tæt op ad hinanden hvad angår leksikalisering, metaforer og billedsprog.

Der er et tydeligt behov for en dansk-hollandsk/hollandsk-dansk ordbog og dette behov er bare blevet større med tiden på grund af de mange politiske, økonomiske og kulturelle kontakter mellem Danmark og Holland/Belgien både på bilateral basis og inden for den europæiske union. Jeg håber, at HDO og DHO vil opfylde dette behov.

Lidt om produktionen af Dansk-Hollandsk Ordbog og Hollandsk-Dansk Ordbog

Af Morten Winkler

For ca. 9 år siden fik det daværende Munksgaard tilbudt at producere de to manuskripter, hvilket man sagde ja og mange tak til, og dette tilbud overgik til Gyldendal, da Munksgaards ordbøger blev overtaget i 1999. I hele det lange redaktionsforløb havde Munksgaard og Gyldendal faktisk ingen berøring med projektet ud over, at vi så nogle prøver på artikler fra begge ordbøger. Da vi nåede 2002, var projekterne begge to fuldendt fra forfatterens og redaktørens side, og så gik jeg i gang med det tekniske arbejde. Den første iagttagelse og praktiske udfordring var det faktum, at ordbøgerne ikke var koordineret, idet de var redigeret ved to forskellige universiteter i to forskellige lande, og de to projektgrupper arbejdede i forskelligt tempo og med forskellige redigeringsystemer og redaktionsregler. I Amsterdam arbejdede man i rå Word-filer, som havde en rudimentær leksikografisk opmærkning. Lemmalisten, eksempelmaterialet og fodudtryk, ordsprog etc. var udarbejdet i samarbejde med Center for Leksikografi, Handelshøjskolen i Århus. Her har man også hentet råd og vejledning mht. artikelstruktur.

Dansk-Hollandsk Ordbog (DHO) er strukturelt meget enkel, og den leksikografiske opmærkning gjorde det relativt let at konvertere til SGML, så ordbogen kunne importeres i Gyldendals redigeringsystem GestorLex. Imidlertid opstår der jo uvægerligt kodningsfejl, når en ordbog bliver skrevet i et tekstbehandlingsprogram. Der var immervæk godt 45.000 artikler, som fx alle skulle afsluttes med tegnet "@" for at markere slut på artiklen. Hvis dette tegn mangler, kan to artikler smelte sammen. Der var også et stort antal eksempelparenteser, hvor både højre- og venstreparentes skal være der! Det er ikke altid tilfældet, når ordbogen er manuelt indtastet, og det kan også resultere i fejlparsning. Men da først Word-filerne var parset og importeret i GestorLex, var det ligetil at ensrette metasproget, ajourføre danske ord med nyeste retskrivning og

at tjekke bøjningsangivelserne. En anden udfordring frembød tilføjelsen af godt 45.000 udtaleangivelser, som Gyldendal stod for (efter aftale med det hollandske forlag Het Spectrum). Som forlæg brugte vi Peter Molbæk Hansens "Dansk Udtale", hvilket krævede en del efterbearbejdelse. Der skal fx træffes nogle valg der, hvor flere varianter er mulige. Udtale flytter sig jo også, og PMH's bog har snart 15 år på bagen. Nye ord er også kommet til. Et ord som *skideren* var ikke på tale dengang, og her er det ikke uvæsentligt, om trykket ligger på første eller på sidste stavelse!

I Gent derimod var manuskriptet til Hollandsk-Dansk ordbog (HDO) blevet udarbejdet i et databasebaseret redigeringsystem, OMBI, som det Nederlandske Sprognævn havde fået udviklet. Det tog afsæt i en hollandsk kildesprogsbase, Referentiebestand Nederlands, som var forberedt til at indgå i ordbogsmanuskripter til flere sprog. Denne ordbase er opmærket med en meget stor grammatisk detaljeringsgrad, som krævede noget efterbearbejdelse for at komme til at ligne vores sædvanlige artikelstruktur. Fx var der for verbernes vedkommende angivet følgende træk: transitivitet, reflektivitet (obligatorisk eller optionel), fuldverbum med *hebben/zijn* som hjælpeverbum, upersonligt, hjælpeverbum. Dette gav anledning til homografi efter kriterier, vi normalt ikke anvender. Det samme gjaldt for adjektivernes vedkommende, hvor følgende træk var angivet: flektionstype, bøjelig/ubøjelig, komparationsmåde, position: attributiv og/eller prædikativ, mulighed for afledning til adverbium. Fx stod adjektivet *afwezig* oprindeligt i to artikler: en med den konkrete betydning 'fraværende' og en med betydningen 'distræt, åndsfraværende'. Forskellen på de to var, at den konkrete betydning ikke kan bruges adverbialt. Denne detaljerighed har stor værdi for erfarne brugere af kildesproget og kunne fx være fin i en konstruktionsordbog, men den svækker overblikket i en oversættelsesordbog. Vi valgte derfor i samråd med ordbogens hovedredaktør at udelade nogle af disse træk, så visse af homograferne kunne slås sammen. I første omgang skete dette ret mekanisk, men det gik op for os, at der måtte semantiske øjne til at gennemgå de sammentrukne artikler, især med henblik på rækkefølgen af betydningsvarianterne. Fx blev substantivet *advocaat* i sin nye form oversat med 1. æggelikør, 2. advokat!

OMBI er indrettet således, at det kan eksportere i SGML, så råmanuskriptet var relativt let at bearbejde, så det kunne importeres i GestorLex. En af tankerne bag OMBI er, at kildesprogsordbasen, når den er linket til oversættelser på målsproget, kan vendes, så der opstår et råmanuskript til den omvendte ordbog. Dette skete, som det ovenfor fremgår, ikke ved dette ordbogspar. Den såkaldte "hub-and-spoke-model" muliggør også, at man fra dansk via hollandsk i teorien fra OMBI vil kunne lave et råmateriale til fx en dansk-finsk ordbog, hvis der også findes en hollandsk-finsk linkning. I et igangværende hollandsk-finsk projekt er noget af materialet fra hollandsk-dansk genanvendt.

Ud over den strukturelle tilretning af SGML skulle artiklerne efter aftale med det hollandske forlag Het Spectrum forsynes med udtaleangivelser ligesom i DHO. Disse blev stillet til rådighed af Het Spectrum, som i den forbindelse nyskrev et par tusinde udtaler, som ikke fandtes i deres base i forvejen.

Ordbøgerne sigter begge mod at være bidirektionelle, dvs. de skal kunne anvendes af brugere med begge sprog som udgangspunkt. Dette er ikke nogen nem opgave, idet der jo er stor forskel på ordbogens opbygning, alt efter om det er en produktions- eller en receptionsordbog. I DHO er der udtaler og morfologi efter lemmaet, som tilgodeser de hollandske og flamske brugere. Til gengæld er stilmarkeringer og semantiske specifikationer angivet på dansk af hensyn til danske brugere. Grammatiske træk ved ækvivalenten såsom køn og bestemthed er angivet med hollandske forkortelser. I HDO er der hollandsk udtale og morfologi efter lemmaet af hensyn til de danske brugere og semantiske betydningspecifikationer på hollandsk (danskerne forstår jo alligevel

ækvivalenterne). Fagkoder og stilmarkører er på dansk. Disse forskelligheder er udtryk for et forsøg på at tilpasse manuskripterne til et standardiseret udtryk og de kompromiser, der kunne indgås med forfatterne på disse områder.

Ud over datakonvertering gennemgik Gyldendal lemmabestanden i DHO og ækvivalenterne i HDO, foreslog nye lemmaer (og ækvivalenter) og andre lemmaer, som kunne udgå. De sædvanlige: geografiske navne, ministerier, forkortelser og den slags, men også en række nye ord inden for it, bioteknologi, slang etc. Alt i alt har det været en glæde for Gyldendal at kunne udgive to så omfattende og moderne ordbøger imellem to små europæiske sprog. Dette var alene muligt, fordi Ordbogskommissionen for det Nederlandske Sprog, CLVV, skaffede de nødvendige midler (fra den hollandske og belgiske stat, Statens Humanistiske Forskningsråd m.fl.).

NFL-konference i Sønderborg 2005

Den 8. Konference om Leksikografi i Norden finder sted i Sønderborg fra den 24. til den 28. maj 2005. Sidste frist for tilmelding til reduceret pris var 1. marts, men for en beskeden merpris (for medlemmer 1200 DKK i stedet for 1000 DKK) kan man stadig tilmelde sig.

Konferencen byder på et alsidigt fagligt program med knap 40 foredrag, hvis indhold vil kunne ses på NFL's hjemmeside (www.nordisk-sprakrad.no/nfl.htm) sammen med 2. rundskrivelse, hvor alle detaljer om tilmelding, betaling og indkvartering fremgår. Spørgsmål angående konferencen kan sendes til nfl05@dsl.dk.

Vi ses i Sønderborg – mojn!

Med venlig hilsen

Organisationskomitéen
v/Carsten Hansen, Henrik Lorentzen og Lars Trap-Jensen