



LEDA

LEKSIKOGRAFER
I DANMARK

LEDA-Nyt nr. 61 – April 2016

ISSN 1603-7006

Redigeret af Sussi Olsen, Anita Ågerup Jervelund og Merete Korvenius Jørgensen

Meddelelser fra bestyrelsen	s. 2
Indkaldelse til medlemsmøde	s. 3
Indkaldelse til den årlige generalforsamling 2016	s. 4
Når en + en + en bliver flere	s. 5
Græsk leksikografi 2: Nygræsk	s.16
Kend din leksikograf	s.25
Nye ordbøger	s.28
Konferencer	s.30

Medlemsmøde

Når ordbøgerne går på nettet i undervisningen

Malte von Sehested, Gyldendal

Onsdag d. 20. april 2016, kl. 15.00

Københavns Universitet Amager
Njalsgade 136
Lokale 27.0.47

Alle er velkomne
Der serveres kaffe/te i pausen

Meddelelser fra bestyrelsen

Der indkaldes til generalforsamling onsdag den 20. april 2016 kl. 17.00
Københavns Universitet Amager
Njalsgade 136, lokale 27.0.47

Se dagsorden mv. på side 4.

Imellem hvert nyhedsbrev lægges nyheder på vores hjemmeside: <http://leksikografer.dk>.

LEDA-Nyt udgives af Foreningen af Leksikografer i Danmark (LEDA) og udkommer to gange om året. LEDA-Nyt udsendes fast til LEDAs medlemmer, og ældre numre af tidsskriftet kan desuden downloades fra LEDAs hjemmeside (www.leksikografer.dk).

Bidrag til næste nyhedsbrev sendes til formanden inden 1. oktober 2016.

Dette nummer er redigeret af Sussi Olsen, Anita Ågerup Jervelund samt Merete Korvenius Jørgensen.

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:

Sussi Olsen (KU, NFI, Center for Sprogteknologi), formand, saolsen@hum.ku.dk

Merete Korvenius Jørgensen (DSL), næstformand, mkj@dsl.sk

Anita Ågerup Jervelund (Dansk Sprognævn), kasserer, aagerup@dsn.dk

Lena Wienecke Andersen (Epic), lena@wienecke.dk

Jacqueline Levin (Ordbogen A/S), jrl@ordbogen.com

Lotte Follin, lotte.follin@gmail.com

Thomas Troelsgård (DSL), tt@dsl.dk

Kontingent (2016) for medlemskab af Foreningen af Leksikografer i Danmark (LEDA) 125,- DKK (kan indbetales via netbank: reg.nr. 1551, kontonr. 1391968).

Kontingent (2016) for medlemskab af Nordisk Forening for Leksikografi (NFL) 160,- DKK (200,- NOK)

Henvendelse vedrørende indmeldelse og adresseændring sker til foreningens kasserer. Øvrige henvendelser rettes til formanden.

Medlemsmøde

Når ordbøgerne går på nettet i undervisningen

ved

Malte von Sehested, Gyldendal

Med udgangspunkt i skoleordbog.dk fortæller *Malte von Sehested fra Gyldendal* om de overvejelser de har gjort sig om den pædagogiske brug af ordbøger i skolen, når ordbøgerne er digitale.

Hvilke krav stiller det til ordbogsdata, og hvilke udfordringer og muligheder giver det i en undervisningsmæssig sammenhæng?

Onsdag d. 20. april 2016, kl. 15.00

Københavns Universitet Amager
Njalsgade 136
Lokale 27.0.47

Alle er velkomne
Der serveres kaffe/te i pausen

Indkaldelse til generalforsamling 2016 i LEDA

Onsdag den 20. april 2016 kl. 17.00

Københavns Universitet Amager, Njalsgade 136

Lokale 27.0.47

Dagsorden ifølge vedtægterne:

1. Valg af dirigent og referent
2. Formanden aflægger beretning
3. Kassereren fremlægger det reviderede regnskab
4. Fastsættelse af kontingent
Bestyrelsen foreslår at kontingentet fastholdes på 125 kr. årligt.
5. Valg af formand
På valg er Sussi Olsen (indvalgt 2010).
6. Valg af to bestyrelsesmedlemmer og en suppleant
På valg er Merete K. Jørgensen (indvalgt 2012), Lena Wienecke Andersen (indvalgt 2010) samt Lotte Follin (indvalgt 2012).
7. Valg af revisor og revisorsuppleant
På valg er Christian Becker-Christensen og Pernille Neergaard-Dibbern.
8. Forslag
9. Eventuelt

Ad 5. Sussi Olsen modtager genvalg.

Ad 6. Merete K. Jørgensen modtager genvalg. Lena Wienecke Andersen og Lotte Follin modtager ikke genvalg.

Ad 7. Christian Becker-Christensen og Pernille Neergaard-Dibbern modtager genvalg som hhv. revisor og revisorsuppleant.

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

Når en + en + en bliver flere

Om projektet *Middelalderens danske sprog og litteratur formidlet til nutiden*

Af Merete K. Jørgensen, David Svendsen-Tune og Thomas Troelsgård

Projektets baggrund

Projektet *Middelalderens danske sprog og litteratur formidlet til nutiden* er finansieret af Velux-Fonden og Det Danske Sprog- og Litteraturselskab (DSL). Det blev afviklet i årene 2013 til 2015 og er en videreførelse af to tidligere projekter på DSL, nemlig 1) digitaliseringen og udlæggelsen af Gammeldansk Ordbogs (GO) seddelsamling og 2) *Studer Middelalder på Nettet* (SMN).

Formålet med projektet er at formidle den danske middelalders sprog og litteratur – og dermed også den middelalderlige tankeverden – for et bredere publikum, men uden at give køb på det videnskabelige niveau. For at opnå dette mål skulle:

- GO's artikler publiceres i elektronisk form,
- GO's artikler linkes til de moderne opslagsord i Den Danske Ordbog (DDO) eller Ordbog over det danske Sprog (ODS) hvor det var muligt, og
- teksterne i det nye SMN, nu kaldet 'middelaldertekster.dk', kobles til GO. Linkningen mellem GO og de moderne ordbøger kunne så fungere som en form for løbende forklaringer.

Ved projektets begyndelse havde vi indset, at ressourcerne ikke rakte til en egentlig færdigredigering af GO. Derfor var det nødvendigt at nytænke projektet og lægge en mere realiserbar plan for færdiggørelsen, eller i hvert fald for en foreløbig færdiggørelse af ordbogen. Ordbogens status i 2013 var, at en delmængde af artiklerne var redigerede efter traditionelle leksikografiske principper, dvs. med anførelse af fx lydlige og bøjningsmæssige varianter og med en meget fintmærkende semantisk beskrivelse, der blev underbygget med mange indskrevne citater. En mindre del var redigeret efter en beskåret model: ingen varianter og uden indskrevne citater, men til gengæld med links til de elektroniske sedler i GO's seddelsamling. Den sidste del af artiklerne – og desværre også den største – var ikke redigerede.

Studer Middelalder på Nettet var et af DSL's allerførste netprojekter. Det blev udarbejdet i perioden 2002-2006, og var ved offentliggørelsen fuldt på højde med tidens teknologiske muligheder. Men den teknologiske udvikling krævede efterhånden, at der blev udarbejdet en

ny, opdateret version, som selvfølgelig skulle udnytte det foreliggende materiale, men også berige og udvide det væsentligt. Projektet *Middelalderens danske sprog og litteratur formidlet til nutiden* blev derfor opdelt i to delprojekter: 1) digital publicering af GO's artikler, og 2) *Studér Middelalder på Nettet* i ny version.

Den digitale publicering af GO's artikler

GO var oprindelig planlagt som en trykt ordbog. Derfor var redigeringsystemet, iLEX, også sat op til at imødekomme en trykt bogs krav. Efter revideringen af redigeringsprincipperne, og med tanke på en elektronisk publicering, måtte såvel artikelstruktur som design lægges afgørende om. Den mest afgørende forskel på artikler før og nu ligger i, at tidligere tiders indskrevne citater nu skal erstattes med links til de digitale sedler. Vi ønskede dog ikke at miste det arbejde, der allerede var lagt i de ”fuldredigerede” artikler, så designet skulle både tage hensyn til dem og til de nye, linksredigerede artikler.

The screenshot shows the GO website interface for the word 'piperkværn'. At the top, the word is written in red, followed by 'sb.' and a box indicating 'KENDT SIDEN 1500-25'. Below this is the section 'Betydning' with a list of entries. The first entry is 'peberkværn' with a quote: 'tagh .. alwn ligæ veckt meth muskathet oc mal thel sammen i en peber qwern Lægeb. (NkS314b). D3r'. The second entry is 'sidhen skal thu haffwe eth lood canipher [se kamfer] oc mole then i en reen piper qwern Lægeråd. 45r,b'. Below the entries is the 'ETYMOLOGI' section, stating 'sammensat af: [piper](#) og [kværn](#)'. At the bottom, there are two boxes: 'Andre ressourcer for gammeldansk' with links to 'Slå piperkværn op i Gammeldansk Seddelsamling' and 'Gå til Gammeldansk Kildefortegnelse', and '“piperkværn” i andre ordbøger' with a link to 'Slå peberkværn op i Den Danske Ordbog'.

piperkværn *sb.*
KENDT SIDEN 1500-25

Betydning

- *peberkværn*
tagh .. alwn ligæ veckt meth muskathet oc mal thel sammen i en peber qwern Lægeb. (NkS314b). D3r
- sidhen skal thu haffwe eth lood canipher [se [kamfer](#)] oc mole then i en reen piper qwern Lægeråd. 45r,b

ETYMOLOGI sammensat af: [piper](#) og [kværn](#)

Andre ressourcer for gammeldansk

Slå **piperkværn** op i Gammeldansk Seddelsamling

Gå til Gammeldansk Kildefortegnelse

“piperkværn” i andre ordbøger

Slå **peberkværn** op i Den Danske Ordbog

Fig. 1: Artiklen *piperkværn* i GO.

Fig. 1 viser artiklen *piperkværn*, sb. i det nye design for ordbogen på nettet. Lige under opslagsordet er der en oplysning om ordets alder i det danske sprog. Næste afsnit er betydningsafsnittet, hvor betydningen og citaterne underbygger hinanden. Et kort etymologisk

afsnit følger, og derefter tre link. Ét til GO's seddelsamling, og ét til artiklen *peberkværn* i DDO. Desuden er der et link til ordbogens kildefortegnelse, hvor hele det gammeldanske kildekorporus nøje er beskrevet.

En artikel kommer til at se ganske anderledes ud nu, hvor vi ikke længere udvælger, afgrænser og indskrifter de enkelte citater. Fig. 2 viser artiklen *brun*, sb. - 'brønd'. I afsnittet "Betydning" ses en række små seddel-ikoner. Det er links til de elektroniske sedler i seddelsamlingen. I stedet for at læse et indskrevet citat, skal man her klikke på ikonet og læse selve sedlen. I artiklen er ikon-billederne delt op på ældste og yngste belæg, betydning og parallel- og forlægsbelæg. Også her er der links til hhv. GO's seddelsamling og til artiklen *brønd* i DDO. Afsnittet indeholder også oplysninger om ordets etymologi og hvilke orddannelser det indgår i.

I. brun sb.

KENDT SIDEN ca. 1250

BØJNINGSKLASSE m. a-st.

BØJNING pl. brunne

Betydning

- brønd; kilde

DATERING Ældste belæg:  • Yngste belæg: 

BELÆG  •  •  •  •  •  •  •  •  •  (afskr. ca. 1500?) •  •  •  • 

PARALLEL graffwe GldBib. 1Mos. 26.18  • hwlæ PLaale (DgO.nr. 652) 

• kieldæ ABrade (Bønneb.2: 126,2)  • watn PLaale (DgO.nr. 401) 

FORLÆG brwn fsu.  • borne mnt.  • sode mnt  • burne mhty. 

• prunn mhty.  • cisterna lat.  • puteus lat. 

luk

[SkL.DgL.kap.99] vm mæn høggiæ træ sammæn.

Høggiæ mæn træ sammæn oc faldæ ofnæ ænnæ theræ. sva at han dør af, tha bôte the mæth hanum høggo frændom hans the marc. [SkL.DgL.kap.100] Dør man j brunni annærs. manz then ær han á ænsammæn. Tha bôte then ær brunni atte oc grauæ lot frændom hins døthæ the marc. Drucnæ man j almænnings brunni then alfi grannæ ægho. [h ov.l.] tho at en groue. tha bôtæs æy hin døthe atær. [SkL.DgL.kap.101] Dør man j lergraum [rett. til lergrafum med sen.hd., ca. 1430] ællær ræfgrauum. // [57] bôte oc hin the marc grauæ lot. Aen tho sighir kunung at fore theskyns mal. ællær grauæ scal man ækki bôte.

FD

SkL. HolmB74.5^f.14.

ETYMOLOGI [Brønd](#)

ORDDANNELSE efterled: [almænningsbrun](#), [bybrun](#), [dravbrun](#)

Fig. 2: Artiklen *brun*, sb., med seddel-ikoner (heraf en aktiveret til visning).

Som det blev nævnt tidligere, så er der stadig mange artikler i GO, som *ikke* er redigerede. Men alle de godt 29.000 ord, som findes i de gammeldanske kilder, har en artikel, der indeholder en række faste informationer om glosens identitet. Som et eksempel på en artikel, der ikke er redigeret ses i fig. 3 artiklen *brotherdotter*, sb.: Der er et opslagsord, (og evt. et homografnr.),

en ordklasseangivelse og en angivelse af ordets ældste dokumenterede forekomst i det danske sprog. Desuden findes tre muligheder: Et link til de elektroniske sedler, så interesserede kan se det samlede materiale på ordet. Et link til GO's kildefortegnelse, og et link til ordet i moderne ordbøger, her altså ordet i den moderne form *brordatter*, som findes i DDO. Der er naturligvis ord på gammeldansk, som ikke længere er levende. I disse tilfælde oplyses det, at ordet er uddødt. Oplysningen om ældste forekomst og oplysningerne om ordet i moderne ordbøger giver en historisk dybdedimension, som det ikke tidligere har været muligt at udtrykke i en ordbog.

The screenshot shows the entry for 'brotherdotter' in the Gammeldansk Ordbog (GO). At the top, the word 'brotherdotter' is written in red, followed by 'sb.' in black. Below this, a grey box contains the text 'KENDT SIDEN ca. 1250'. There are two main sections, each with a blue header and a light blue background. The first section is titled 'Andre ressourcer for gammeldansk' and contains two links: 'Slå **brotherdotter** op i Gammeldansk Seddelsamling' and 'Gå til Gammeldansk Kildefortegnelse'. The second section is titled '“brotherdotter” i andre ordbøger' and contains one link: 'Slå **brordatter** op i Den Danske Ordbog'. At the bottom left, there is a link 'Rapportér et problem' and at the bottom right, the text 'fra Gammeldansk Ordbog'.

Fig. 3: Artiklen *brotherdotter* (ikke redigeret) i GO.

Linkningen til de moderne former af ordet betyder også, at det er blevet muligt at søge efter gammeldanske ord med ordets moderne form. Hvis man er interesseret i sundhedsforhold i middelalderen, kan man fx søge på *helbred*, og så kommer man til den gammeldanske form *helbræghthe*. Det er altså ikke længere nødvendigt at kende ordbogens lemmatiseringsprincipper for at få udbytte af den. GO's opslagsformer afspejler en sprogform, der kan dateres til ca. 1350, og for ikke-sprogfolk har det derfor været vanskeligt at anvende ordbogen og dens seddelsamling.

Koblingen til de moderne ordbøger formidler dog ikke blot søgemuligheden med moderne former. Den muliggør også, at opslag i de moderne ordbøger virker som en løbende kommentar eller forklaring til de gammeldanske tekster.

Studér Middelalder på Nettet i ny version

Det gamle projekt *Studér Middelalder på Nettet* måtte nødvendigvis opdateres til en ny og mere tidssvarende version. De tekster, der allerede var udlagt, var repræsentative for middelalderens dansksprogede litteratur med hensyn til genre, periode og sprogform, og emnemæssigt omfattede de jura, lægekunst, pædagogik, historieskrivning, geografiske forestillinger, digtning og religiøs litteratur. Teksterne blev fremlagt i en nøjagtig afskrift af deres originale form, og som i andre videnskabelige udgaver ledsages teksterne af tekstkritiske oplysninger om deres tilblivelse og den senere brug af dem. Også deres litteratur- og kulturhistoriske baggrund opridses, og desuden offentliggøres digitale fotografier af en del af originalerne. Den nye version, *middelaldertekster.dk*, udvider tekstudvalget kraftigt og – vigtigere – tilføjer en længere række avancerede søgemuligheder. Det vil blive muligt at finde enkelte gloser i teksterne, hvad enten man går ud fra deres moderne eller deres middelalderlige form. Man kan definere søgninger med flere ord, med ordklasser og med kombinationer af disse parametre.

Den nye version bevarer det videnskabelige niveau og gør samtidig teksterne lettere tilgængelige ved at sammenkoble dem med GO. Der linkes fra hvert enkelt ord i teksterne til de relevante ordbogsartikler, og takket være linkningen fra GO til de moderne danske ordbøger vil denne kobling som tidligere nævnt reelt fungere som en løbende forklaring af teksterne.

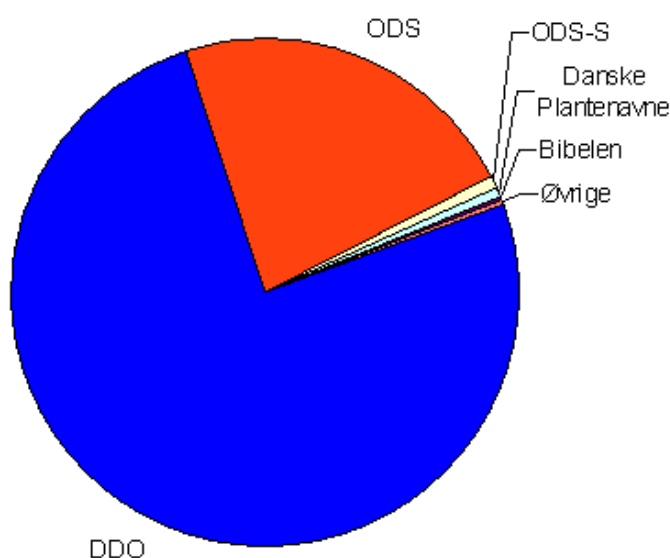


Fig 4: Fordelingen af ressourcer der er linket til fra GO.

Koblingen mellem GO og moderne ordbøger

Koblingen til de moderne ordbøger – eller rettere til ordbøger, der beskriver sprogtrin yngre end gammeldansk – udføres i GO's database. Der linkes til DDO, hvis dette er muligt, ellers til ODS. Herudover er der blandt DSL's ordbogsressourcer linket til Meyers Fremmedordbog og Moths Ordbog. Desuden er der i enkelte tilfælde linket til Retskrivningsordbogen, Bibelen og J. Langes *Ordbog over Danske Plantenavne*. Fordelingen af brugte målordbøger fremgår af fig. 4.

Ved koblingen ses først på match af hele opslagsordet, både for simpleksord og for sammensætninger og afledninger. Således kan man fx koble:

brun (sb.) → *brønd* (DDO)

piperkværn → *peberkværn* (DDO)

hjalperinne → *hjelperinde* (ODS)

For en del sammensætninger og afledninger findes opslagsordet i sin helhed ikke i de nyere ordbøger, mens de enkelte elementer gør; i disse tilfælde oprettes tilsvarende to eller flere koblinger, fx:

kastaneabark → *kastanje* (DDO) og → *bark* (DDO)

med en tilknyttet oplysning ved den enkelte kobling om, at den kun vedrører en del af ordet. Eksempler på mere komplekse tilfælde er fx:

allehælghenmissedagh → *allehelgen* (DDO) og → *messe* (DDO) og → *dag* (DDO)

hvor tre mål-lemmaer er i spil, eller:

botgørende → *bod* (DDO) og → *gøre* (DDO)

hvor sidsteleddet er en bøjet form af mål-lemmaet.

Som nævnt ovenfor findes der GO-ord, der må anses for uddøde (i denne sammenhæng forstået som fraværende i de ressourcer, der kobles til), fx *garnalie* ('finkornet metalgranulat'), der så i stedet for en kobling forsynes med et "uddødt"-mærke. For sammensætninger og afledninger hvoraf et element er uddødt, mens andre kan kobles til nyere former, anføres koblinger for hvert element, fx:

dryllegæst → *drylle* "uddødt" og → *gæst* (DDO)

idet *drylle* ikke findes i de nyere ordbøger. Fig. 5 viser fordelingen mellem forskellige typer af koblinger.

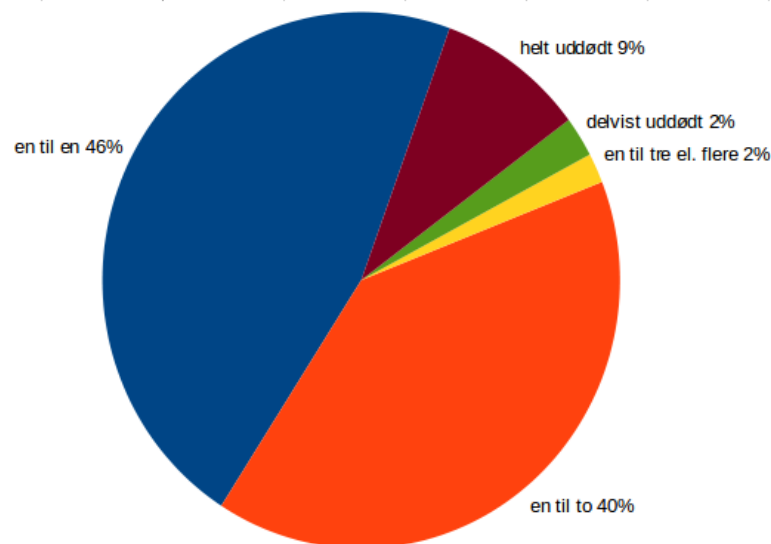


Fig. 5: Fordelingen af typer af koblinger mellem GO og andre ordbøger m.v.

Selve koblingen foretages i ordbogsredigeringsprogrammet iLEX, hvor man, efter at have indtastet den formodede moderne form af opslagsordet i det element, der rummer koblingen, automatisk kan slå formen op i en særlig hjælpebase, der samler opslagsordene i de større DSL-ordbøger; hvis det ønskede mål-lemma er på resultatlisten, vælges det, og oplysninger om kilde og id-nummer indhentes i GO-basen. Som eksempel på arbejdsgangen kan vi tage ordet *ædelik*; redaktøren indtaster den formodede moderne form, *ædelig*, og præsenteres så for en resultatliste som vist på fig. 6. Der er ingen DDO-forekomster, så andenprioriteten, ODS, må vælges. Her er der tre homografer at vælge imellem, og da der er tale om et adjektiv med betydningen 'ædel', vælges – efter kontrol af ODS-artiklerne – homograf nummer 3, idet nummer 2 dækker en afledning af verbet *æde*, altså betydningen 'spiselig'.

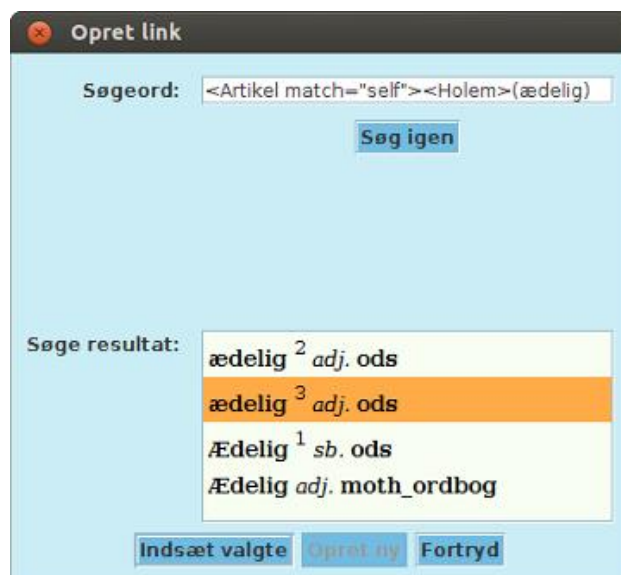


Fig. 6: Resultatliste for en søgning på ædelig i DSL's "samleordbog".

Lemmatisering af de gammeldanske tekster

En central del af projektet er lemmatiseringen af teksterne på middelaldertekster.dk – dvs. etableringen af en entydig forbindelse mellem hver forekommende ordform i teksterne og det relevante lemma i GO. Koblingen mellem tekst og ordbog er forudsætningen for at berige teksterne på middelaldertekster.dk med forklaringer på ordniveau, men hvor man for moderne tekster ofte vil overlade denne opgave til en automatisk lemmatiser, ville det for gammeldansk ikke give et tilfredsstillende resultat; teksterne overholder ingen normer for retstavning, og formvariationen er enorm, idet stavningen varierer afhængigt af tekstens periode, genre, oprindelse, forfatter, afskriver og mange andre faktorer. Hertil kommer, at mange af de bøjningsformer, der er forsvundet i moderne dansk, optræder i periodens tekster. Selv ved applicering af blot de mest basale regler for bøjning og stavevariation vil en lang række former derfor ikke kunne identificeres entydigt.

For at opnå en brugbar lemmatisering kræves således en manuel indsats, nemlig identificering af korrekt lemma for hver ordform. Til brug for projektet er der derfor udviklet et værktøj, der for hvert løbende ord genererer en liste af mulige lemmakandidater, hvorfra en redaktør så vælger den korrekte løsning, hvorefter lemma-id'et tilføjes tekstens XML-data. I tilfælde hvor det korrekte lemma ikke optræder blandt kandidaterne, kan det indtastes manuelt.

Genereringen af lemmaforslag er baseret på en række transformationsregler, der afspejler de mest udbredte stavevariationer i perioden. Fig. 7 viser tre af de mere end 200 regler der p.t. anvendes i systemet.

Variantform	Normaliseret form
-dh-	-th-
-rræ	-rre
[vokal]k[vokal]	[vokal]kk[vokal]
...	...

Fig. 7: Eksempler på lemmatiserens transformationsregler.

Transformationsreglerne anvendes rekursivt på hver ordform, dvs. efter følgende princip:

Input $\leadsto$ output1 output1 $\leadsto$ output2 output2 $\leadsto$ output3 osv.

Denne fremgangsmåde overgenererer naturligvis en stor mængde former, men ved match mod GO's normaliserede lemmaformer er dette nødvendigt, da de korrekte lemmaer ellers ofte ville havne ”under radaren”, og altså slet ikke komme med blandt lemmaforslagene. Fig. 8 viser, hvordan formen *fadher* ved applicering af transformationsreglerne giver seks

forslag, hvoraf kun et, *father*, matcher et GO-lemma, og således bliver foreslået som kandidat i lemmatiseringsværktøjet.

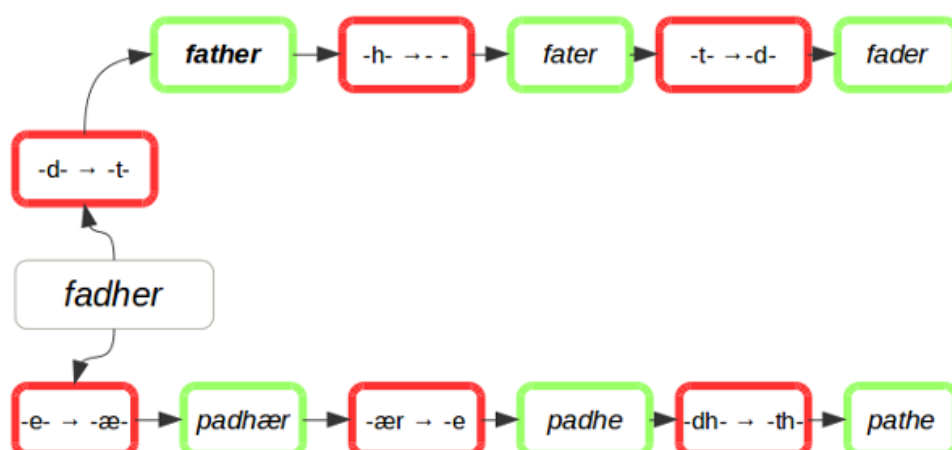


Fig. 8: Eksempler på automatisk genererede variantforslag til formen *fadher*.

Lemmakandidaterne præsenteres for redaktøren i en prioriteret rækkefølge, der bl.a. fastlægges ud fra følgende forhold:

1. Set tidligere i teksten
Lemmaer, der før er valgt for lignende former i den aktuelle tekst foretrækkes.
2. Set i en anden tekst
Lemmaer, der før er valgt for lignende former i de allerede lemmatiserede tekster foretrækkes. Denne type forslag anføres desuden efter faldende frekvens.
3. Dannet vha. transformationsregler
Lemmaer, der matcher formen direkte, foretrækkes i forhold til match med en genereret form.

Systemet er ”intelligent”, forstået på den måde at listerne over registrerede matches (svarende til punkt 1 og 2 ovenfor) opdateres løbende, og således løbende forbedrer prioriteringens nøjagtighed.

Udover at understøtte selve bindingen af ordform til lemma rummer værktøjet en række faciliteter til brug i særlige tilfælde, herunder:

- Mulighed for at knytte to løbende ord til ét lemma (hvor de to ord læses som en særskrevet sammensætning – og matcher ét GO-lemma).
- Mulighed for at knytte ét løbende ord til to lemmaer (hvor en sammenskrivet form ikke er GO-lemma, mens dens to led er).
- Mulighed for at foretage manuel lemmatilknytning (hvor det korrekte lemma ikke er blandt de genererede forslag).

- Mulighed for manuelt at vælge, hvilke af de allerede lemmatiserede tekster der skal bruges ved fastlæggelse af lemmaforslagenes rækkefølge i forslagslisten. Ved tilvalg kun af tekster af samme type eller fra samme periode som den aktuelle tekst kan andelen af korrekte førstevalg i mange tilfælde forøges væsentligt.

Og så til lidt statistik: Fig. 9 viser dels succesraten for lemmatiseringsværktøjets fund af korrekt lemma (sort søjle), dels delmængden deraf hvor det korrekte lemma desuden har første prioritet på kandidatlisten (rød søjle). Fig. 9 viser data for de otte først lemmatiserede tekster, i den rækkefølge de er kørt igennem systemet. Ser man på de sorte søjler, der altså viser om systemets lemmaforslag rummer det korrekte lemma, ses et spring fra de første tekster, der ligger på lidt over 60 %, til de næste fire, hvor lemmaet er fundet i mere end 90 % af tilfældene. Teksterne er fra samme kilde, så springet skyldes primært at systemet efter de to første tekster har ”lært” næsten hele denne kildes ordforråd (eller snarere ”formforråd”). For de to sidste tekster falder tallet lidt, hvilket må skyldes, at der her er tale om andre kilder; selvom systemets basisordforråd her er udvidet betragteligt, skal der altså i et vist omfang begyndes forfra, hvilket naturligvis vil medføre en lavere succesrate.

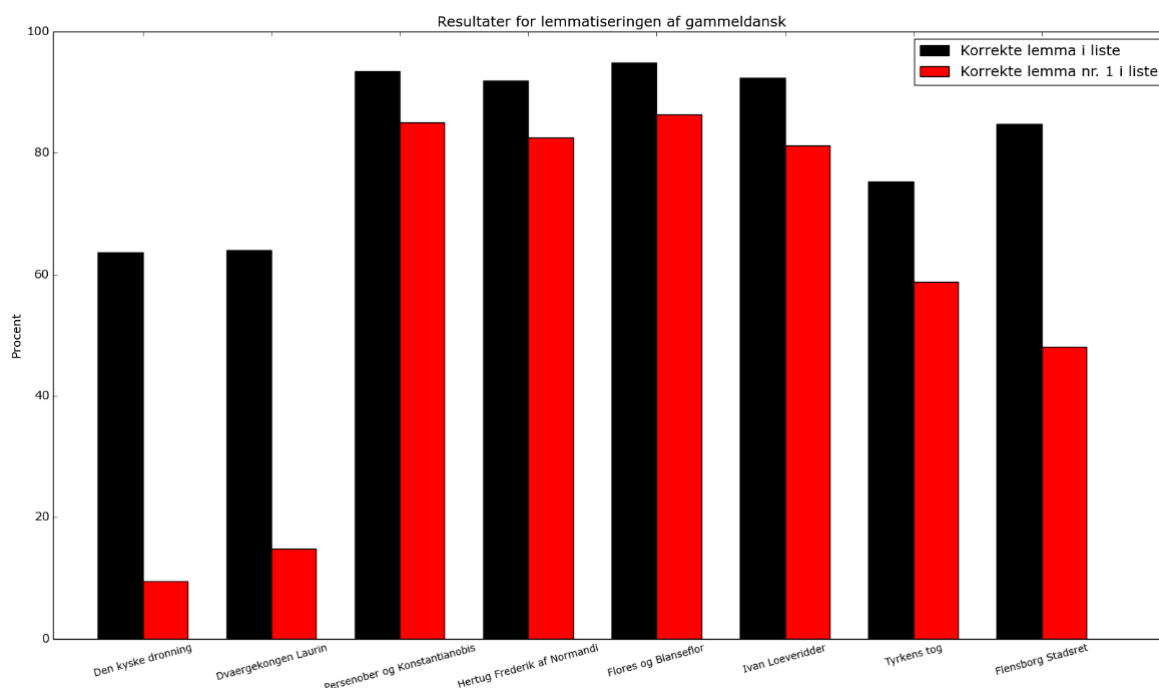


Fig. 9: Succesraten for systemets fund af dels korrekt lemma (sort søjle), dels korrekt lemma som førstevalg (rød søjle).

De røde søjler i fig. 9 viser, hvor ofte systemets prioritering har ramt plet, så det korrekte lemma optræder som nummer et på forslagslisten. Også her scorer de to første tekster markant ringere end de øvrige, men der ses dog en stigning fra første til anden tekst, som en følge af systemets dynamiske opprioritering af tidligere matches. For de næste fire tekster ligger tallet for korrekt førstevalg stabilt på lidt over 80 %, mens det for de to sidste (der som nævnt repræsenterer andre teksttyper) falder lidt, men dog ligger væsentlig højere (50-60 %) end for de to første tekster (10-15 %).

I de otte lemmatiserede tekster er 97.002 ordforekomster blevet koblet til et lemma i GO. Formerne er fordelt på 2.791 af ordbogens godt 29.000 lemmaer, hvilket altså vil sige at kun ca. 9,6 % af GO's lemmaer forekommer i disse tekster. De 10 hyppigste lemmaer er (i moderne form): *den*, *han*, *og*, *jeg*, *være*, *i* (pron.), *da*, *hun*, *så* (konj.) og *til*. Forholdet mellem lemmaerne og deres frekvens viser en typisk ”long tail”-fordeling, hvor ganske få lemmaer forekommer hyppigt, mens rigtig mange forekommer sjældent. Således optræder 36 % af lemmaerne kun én gang, og ca. 13 % kun to gange.

Fremtidsplaner

Denne artikel kunne godt have haft mange flere ”+ en'ere” i titlen, for der er jo i princippet ikke noget i vejen for at inddrage flere elektroniske ressourcer i linkningen. Indtil videre er der linket til DDO, ODS, nogle få links til Moths ordbog og Meyers fremmedordbog, men vi kunne også have tilføjet Holberg-Ordbog – og til næste år vil det også blive muligt at linke til Otto Kalkars *Ordbog til det ældre danske Sprog*. Og middelaldertekster.dk kan jo udvides med endnu flere tekster. Og fra teksterne vil det blive muligt at samle formkataloger af forskellig art. Og det hele skal jo nødvendigvis ikke kun appliceres på middelaldertekster.

Inden alle æggene ligger på jorden, skal det tilføjes, at der skal tilvejebringes flere ressourcer, end vi råder over for nuværende, så det sidste er foreløbig kun fremtidsdrømme, men drømme, som vi har medtaget i vores planer for fremtiden.

Medvirkende ved projektet:

Thomas Hansen, it-koordinator
Merete K. Jørgensen, leksikograf
Simon Skovgaard Boeck, editionsfilolog
Jonas Nielsen, datalog, webudvikler
Marita Akhøj Nielsen, projektleder, filolog
David Svendsen-Tune, datalingvist
Nicolai Hartvig Sørensen, datalingvist
Thomas Troelsgård, leksikograf, datalingvist

Merete K. Jørgensen
Seniorredaktør
Det Danske Sprog- og
Litteraturselskab
mkj@dsl.dk

Thomas Troelsgård
Seniorredaktør
Det Danske Sprog- og
Litteraturselskab
tt@dsl.dk

Græsk leksikografi 2: Nygræsk

Af Rolf Hesse

I forrige nummer af LEDA-Nyt (nr. 60) fortalte jeg om oldgræsk leksikografi, dvs. ensprogede græske ordbøger og encyklopædier ("leksika") fra omkring 300 f.Kr. og op gennem den byzantinske periode indtil ca. 1400 e.Kr. Det er receptionsordbøger, der hjælper til at forstå de klassiske tekster, med undtagelse af den såkaldte atticismebevægelse i slutningen af oldtiden, som med sine ordlister over "uacceptable" ord i samtiden ønskede at korrigere den aktuelle sprogbrug og genetablere de klassiske normer. Nu er vi så nået til *nygræsk* (engelsk: Modern Greek, tysk: Neugriechisch, græsk: Νέα Ελληνικά (néa elliniká)).

Ved *oldgræsk* forstår jeg her de former for græsk som er overleveret i antikke tekster og som vedblev – både i byzantinsk tid og senere – i stort omfang at tjene som forbillede for skriftsproget, lidt ligesom latin i Vesteuropa, mens *nygræsk* er den sprogform der anvendes i dag og hvis grundstrukturer allerede ses i tekster fra omkring 1000 e.Kr.¹ Den nygræske litteratur starter med en digtcyklus om helten Digenis Akritas, søn af en arabisk emir og en kristen prinsesse, og derefter kan sproget følges op til nutiden.

Græske ordbøger 1500-1830

Omkring år 1500 er bogtrykkerkunsten etableret i Vesteuropa, men ikke i det osmanniske rige, så de første græske bøger blev trykt i Vesten med Venedig som vigtigste sted. Nu begyndte der også at trykkes græske ordbøger, vist altid flersprogede, især med latin som andet sprog. De lærde i Vesteuropa skulle nu lære græsk, dvs. det sprog de klassiske græske tekster og – ikke mindst efter Reformationen – Det Ny Testamente var skrevet på. Græsk opfattedes stadig i det store og hele som ét sprog, selv om afstanden mellem daglig tale og de klassiske normer efterhånden blev tydeligere og der opstod behov for ordbøger der specielt registrerede samtidens græsk. Den tidligste ordbog der rummer nygræske ord hedder *Μέγα και πάνν ωφέλιμον Λεξικόν* (Mégħa kai páni ofélimon leksikón²) "Stor og særdeles nyttig ordbog" af Guarino Favorino, Rom 1523.

¹ De vigtigste grammatiske ændringer fra oldgræsk til nygræsk er bortfald af infinitiv og dativ. Infinitiv er stort set erstattet af finite konstruktioner, mens dativens funktioner er overtaget af genitiv og præpositionsforbindelser. Desuden har fonetikken udviklet sig; antallet af vokaler er reduceret til fem (især ved den såkaldte itacisme), og betoningen er ændret fra musikalsk til trykstærk/trykssvag. Den ændrede fonetik skaber store problemer ved transskriptionen af græske ord, når det drejer sig om den mellemliggende periode: I århundrederne op til Renæssancen udtaler man oldgræsk som samtidsgræsk, dvs. stort set som i dag. Men så konstaterede Erasmus fra Rotterdam o.a. at især vokaltegnene havde helt andre værdier i oldtiden, f.eks. var H (eta) et [e:] og ikke et [i], OI et [oi] og ikke et [i], B repræsenterede et [b] og ikke et [v], og det såkaldte "spiritus asper"-tegn foran begyndelseskarakter angav et [h], som siden er forstummet. Således kan 'Ελληνικόν ('græsk') enten transskriberes *Hellēnikón* ("erasmeisk" udtale) eller *Ellinikón* (nygræsk).

² Dansk transskription med forenklet udtaleangivelse: dh repræsenterer bogstavet δ [δ], gh γ [γ],

Et tidligt eksempel på en egentlig nygræsk ordbog er Ioannes Meursius' *Glossarium graecobarbarum*, Leiden 1610 (fig. 1³), ordbog over "barbarisk græsk" med forklaringer på latin; den indeholder ord der ikke findes i oldgræsk, som f.eks. adjektivet *ἄσπρος* (*áspros*), det nygræske ord for 'hvid', der på oldgræsk hedder *λευκός* (*leukós*, nygræsk: *lefkós*), et ord, der stadig bruges i dag parallelt med *áspros*.

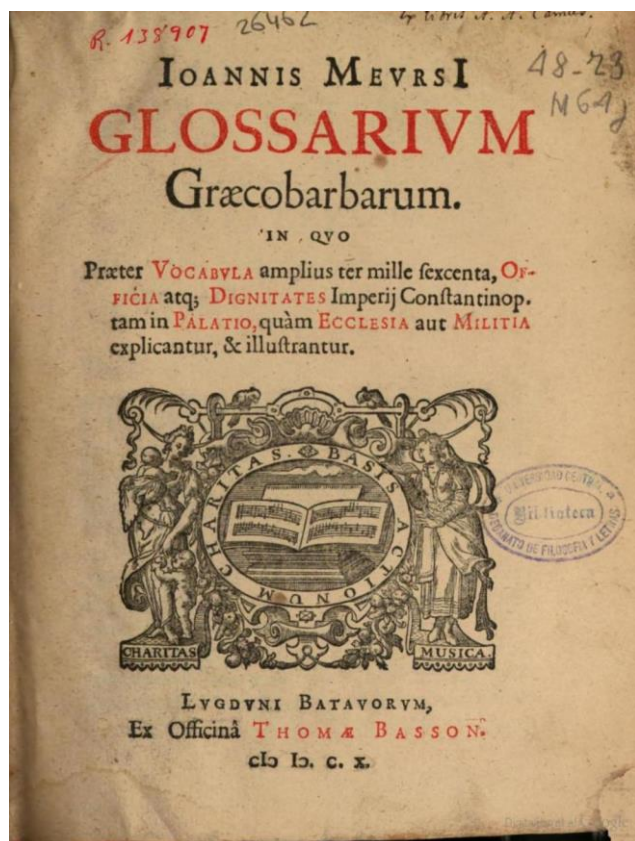


Fig. 1. Ioannes Meursius' *Glossarium graecobarbarum*, Leiden 1610.

I 1635 udkommer i Paris den katolske teolog Simon Portius' *Λεξικόν Λατινικόν, Ρωμαϊκόν και Ελληνικόν* ("Leksikón latinikón romaikón kai hellenikón") – *Dictionarium Latinum, Graeco-barbarum et Litterale* (fig. 2). Bogen består af to trespaltede dele: en ordbog på 512 sider med rækkefølgen latin – "barbargræsk" – "skriftgræsk", og en "lille-ordbog"⁴ på 247 sider "barbargræsk" – "skriftgræsk" – latin (se fig. 3).

Titlen kræver en nærmere forklaring: Adjektivet *romaikós* 'romersk' henviser oprindeligt til det østromerske rige, det vi kalder det byzantinske rige, der var en fortsættelse af Romerriget. *Romaikós* refererer til samtidens sprog/mennesker osv.⁵ i modsætning til *hellenikón/ellinikón*, der betød 'oldgræsk' 'klassisk græsk' (i tidlig byzantinsk tid også 'førkristen', 'hedensk'). *Romaikós* i den græske titel svarer til

th θ [p]; o og e er korte åbne vokaler. Betonet vokal får accent: Elládha, ou repræsenterer ou [u].

³ De afbildede ordbøger er gengivet i formindsket størrelse.

⁴ Λεξικόπουλο (leksikópoulo), ordret: "ordbogsunge".

⁵ Substantivet Ρωμιός (romiós) betyder tilsvarende 'græker', 'den jævne græker', og bruges stadig i denne betydning,

Graeco-barbarum i den latinske, mens ellinikón svarer til (Graeco-)Litterale, 'skriftsprog', 'formelt/lærd sprog'. Undertiden er det "barbargræske" og det "skriftgræske" ord dog identiske. Det er i øvrigt overraskende at se hvor mange af ordene i Simon Portius' ordbog (eller ordliste) der stadig er i brug i nutidig græsk, både af de "barbariske" og af de "lærde" ord (se forklaring til fig. 3).

De latinske betegnelser graecus og Graecia brugtes og bruges kun meget sjældent på græsk. Endnu da Grækenland skulle optages i EF, løb der en langvarig diskussion om de officielle betegnelser for land, sprog og tilhørsforhold. Resultatet varierer fra medlemsstat til medlemsstat, således hedder det på dansk Den *Hellenske* Republik, men den *græske* regering, mens sproget forkortes som **EL** (= elliniká).

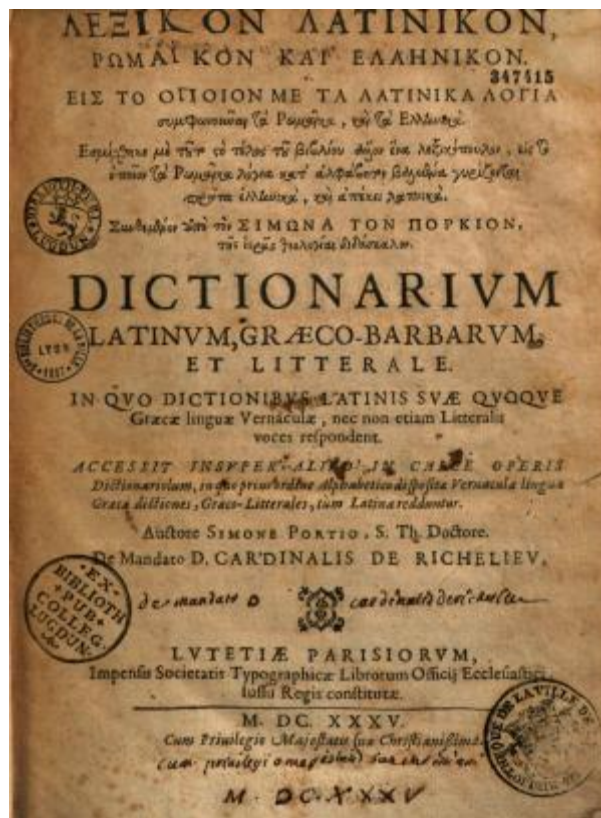


Fig. 2. Simon Portius' Dictionarium Latinum, Graeco-barbarum et Litterale, Paris 1635.

Fig. 3. (næste side) Simon Portius' Dictionarium, 1. del s. 136. *Dominus* (herre) er oversat til dagligsprog ("barbargræsk") som αφέντης, αφθέντης (aféntis, afthéntis, jf. autentisk, et opr. oldgræsk ord), mens det på "Graeco-Litterale" oversættes som κύριος, δεσπότης, άναξ (kírios, dhespótis, áanax) – kyrios betyder 'herre' despotis bruges bl.a. om højere gejstlige; áanax betyder 'konge', 'fyrste'. Domus (hus) hedder nøjagtig som i dag σπίτι (spíti, af latin hospitium), men også alle fire skriftsprogsord οίκία, οίκος, δῶμα δόμος (oikía, oíkos, dhóma, dhómos – oi udtales [i]) bruges den dag i dag. οίκία: jf. økonomi og økologi.

136 *Latine.**Græco-Barb.**Græco-Litt.*

Documentum.

μαθήμα, διδασκαλία, δοκιμα-
σία, παράδειγμα.

δῖσμα, διδάσμα, ἐπίδειξις, πῖρα.

Dochimus.

ἔνα ποδάρι με πέντε συλλαβῶν,
μία βραχὴ δύο, μακρᾶς, μία
βραχὴ, ἢ μία μακρῇ.

δὲχος.

Doctor.

διδάσκαλος.

καθηγητής, διδάσκαλος, καθηγημάτων

Dodrans.

τὸ τρίτον μέρος κενῆν ἀρχήματος
βουλή.

τεμνομερία, τεμνομερίον.

Dogma.

βουλή.

δῶγμα, δόξα.

Dolabra.

ἀξίνη, τζικούρι.

πέλεκυς, ἀξίνη, σμίλη, ξυίνη.

Dolabella.

τζικουράκι.

ἀξίναριον.

Doleo, dolito.

λυπῶμαι, πονῶ, παρὰ πονοῦμαι.

ἀλγέω, ὠ, ἀλγίζομαι, ἀνῶμαι.

Dolichus.

τὸ διάστημα δύο παδίων, ἢ τὸ φα-

δολιχός.

Dolium.

ποδάρι, βουτζί.

ζύλι.

πίδος, κέρας.

Doliarins.

βουτζάς, λαϊνάς.

πιδοποῖός, κερματοποῖός.

Dolo.

ξυῶ, ἰσιῶν.

πτεκίζω, ἀποξίζω, σκεδιάζω.

Dolon.

δολοῖμι μέσα εἰς τὸ ὅποιον ἦταν
ἔνα σπυρτί χρωσμοῦν.

δολῶν.

Dolor.

πίκρα, πόνος, βέσανον.

λύπη, ἄλγος, δόληδον, ἀγία.

Dolorificus, doloro-
fus.ὅπου χάνει πόνον, κοπιαστικός,
λυπητέρος.ἐπιδυμός, δόλινος, ὀδυνητός,
ὀπιπνός.

Dolofus.

δολερός, γέλαστος, ποτηρός.

δολιός, δολερός, χακούρτος.

Dolus.

γέλασμα, δόλος, πονηρία.

ἀπάτη, δόλος, ὑπαγῶν, ἐγκέρημα

Doma.

ἀπίτι.

δῶμα, οἶκος.

Domator, domitor.

ὅπου ἡμερώνει, ἡμερωτής.

δομαστής, δομάτωρ.

Domesticatim, dome-

ἀπίτι σὲ ἀπίτι, θάρετα.

κατ' οἰκηματα, οἰκείως, κατ' οἴκους.

Domesticus. [mesticē.

ἀπιτικός, τὸ ὁμοίου ἀπιτοῦ.

συνοίκος, οἰκείος, ἑνοίκος, σύσκητος.

Domicilium.

κατοίκημα, κατοίκησης.

οἴκησις, οἰκηθῆλον, κατοίκησις.

Dominus.

ἀφέντης, ἀφέντης.

κύριος, δεσπότης, ἀνάξ.

Domina.

ἀρχόντισσα, κυρία.

δεσποτὶς, δεσποῖνα, κυρία.

Dominicus.

τὸ ἀφέντης, ἀφεντικός.

κυριακός, δεσποτικός.

Dominor.

ἀφέντευω, εἰρίζω.

κατακυριεύω, δεσπόζω, ἀρχώ, κυ-
ριεύω.

Dominans, dominator

ὅπου ἀφέντευει.

ἀρχων, κυριεύων, ἀρχηγός, ἡγεμὸς.

Dominatus, us, domi-
natio, dominium.

ἀφέντια.

δυναστεία, τυραννίς, δεσποτεία,
ἀρχή.

Domiportus.

ὅπου φέρει τὸ ἀπίτι του μαζί, ὁ σά-
λιαχος.

φειροίκος.

Domito, domo.

ἡμερώνω.

ημεροσύνω, δομάζω, δομήμι.

Domitura, domitus, us.

ἡμέρωσις.

ἡμερωμα, δομήσις, δομασις.

Domus.

ἀπίτι.

οἶκος, δῶμα, δόμος.

Domitio.

τὸ γυμνασία εἰς τὸ ἀπίτι.

οἰκοφροσύνη.

Donarium, donatio,
donatium, donum.

δῶρον, χάρισμα, χάρις.

δωρεά, ἀνὰ δῶμα, δῶσις, δῶρον.

Dona-

I de følgende århundreder udgives en lang række to- og flersprogede ordbøger, alle i Vesteuropa, ofte med henblik på katolsk mission. Et eksempel er kapucinermunken Alessio da Somaveras [...] *Tesoro della lingua italiana e greca-volgare*, Paris 1709, en meget righoldig ordbog på ca. 1000 trespaltede sider. Efterhånden udkommer også mange mindre ordbøger, til brug både for grækere der vil lære fremmedsprog og for udlændinge der vil lære græsk. Ikke mindst efter at oprøret mod »tyrkerne« (det osmanniske rige) er startet i 1821, stiger antallet af tosprogede ordbøger og grammatiske fremstillinger af nygræsk til praktisk brug. Samme år kom den første trykpresse til Grækenland, hvor der i første omgang udelukkende tryktes aviser og andet til brug for frihedskampen.

1830: Den nye græske stat – dobbeltsprogethed

I 1830 blev Grækenland oprettet som den første suveræne nationalstat på Balkan, og dermed opstod behovet for et officielt nationalsprog. Græsk blev talt og skrevet i forskellige varianter i hele den vestlige del af det osmanniske rige, på de joniske øer i Adriaterhavet (Korfu, Zakynthos osv.) og blandt grækere i mange af Europas store byer. Sprogspørgsmålet havde længe været diskuteret, og svarene gik i tre retninger, der kan sammenfattes i at målet var:

- 1) at genoplive klassisk attisk græsk (med dativer, infinitiver osv.), en tanke der fik støtte fra nogle filhellenske kredse i Europa, men naturligvis ikke var fuldt realisabel;
- 2) at bygge på folkesproget – men forskellene fra egn til egn (f.eks. flere italienske ord i vest⁶ og mange tyrkiske elementer mod øst) stod i vejen, og folkesproget opfattedes af Vestens filhellenere, hvis støtte var af afgørende betydning for oprettelsen af den nye stat, som ”barbarisk” og ”fordærvet”
- 3) at tage udgangspunkt i det aktuelle sprog og gradvist ”forædle” det i retning af klassisk græsk.

Det var den tredje løsning der sejrede i første omgang. Dermed videreførtes skellet mellem dagligsprog og en tilnærmelse til ældre sprogformer. Grækenland fortsatte altså sin nedarvede dobbeltsprogethed: på den ene side en officiel, puristisk sprogform (katharévousa), der mere eller mindre lignede oldgræsk, og på den anden side folkesproget (dhimotikí), der i første omgang ikke fik nogen fast etableret form.

I 1835 udkommer den første nygræske ordbog, trykt i Athen og forfattet af den autodidakte leksikograf Skarlatos Byzantios, et stort værk på 446 sider, hvor de nygræske lemmata og udtryk forklares på katharévousa og desuden oversættes nødtørftigt til fransk. Senere udgav Byzantios også en oldgræsk og en græsk-fransk ordbog.

Først 50 år senere blev folkesproget for alvor anerkendt i (litterær) prosa med lingvisten Yannis Psycharis (Jean Psychari, 1854-1929) som drivende kraft. Katharévousa vedblev dog at være landets officielle sprog, og først da militærdiktaturet (»juntaen«) 1967-1974 brugte dette i en ekstrem, ofte latterlig

⁶ Stærkest på de joniske øer, men de kom først under Grækenland i 1864.

oppustet form, blev dets officielle status undermineret og derpå afskaffet i 1976, hvor dhimotikí fik status af nationalsprog. Dermed ophævedes forskellen mellem de to sprogformer, således at grækerne i dag frit kan anvende lærde eller formelle udtryk, ord og former, omtrent som vi kan vælge mellem danske ord og fremmedord.

Forskellen mellem de to sprogformer (katharévousa/dhimotiki) kunne være:

- ordets (bøjnings)endelse: Ελλάς/Ελλάδα (Ellás, Elládha) 'Hellas' ('Grækenland'), μουσείον/μουσείο (músíon/musío) 'museum' (med og uden slutnings-n)
- fonetisk: ir/er: σίδηρος/σίδερο (sídhiros/sídhero) 'jern'⁷
- helt forskellige ord: ύδωρ/νερό (ídhōr/neró) 'vand'⁸
- syntaks: katharévousa kunne være alt lige fra noget nær oldgræsk (ikke ulig tysk forvaltningssprog) over til talesprogsnær prosa.

Men katharévousa blev ikke en fast forankret skriftsprogform og endnu mindre et naturligt talesprog i spontan tale. Katharévousa varierede også efter periode, situation og politisk holdning⁹ på en skala fra 'meget oldgræsk' til noget tæt på et mere naturligt eller folkeligt sprog.

Denne sprogsituation afspejles naturligvis i ordbøgerne, både de en- og tosprogede: Indtil 1976 omfatter græsk-græske ordbøger i princippet sproget i hele dets lange historie, med katharévousa (eller oldgræsk) som grundform. Et heroisk, men velgennemført projekt, der dækker flere årtusinders sprog med oldgræsk som det bærende element, er de ni (senere femten) bind *Μέγα Λεξικόν Όλης της Ελληνικής Γλώσσης* (Mégha leksikón ólis tis ellinikís ghlóssis) "Stor ordbog over hele det græske sprog", af forlæggeren D. Dhimitrákos, der udkom 1933-1950. I samme periode udkom en mindre, men fremragende ordbog over det moderne sprog i to (senere tre) store bind, *Της Πρωίας*, (Tis Proías, navnet på en avis, "Morgenbladet"). Et eksempel til illustration af forholdet mellem de to sprogformer: 'en dør' hedder i dagligsprog πόρτα (pórta), på katharévousa θύρα (thíra). I Proias ordbog findes begge ord som opslag. Pórta fylder 14 linjer, og thía 22. Pórta defineres som synonym til thía, mens definitionen på thía er: "en åbning, hvorigennem man har adgang til et lukket eller adskilt rum, i dagligsprog pórtá". Begge ord er stadig i brug, men i dag er forklaringerne vendt om, f.eks. i Akademiets Ordbog (2014): thía (formelt sprog) = pórtá (ca. 32 linjer). Pórta (ca. 120 linjer) defineres (forkortet af mig) som "bevægelig plade, der lukker en indgang; selve indgangen".

Ét nationalt sprog

Efter afskaffelsen af dobbelsproget i 1976 kunne man endelig begynde at udgive ordbøger over det synkrone sprog inklusive de mange ældre ord og former der stadig anvendes. I de følgende år udkom en række gode etbinds ordbøger, dog ofte

⁷ Det er karakteristisk for udviklingen at de lærde former i dag ofte har fortrængt eller bruges parallelt med folkesprogsformen. 'Jern-' hedder til daglig σίδερο (sídhero), men som grundstof σίδηρος (sídhiros) og i sammensætninger næsten altid σίδηρο- (sídhiro-).

⁸ Den (erasmeiske) oldgræske udtale af ύδωρ var hýdōr.

⁹ F.eks. brugte konservative aviser udpræget katharévousa, venstreorienterede aviser dhimotikí, og centrumaviser et mere midtsøgende sprog. Alle aviser brugte dog dhimotikí i spalter med small-talk, vittigheder, opskrifter og lign.

med det formål at styrke dhimotikís overherredømme ("omvendt purisme"), uanset den faktiske sprogbrug. Antallet af tosprogede almene og fagspecifikke ordbøger for denne periode er helt uoverskuelig, ikke mindst efter at Grækenland blev medlem af EF i 1981.

Efter 1990 udkommer så tre store etbinds ordbøger efter moderne leksikografiske principper på ca. 2.000 sider hver, nemlig (i oversættelse):

- Ordbog over det almindelige nygræske sprog, Centre for the Greek Language/Manolis Triantafyllidis Foundation, Thessaloniki 1998¹⁰.
- Ordbog over det nygræske sprog, af Georgios Babinotis, Athen 1998, 2012.
- Athens Akademis ordbog, hovedredaktør Chr. Charalambakis, Athen 2014¹¹.

I dag er græsk leksikografi fuldt ud på niveau med andre europæiske sprog, samtidig med at den bygger på en meget lang tradition. Jeg vil derfor ikke beskrive dem nærmere, men til slut blot fremhæve tre særlige træk ved græsk, som spiller en rolle for leksikografien, nemlig fremmeordsproblematikken, oldgræsk og "modlån".

1. Fremmedord

Græsk har sit eget alfabet¹², og transskriptionen af fremmede ord kræver derfor altid at der træffes valg. Hovedspørgsmålet – men langt fra det eneste – er: Skal stavemåden eller udtalen gengives? Eksempelvis "Søren Kirkegaard" eller "Seren Kirkegor"?

Ældre lån er for længst naturaliserede og skrives derfor altid med græske bogstaver: μπαλκόνι (balkóni), μαγιονέζα (majonéza), μουσακάς (musakás, fra tyrkisk). Men hvad med de mange nye (især engelske) fremmedord? Skal man skrive dem med græske bogstaver eller med latinske? 'Computer' skrives normalt som κομπιούτερ (kompióuter), men langt hyppigere anvendes det græske ord υπολογιστής (ipolojistís). 'Lommeregner' hedder dog næsten altid κομπιουτεράκι (kompiuteráki), en diminutivform af κομπιούτερ. Akademiets ordbog har medtaget en lang række fremmedord i græsk ikklædning som f.eks. e-mail: ιμέιλ (iméil), men ordet skrives nok hyppigere med latinske bogstaver.

Transskriptionen af fremmedord har ændret sig over tid. Hvor man tidligere lagde sig tæt op ad den fremmede stavemåde, søger man i dag især at forenkle og ensrette stavningen, f.eks. gengives alle i'er som iota (i). 'Berlin' hedder fra gammel tid Βερολίνο (Verolíno), idet latinsk B skrives som græsk Β der repræsenterer [v], mens 'Biennale' skrives Μπιενάλε (bienále), idet ΜΠ (mp) udtales [b]. Nogle ord har begge muligheder: 'bombe' kan både hedde βόμβα (vómva) og μπόμπα (bómba), men bómba er på vej ud af sproget.

¹⁰ Den eneste af de tre, der (indtil nu) er tilgængelig elektronisk:

http://www.greek-language.gr/greekLang/modern_greek/tools/lexica/triantafyllides/

¹¹ Anmeldt af mig i SCANDINAVIAN JOURNAL OF BYZANTINE AND MODERN GREEK STUDIES, nr. 1, 2015, s. 131-141, ISBN 978-960-404-278-4.

¹² Mig bekendt bruges det græske alfabet i dag ikke til noget andet sprog. Det græske alfabet er et barn af fønikisk (jf. navnene på bogstaverne: alfa, beta, gamma ...) og moder til det latinske og til det kyrilliske, som primært anvendes i de slavisktalende lande der blev kristnet fra Byzans.

2. Rent oldgræske ord og udtryk

Mange udtryk både fra Biblen, fra de klassiske forfattere og rent oldgræske ord er stadig i brug i faste udtryk og citater. Således findes det oldgræske verbum οἶδα (oida, nygræsk udtale: ídha) 'jeg ved' ikke på nygræsk. Alligevel (og med rette) optræder det som lemma i alle tre ordbøger for at rumme et par faste udtryk som τις οἶδε (tis ídhe) 'hvem ved?' og Sokrates' berømte: "Jeg ved at jeg intet ved".

Den oldgræske flertalsform ὅτα (óta) 'ører' (jf. otorhinolaryngolog) står i to af ordbøgerne med henvisning til entalsformen οὖς (us), hvor de få faste (oldgræske) udtryk er forklaret. 'Øre' hedder i dag kun αὐτί (aftí).

3. "Modlån" (antidáneia)

En række ord er overtaget fra vesteuropæiske sprog (italiensk, fransk, engelsk), som (typisk via latin) har lånt dem fra oldgræsk. De mange ord af denne type giver et falsk indtryk af kontinuitet i ordforrådet over århundrederne og skaber problemer ikke mindst ved angivelsen af etymologi. Det gælder f.eks. λάμπα (lámpa) 'lampe', '(elektrisk) pære', der er lånt fra fransk/italiensk, men går tilbage til oldgræsk λαμπάς (lampás) 'fakkel'; samtidig har verbet λάμπω (lámpo) 'skinne' og en række afledninger heraf eksisteret lige siden oldtiden. Mange termer i videnskab, teknik, økonomi osv. er dannet af oldgræske elementer, såsom τηλέφωνο (tiléfono) 'telefon', egl. "fjernstemme" og volder ingen problemer. Græsk-latinske hybrider som 'television' eller 'automobil' er derimod for længst erstattet af rent græske sammensætninger: τηλεόραση (tileóراسι) og αυτοκίνητο (aftokínito), men fjernsyn kaldes også ofte TV (udt. tiví). Endelig er ikke så få omdefinerede eller nydannede ord af oldgræsk oprindelse i virkeligheden oversættelseslån, f.eks. εφάληριο (efaltírio) 'springbræt' efter fransk tremplin.

Undertiden opstår der konflikt mellem modlån og gammelt græsk ord: 'empati'¹³ stammer fra græsk, men det etablerede græske ord εμπάθεια (empáthia) betyder 'lidenskabeligt had', 'animositet', altså nærmest det stik modsatte. Psykologerne har derfor måttet danne et nyt kunstigt ord, ενσυναίσθηση (ensinésthisi) af *en* = 'i' og *syn-aisthesi* = 'med-føling', som imidlertid ikke er almindelig kendt og anvendt. Tilsvarende kan 'semantisk' ikke hedde σημαντικός (simantikós), for det betyder 'betydelig', 'vigtig'. Derfor har man dannet σημασιολογικός (simasiolojikós) og σημασιολογία (simasiolojía) hhv. 'semantisk' og 'semantik'.

Slutbemærkninger

Græsk leksikografis årtusindlange historie kan naturligvis ikke fremstilles blot nogenlunde fyldestgørende så kort som jeg her har forsøgt. Jeg har i denne anden del af oversigten helt måttet udelade en række felter, nemlig:

¹³ "Empathy: 1908, modeled on German Einfühlung [...], which was coined 1858 by German philosopher Rudolf Lotze (1817-1881) as a translation of Greek empathia "passion, state of emotion" [...] (Online Etymology Dictionary).

1. "Middelalderordbøger": værker fra nyere tid, der dækker (dele af) perioden fra slutningen af oldtiden op til det 18. århundrede¹⁴. Tekstmængden er gigantisk.
2. Etymologiske ordbøger, senest G. Babiniotis' værk fra 2009.
3. Onomasiologiske ordbøger, synonymordbøger osv. Berømt er Th. Vostantzoglous store "Antileksikon" fra 1962, der opstiller ordforrådet så vidt muligt i to spalter, f.eks. varme/kulde, forsoning/brud.
4. Dialektordbøger, herunder også om sproget på Cypern. Det græske Akademis store dialekt- og historiske ordbog blev påbegyndt i 1908, men indtil i dag er kun de fire første bogstaver (Α-Δ) publiceret.

Jeg håber hermed at have givet et lille indblik i de græske ordbøgers historie fra oldtiden indtil i dag. Mine egne *dansk-nygræsk* og *nygræsk-dansk ordbog* fra 1990'erne ville have haft stor gavn af de videnskabeligt baserede værker der er udkommet siden.

En tak skal her lyde til Kretas Universitet, som har gjort mange af de gamle ordbøger let og gratis tilgængelige på nettet (<http://anemi.lib.uoc.gr>). Endvidere har jeg haft stor nytte af <http://www.greek-language.gr> fra Centre for the Greek Language, Thessaloniki og af <http://dimartblog.com/2014/04/24/history-of-greek-typography>.

Rolf Hesse
dagr@rhesse.dk

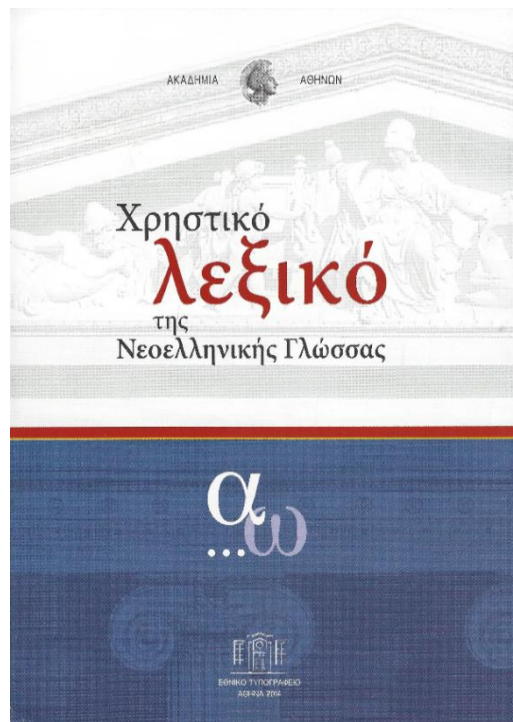


Fig. 4 "Brugsleksikon over det nygræske sprog" udgivet af Akademiet i Athen 2014.

¹⁴ F.eks. Ch. du Cange, *Glossarium ad scriptores mediae et infimae graecitatis duos in tomos digestum*, Lyon 1688, og E. Kriaras, *Λεξικό της μεσαιωνικής ελληνικής δημόδους γραμματείας 1100-1669*, 14 bind, Thessaloniki 1968-1997.

Kend din leksikograf! – Carsten Hansen

I dette nummer giver vi ordet til en alsidig leksikograf der har været involveret i mange forskelligartede leksikografiske projekter gennem årene, men som de seneste år har helliget sig talesproget. Og som desuden er en central figur ved diverse sociale arrangementer i leksikografiske kredse.

Hvordan og hvornår blev du interesseret i faget?

Det hele startede da jeg som 16-årig gymnasieelev fandt gloselisterne i tysk, engelsk, fransk og latin utilstrækkelige og til min 17-års fødselsdag ønskede mig ordbøger. Jeg fik kun Gyldendals røde engelsk-dansk, men supplerede selv næste gang jeg fik et mønsterbryderlegat: Det røde arsenal var komplet, men det bedste var at lærernes vidensmonopol var brudt! Da vi tyve år senere i 1989 stiftede LEDA, betroede Ole Norling mig: - *Nu skal du høre, Carsten. Udgivelsesåret for de røde ordbøger i de danske hjem er meget nemt at fastslå. Man ta_r ejerens fødselsår og lægger 17 år til. Boghandlerne solgte i de år så mange røde ordbøger at de altid havde den sidste udgave eller det næste oplag på hylden.* Jeg er født i 1951 og tjekkede straks mine syv røde første ordbøger. Ole havde ret.

Hvad kan du lide ved leksikografi?

Min fascination af ordbøger fik i firserne tilført nye teoretiske og metodisk-praktiske dimensioner.

Wiegands leksikografiseminar som jeg deltog i ved Germanistisches Seminar ved Universitetet i Heidelberg i sommersemestret 1980, faldt sammen med mit forskningsophold ved Institut für deutsche Sprache i Mannheim hvor jeg var i dialog med forskerne omkring det banebrydende leksikografiske værk Deutsche Wortbildung. Det part-of-speech-strukturerede firebindsværk var en stor åbenbaring for mig. For det første var det konsekvent korpusbaseret, for det andet var det både semasiologisk og onomasiologisk organiseret, og i tilgift var de enkelte lemmaers konkurrenceformer angivet.

Da det leksikografiske felt i Danmark i 1988 på Bergenholtz' og vel også Ole Norlings initiativ blev velsignet med DK87 og i de følgende år med yderligere 3x1 mio. tokens i repræsentative tekstuddrag, konverterede jeg definitivt fra litterat til leksikograf. Sikke et systemiseret domænespecifikt univers man helt efter egne lyster og orienteringer kunne udfinde på søgninger i Ugens Rapport.

En suppleringsuddannelse i datalingvistik i årene 87-89 hvor omdrejningspunktet var sproglige data i leksikografisk perspektiv, kvalificerede mig til først i 90'erne sammen med Jan Daugaard at parse Gads tyske WordPerfect-ordbøger og databaseficere dem så de kunne redigeres kategorivis og globalt, og ikke mindst så de nu også kunne rapporteres på alle tænkelige måder – også som færdig sats. Gads ledelse var næsten lige så begejstret som mig og Jan for dette kvalitetsløft af redigering og billiggørelse af korrektur og trykning.

Hvad bryder du dig ikke om ved feltet?

1. Som så mange andre forskningsområder er også leksikografien, det være sig inden for rammerne af forlagsverdenen såvel som i universitetsregi, ramt af tidens overordnede kapitalistiske maksimer

om profitmaksimering. Denne ramme spænder ben for kreativitet og nytænkning og udvirker dårlige arbejds- og ansættelsesforhold.

2. Korte deadlines, lange deadlines, små deadlines, store deadlines.

Hvilken ordbog kan du ikke undvære?

Den Danske Ordbog. Hver gang jeg frekventerer den, opfylder den mit behov. Når jeg som almindelig bruger søger efter en betydningsdefinition, går jeg sjældent forgæves. Selve det leksikografiske håndværk aftvinger altid respekt, og når redaktøren har valgt en mindre heldig løsning, kan jeg forstå hans tankegang. På ODT (Ordbog over Dansk Talesprog) bruger vi den også som frugtbar inspiration til vores egne artikler om interjektioner, og i denne brugsfunktion har vi for det meste den store glæde at kunne konstatere hvor meget stærkere vi står, fordi vi konsekvent og udelukkende opererer på autentiske sociolingvistisk annoterede talesprogsdata.

Fortæl en leksikografisk anekdote

I får sgu to:

For en del år siden spurgte jeg Poul Zerlang der netop havde redigeret en ny udgave af Bevingede Ord, om han nogen sinde havde plukket lidt i andre ordbøger. – *Nej. Det har jeg aldrig gjort, det går alt for langsomt. Jeg har altid skovet (smiling voice).*

Og den anden: Da jeg var på Gads Forlag, sad jeg engang først i 90'erne med direktøren i hans store, smukke, klassisk indrettede kontor med direkte elevatoradgang. Vi havde netop afsluttet drøftelserne omkring etableringen af min autonome ordbogsrepublik i koncernen da hans blik saligt drømmende vendte tilbage fra de forbipasserende nede på Strøget. – *Carsten, du komet på den leksikografiske stjernehimmel. Når vi når årtusindeskiftet, så går vi to ned på Strøget og spørger en række forbipasserende om de kender de røde ordbøger. De fleste vil svare nej, men enkelte fra de ældre årgange vil efter en tænkepause svare: – Nåååh ja, var det ikke Gyldendal der i sin tid udgav dem? I vore dage er ordbøger jo stribede.*

Hvad mener du om Leda og om Nordisk Forening for Leksikografi?

Både Leda og NFL har siden vi stiftede foreningerne, dannet rammer om et potent fagligt og socialt netværk der har haft umådelig stor indflydelse på mit leksikografiske virke og bidraget til min identitet som leksikograf. Jeg vil gerne benytte lejligheden til at udtrykke min dybeste respekt for alle de mange medlemmer der i årenes løb har investeret tid og kræfter i at arrangere konferencer og møder og i at redigere nyhedsbreve og rapporter. Tak skal I have!

Hvad er det værste du har oplevet som leksikograf?

Politikens Forlag engagerede mig i slutningen af 00'erne som freelancer til at redigere en tysk-dansk-tysk ordbog på et nyt og mere avanceret niveau end forlaget hidtil havde haft i sin profil. Desværre endte mit leksikografiske virke ved forlaget i en retssag anlagt af forlaget for misligholdelse af kontrakten. Sagen sluttede med fuld oprejsning, men den langvarige proces var et mareridt at håndtere undervejs. Forlagets egentlige motivation for opsigelsen af min kontrakt viste sig at være salget af hele forlagets ordbogsprogram til Ordbogen.com.

Hvad er det bedste du har oplevet som leksikograf?

Uden tvivl arbejdet med Ordbog over Dansk Talesprog.

For det første at jeg blev inviteret til at forestå den leksikografiske frugtbargørelse af Lanchartkorpusset. Efter det lettere rutineprægede arbejde på Gads Forlag og det ganske vist sjove,

omend traditionelle arbejde på Skoleordbog Tysk med Katrine Wohllebe, men frem for alt efter den traumatiske erfaring med ordbogsredaktionen på Politikens Forlag var det som at komme i den leksikografiske himmel da Frans Gregersen inviterede mig ind i varmen på Sprogforandringscentret:

- Carsten, vil du lave en talesprogsordbog over de 1600 samtaletimer i vores korpus? Du bestemmer selv hvordan den skal konciperes og hvor stor den skal være.
- Øh, Frans altså, hvornår skal den være færdig?
- Aldrig! Men du skal demonstrere og dokumentere hvordan en talesprogsordbog kan se ud og formidle det til feltet.

For det andet selve processen med at lave ordbogen. Det har været en meget stor oplevelse at arbejde i team med de mange kompetente studenterredaktører og andre medvirkende, ikke mindst Martin H. Hansen og Jan Daugaard. Alle er ligesom mig vokset i takt med de enkelte arbejdsopgaver og med ordbogen som helhed. Uden deres kreative engagement i ordbogen og dens vækst var ODT aldrig blevet til den ordbog det leksikografiske felt er blevet beriget med i dag.

Den sidste årsag til at arbejdet med ODT har været en stor oplevelse, er den store, ofte praktiske, opbakning og anerkendelse projektet har mødt fra det leksikografiske felt.

Har du en våd leksikografisk drøm?

Ordbog over Dansk Talesprog opererer i sin nuværende skikkelse på et udskrevet ortografisk korpus. De originale lydfiler tjener således som illustration af et skriftsprogskorpus. Min drøm er at vende forholdet om så det bliver muligt at søge direkte i lydfilerne med lyd som søgestreng. Udskrifterne kan så bruges til at facilitere talesproget i dets genuine form, nemlig lyd.

Carsten Hansen
Lektor, projektleder (ODT)
Nordisk Forskningsinstitut/Sprogforandringscentret, KU
carhan@hum.ku.dk



Nye ordbøger 2015

Af Merete K. Jørgensen

Anatomisk billedordbog af Heinz Feneis, oversat og bearbejdet af Peter Holm-Nielsen. 13. oplag. 2015

Beskæftigelsespolitisk ordbog. 3. udgave. Beskæftigelsesregion Midtjylland 2014

Dansk-bulgarsk ordbog (Datsko-bälgarski retjnik) af Lise Bostrup, Vladimir Stariradev. 4. udgave. 2015

Dansk-arabisk ordbog af Ellen Wulff, Noman Kanafani (Special-pædagogisk forlags indvandrerordbøger). 19. oplag. 2015

Dansk-engelsk ordbog for udlændinge (Danish-English Dictionary for Foreigners) af Lise Bostrup, Kirsten Gade Jones, W. Glyn Jones . 3. oplag. 2015

Dansk-kurdisk ordbog af Adel Saleh og Kim Vilstrup. 4. oplag. 2015

Dansk-somali ordbog (Qaamuus deenish-soomaali) af Abdulahi Mukhtar Hussein, Mohamed Abdulahi, A. Macaani, Mohamed Hassan. 5. oplag. 2015

Dansk-tamilsk ordbog af Annamalai Balamanoharan; dansk ordliste Lillian Fannikke Hansen, Hanne Schyth. 7. oplag. 2014

Den danske begrebsordbog af Sanni Nimb; medredaktører: Henrik Lorentzen, Liisa Theilgaard, Thomas Troelsgård. Det Danske Sprog- og Litteraturselskab. 2014

Engelsk-dansk bridgeordbog af Gunna Steffensen. Bridgebutikken, Møntergade 12, 1116 K. 2014

Engelsk-dansk ordbog for udlændinge (English-Danish Dictionary for Foreigners) af Lise Bostrup, Kirsten Gade Jones, W. Glyn Jones. 2. udgave 2015

Entrepriseretlig ordbog - dansk-tysk (Baurechtliches Wörterbuch - Deutsch-Danish) Maqs Law Film. 2015

Gammeldansk Ordbog, ved Det Danske Sprog- og Litteraturselskab
<http://gammeldanskordbog.dk>. 2015

Lexicon mediae latinitatis Danicae (Ordbog over dansk middelalderlatin) af Franz Blatt, Bente Friis Johansen, Holger Friis Johansen, Peter Terkelsen m.fl. Hæfte 7 med dansk og tysk indledning, Hæfte 8. 2014

Medicinske fagudtryk. En klinisk ordbog med kommentarer af Jan O. Rytter Nørgaard, Klaus Levin; redaktør: Ib Sewerin. 4.udgave. Dansk Sygeplejeråd. 2015

Min første visuelle ordbog med engelske ord oversat af Fanny Bruun. 2015

Morsingbo-ordbog: de ær sø'en, vi sæjje af Hanne Glintborg. Morsø. 2015

Politikens krydsogtværsordbog af Holger Scheibel og Henriette Sonne. 7. oplag. 2015

Politikens retskrivnings- og betydningsordbog. 2 oplag. 2014

Psykologisk pædagogisk ordbog af Mogens Hansen, Poul Thomsen, Ole Varming. 17. udgave. 2015

Russisk-dansk ordbog af Jens Kure (f. 1877), genoptryk. Arkona. 2015

Ungarsk-dansk ordbog (Magyar-dán szótár) af Ágnes Sitkeiné Szira, Margit Nielsen, József Szira. 2. reviderede udgave. 2014

Merete K. Jørgensen
Seniorredaktør
Det Danske Sprog- og Litteraturselskab
mkj@dsl.dk



The LREC Conference with workshops and tutorials will be held in Portorož, Slovenia during the week of 23 to 28 May 2016.

Since the first LREC held in Granada in 1998, LREC has become the major event on Language Resources (LRs) and Evaluation for Language Technologies (LT). The aim of LREC is to provide an overview of the state-of-the-art, explore new R&D directions and emerging trends, exchange information regarding LRs and their applications, evaluation methodologies and tools, ongoing and planned activities, industrial uses and needs, requirements coming from the e-society, both with respect to policy issues and to technological and organisational ones.



The XVII EURALEX International Congress will be held 6 - 10 September 2016 in Tbilisi, Georgia.

The Congress will be organized by the Lexicographic Centre at Ivané Javakhishvili Tbilisi State University.

The EURALEX Congresses bring together professional lexicographers, publishers, researchers, software developers and others interested in dictionaries of all types.

The programme will include plenary lectures, parallel sessions on various topics, software demonstrations, a round-table discussion, pre-congress tutorials and specialized workshops, a book and software exhibition as well as social events for participants and their guests.