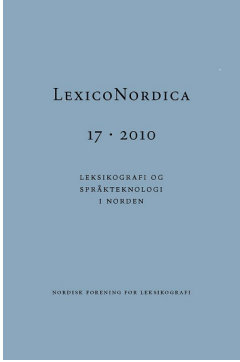


# LexicoNordica

Titel:	Nogle bemærkninger til Henning Bergenholtz: "Hurtig og sikker tilgang til informationer om ordforbindelser" i LexicoNordica 16	
Forfatter:	Christian Becker-Christensen	
Kilde:	LexicoNordica 17, 2010, s. 395-403	
URL:	<a href="http://ojs.statsbiblioteket.dk/index.php/lexn/issue/archive">http://ojs.statsbiblioteket.dk/index.php/lexn/issue/archive</a>	

© LexicoNordica og forfatterne

## Betingelser for brug af denne artikel

Denne artikel er omfattet af ophavsretsloven, og der må citeres fra den. Følgende betingelser skal dog være opfyldt:

- Citatet skal være i overensstemmelse med „god skik“
- Der må kun citeres „i det omfang, som betinges af formålet“
- Ophavsmanden til teksten skal krediteres, og kilden skal angives, jf. ovenstående bibliografiske oplysninger.

## Søgbarhed

Artiklerne i de ældre LexicoNordica (1-16) er skannet og OCR-behandlet. OCR står for 'optical character recognition' og kan ved tegngenkendelse konvertere et billede til tekst. Dermed kan man søge i teksten. Imidlertid kan der opstå fejl i tegngenkendelsen, og når man søger på fx navne, skal man være forberedt på at søgningen ikke er 100 % pålidelig.

# Nogle bemærkninger til Henning Bergenholtz: “Hurtig og sikker tilgang til informationer om ordforbindelser” i LexicoNordica 16

*Christian Becker-Christensen*

## 1. Indledning

LexicoNordica 16 havde som tema ordforbindelser i nordiske ordbøger. Blandt de behandlede emner var klassifikationen af ordforbindelser og præsentationen af de forskellige typer i diverse ordbøger.

Et af bidragene var Henning Bergenholtz’ *Hurtig og sikker tilgang til informationer om ordforbindelser* (Bergenholtz 2009) der har brugertilgangen til ordforbindelser som emne. Dette specificeres nærmere sådan at “Tilgangen og ikke tilgangsstrukturen og heller ikke makrostrukturen er temaet” (op.cit.: 30). Artiklen omhandler en test (et “tilgangseksperiment”) der registrerer en forsøgspersons “søgevej” og “tidsforbrug” ved søgning på 10 faste vendinger (idiomer og ordsprog) i otte forskellige danskordbøger hvoraf de fem er papirudgaver som angives at være:

*Nudansk Ordbog* (2005)

*Den Danske Ordbog* (2003-2005)

*Ordbog over det danske Sprog* (1918-1956)

*Talemåder i dansk* (1998)

*Danske Talemåder* (1998)

og de tre er elektroniske udgaver, nemlig:

netudgaven af *Ordbog over det danske Sprog* (2008)

og Henning Bergenholtz mfl.s to netordbøger:

*Den Danske Netordbog* (2008)

*Betydning af faste vendinger* (2009)

Opgaven der stilles, er at forsøgspersonen bliver bedt om “at finde ud af, hvad bestemte faste vendinger betyder, hhv. om probanden kan finde vendingen i udvalgte opslagsværker” (op.cit.: 34, 36). Alle de valgte ordbøger er blevet testet i formentlig en bestemt rækkefølge; det indgår altså ikke i testen at forsøgspersonen har skullet vælge hvilken ordbog der ville være mest relevant for besvarelsen af opgaven.

Som resultater af undersøgelsen kan udledes:

- at søgetiden er mindre når søgningen på en ordforbindelse giver positivt resultat end når den giver negativt resultat (jf. op.cit.: 44, 49),
- at elektroniske versioner (med direkte søgning på en ordforbindelse eller en del af den) er hurtigere at søge i end papirudgaver (og den elektroniske version af *Ordbog over det danske Sprog* hvor man skal scrolle gennem den fundne artikel) (jf. op.cit.: 48, 51),
- og at de involverede ordbøger fordeler sig i en rangordning efter “positivt resultat + tidsforbrug i gennemsnit” hvor *Betydning af faste vendinger* og *Den Danske Netordbog* får den bedste score, mens *Nudansk Ordbog* og *Talemåder i dansk* ligger i bunden (op.cit.: 50).

Bergenholtz’ testbeskrivelse er gengivet i *Hermes* 44, 2010, i artiklen Bergenholtz & Gouws: “A New Perspective on the Access Process”.

Jeg har disse bemærkninger til testens resultater og tilrettelæggelse:

## 2. Søgning med negativt resultat

Undersøgelsen omfatter opgørelse af forsøgspersonens tidsforbrug ved søgninger med såvel positivt som negativt resultat, jf. definitionen af søgetid som “tidspunktet, hvor behovet for en ordbogskonsultation opstår, hhv. når eksperimentets indhold fremlægges, indtil det herudfra opståede spørgsmål er blevet besvaret, hhv. at søgningen opgives resultatløs” (op.cit.: 33). Men i det sidste tilfælde, søgning med “negativt resultat”, inkluderer undersøgelsen også søgning efter ordforbindelser som ikke forekommer i undersøgelsesmateriale.

Fx bruger forsøgspersonen 2 minutter og 17 sekunder minutter på at lede efter vendingen “barn på gule plader” som ikke figurerer i *Nudansk Ordbog*, og 4 minutter og 36 sekunder på at lede forgæves i *Ordbog over det danske Sprog* hvor vendingen naturligt nok heller ikke forekommer. I de to specialordbøger *Danske Talemåder* og *Talemåder i dansk* som eksplicit ikke medtager ordsprog, giver søgning på “liden tue kan vælte stort læs” og “hvor godtfolk er, kommer godtfolk til” ligeledes lange søgetider.

I *Betydning af faste vendinger*, hvor de nævnte ordforbindelser findes, er søgetiden henholdsvis 15 sekunder, 16 sekunder og 16 sekunder.

Der redegøres ikke i undersøgelsen for i hvilke tilfælde de søgte sekvenser forekommer i undersøgelsesmateriale, og i hvilke tilfælde de ikke eksisterer i materialet. Det er på denne baggrund uklart hvad der menes med “tilgang til information om ordforbindelser”. Har “tilgang” at gøre med strukturen i ordbøgerne, altså på hvilken måde genstanden for en søgning præsenteres i ordbogen? Eller betyder “tilgang” at undersøgelsen drejer sig om ordbøgernes repræsentativitet: hvornår ordbogen er udkommet, dens størrelse,

hvilken type der er tale om, og hvilken brugergruppe den er tiltænkt?

Hvis man mener at en test af søgetider giver bedst mening i forhold til det første, bedømmelsen af og deraf videreudviklingen af ordbøgers tilgangsstruktur, så bør en test af søgetider være begrænset til sekvenser som findes i de testede ordbøger og bør ikke sammenblandes med en undersøgelse af om en bestemt sekvens forekommer i en bestemt ordbog eller ej.

Et andet forhold er at der med "tilgang til information om ordforbindelser" måles på om en ordforbindelse har en betydningsforklaring og ikke blot om man kan finde den pågældende ordforbindelse, dvs. fundne ordforbindelser uden betydningsforklaringer tæller negativt. Hvis man igen – trods Bergenholtz' indledende bemærkning om at temaet er "Tilgangen og ikke tilgangsstrukturen" (op.cit.: 30) – tænker sig at testen skulle have noget at gøre med tilgangen til ordforbindelser, så burde det alene være forekomst af sekvensen der var relevant for søgningen og ikke om der var angivelser knyttet til den.

Man kan venligt bemærke at disse forhold udgør metodiske uklarheder. Eller man kan spørge hvad man egentlig skal bruge testen til. Det er under alle omstændigheder vanskeligt at se værdien af den samlede opgørelse af gennemsnitssøgetider hvori indgår negativt resultat: "tidsforbrug i gennemsnit" og "negativt resultat + tidsforbrug i gennemsnit" (tabellen op.cit.: 50).

### 3. Udvalget af testede ordbøger

*Ordbog over det danske Sprog* er en historisk ordbog i 28 bind som spænder over perioden fra ca. 1700 til redaktionstidspunktet (1918-1956) med tilsvarende lange og layoutmæssigt komprimerede artikler; *Betydning af faste vendinger* er en elektronisk ordbog over netop ordforbindelser som dem der testes for. Sammenstillingen

og sammenligningen af så forskellige ordbogstyper i en og samme tilgangsundersøgelse er besynderlig. For at en test af søgetider skal have nogen mening, bør undersøgelsen forholde sig til typer af ordbøger ved at holde sig til en afgrænset type eller ved eksplicit at sammenholde forskellige typer og skal ikke sammenblende forskellige ordbogstyper på den måde som det gøres i testen. Inden for testens rammer kunne testmaterialet lige så vel have medinddraget Leths *Dansk Glossarium* (1800), som nævnes som eksemplet på en specialordbog til brug for tekstreception (op.cit.: 31).

#### 4. Angivet og faktisk testede ordbøger

To ordforbindelser “Liden tue kan vælte stort læs” og “Hvor godtfolk er, kommer godtfolk til” giver negativt resultat for *Nudansk Ordbog* (med deraf følgende lange søgetider, henholdsvis 1 minut og 26 sekunder og 2 minutter og 12 sekunder) (op.cit.: 44-45).

Men begge står i den udgave af *Nudansk Ordbog* som ifølge præsentationen af forsøgsmaterialet (op.cit.: 35) og ifølge litteraturlisten (op. cit.: 53) er genstand for testen, nemlig 19. udgave, 2005:

**tue** [*tu·ə*] sb. -n, -r, -rne

□ en lille forhøjning i jordoverfladen ◇ *engen var fuld af tuer · de måtte springe fra tue til tue for ikke at få våde fødder* ◇ **græstue** · **myretue**

ORDSPROG: **liden tue kan vælte stort læs** ◇ *sagen kan blive den tue der vælter organisationens læs*

**godtfolk** sb.

□ (ældre) jævne, hæderlige mennesker ◇ *til stede var læger, sygeplejersker og andet godtfolk*

- **være kommet af godtfolk** (ældre) være af god familie

ORDSPROG: **hvor godtfolk er, kommer godtfolk til**

Nudansk Ordbog, 19. udgave, 2005

I stedet for den angivne 19. udgave er forsøgspersonen blevet forelagt en tidligere udgave af Nudansk Ordbog hvori ikke er medtaget informationstypen ordsprog, og hvor de to ordforbindelser derfor ikke forekommer. At der i testen ikke er tale om den angivne 19. udgave af *Nudansk Ordbog* fremgår desuden af side 44 hvor det siges at “probanden ikke kunne forstå, at der stod en betydning 1, men ikke nogen betydning 2”. Denne praksis for markering af overgangen mellem hoved og krop med tallet “1” uanset om der fulgte flere betydninger efter eller ej, blev benyttet i Nudansk Ordbog 17. udgave, 1999, og 18. udgave, 2001. I Nudansk Ordbog 19. udgave markeres overgangen ved en firkant når artiklen ikke er opdelt i talnummererede betydninger.

Ud fra forsøgets præmis, at fundet af betydningen af ordforbindelserne er kriterium for måling af tidsforbruget ved søgning på ordforbindelserne, vil rettelse af denne fejl i testen dog ikke ændre på resultatet “negativ” i de to tilfælde.

## 5. Rangordning mellem ordbøgerne

Bergenholtz’ sammenfattende tabel viser “en rangordning mellem ordbøgerne, som har et positivt resultat som vigtigste og søgetiden som næstvigtigste argument” hvilket “giver en rangordning, hvor ordbøgerne er anført med den – i denne test – bedste ordbog øverst” (op.cit.: 49-50). Her lander *Betydning af faste vendinger* som den bedste ordbog med 10 positive søgninger og et gennemsnit på 24 sekunder. *Nudansk Ordbog* placeres på en syvendeplads

med 5 positive søgninger og en gennemsnitssøgetid på 36 sekunder lige over *Talemåder i dansk* som har 3 positive søgninger med et gennemsnit på 30 sekunder.

Men hvis i stedet søgetid med positivt resultat, dvs. søgetid på de faktisk forekommende ordforbindelser i ordbøgerne, blev lagt til grund for en lidt mere meningsfuld opgørelse, ville ordningen efter gennemsnitlige søgetider være denne:

<i>Betydning af faste vendinger</i>	20"
<i>Den Danske Netordbog</i>	24"
<i>Talemåder i dansk</i>	30"
<i>Nudansk Ordbog</i>	36"
<i>Den Danske Ordbog</i>	40"
<i>Ordbog over det danske Sprog</i> (netudgave) <sup>1</sup>	41"
<i>Danske Talemåder</i>	48"
<i>Ordbog over det danske Sprog</i> (papirudgave)	2' 35"

Bergenholtz viderebringer resultatet af sin test i den efterfølgende artikel i *LexicoNordica* 16: Bergenholtz & Bjærge (2009:59) med denne præsentation: "resultaterne af et eksperiment med søgninger i otte danske ordbøger efter 10 faste vendinger". Men nu citeres testen alene med en rangordning der viser gennemsnitlige søgetider der (ukommenteret) inkluderer søgninger med negativt resultat, altså søgetider for søgninger på ikkeforekommende ordforbindelser. I rangordningen her står *Nudansk Ordbog* på en sjetteplads med en gennemsnitssøgetid på 1 minut og 3 sekunder, et resultat der er baseret på en metodisk dubiøs inkludering af resultater af søgning på ikkeeksisterende forekomster i testmaterialet.

1 At netudgaven af *Ordbog over det danske Sprog* har en så markant bedre søgetid end papirudgaven, skal nok ses i lyset af at forsøgspersonen har fået forelagt papirudgaven først, før versionen som netudgave, og dermed har fået kendskab til hvor i ordbogen de testede ordforbindelser findes (jf. rækkefølgen i eksperimentbeskrivelsen op.cit.:37 og i tabellerne for søgning på de enkelte ordforbindelser).



## 6. Afsluttende bemærkninger

Bergenholtz slutter sin artikel med at der i bidragene i LexicoNordicas temadel “er for meget fokus på selektering, da faste vendinger ikke kan selekteres fuldstændigt, hvis man da ikke kommer op på hundrede tusinde lemmatiserede eller kodificerede faste vendinger i den enkelte ordbog. I stedet bør der lægges større vægt på at forske i tilgang, særligt i søgetider, og hvordan en ordbogs layout kan bidrage til lavere søgetid, så der kan udarbejdes mere hensigtsmæssige ordbøger med en hurtigere og sikrere tilgang” (Bergenholtz 2009:51).

Men Bergenholtz’ løst tilrettelagte tilgangseksperiment i LexicoNordica 16 giver netop ikke afsæt til dette her erklærede formål da han ikke (ud over sporadiske bemærkninger) forholder sig til tilgangsstrukturen i de testede ordbøger (fx placering, ordning og markering af ordforbindelserne) og søgetiderne derfor ikke kan relateres til forskelle i ordbøgernes layout.

## Referencer

- Bergenholtz, Henning 2009: Hurtig og sikker tilgang til informationer. I: *LexicoNordica* 16, 29-54.
- Bergenholtz, Henning & Esben Bjærge 2009: Konception af fire monofunktionelle ordbøger med faste vendinger. I: *LexicoNordica* 16, 55-73.
- Bergenholtz, Henning & Rufus Gouws 2010: A New Perspective on the Access Process. I: *Hermes – Journal of Language and Communication Studies* no 44-2010.
- Betydning af faste vendinger* 2009. [www.idiomordbogen.dk](http://www.idiomordbogen.dk).
- Danske Talemåder 1998 = Allan Røder: *Danske Talemåder*. København: Gad.

- Den Danske Netordbog* 2009. [www.ordbogen.com](http://www.ordbogen.com).
- Den Danske Ordbog* 2003-2005. København: Det Danske Sprog- og Litteraturselskab & Gyldendal.
- Leth, Jens Høier 1800: *Dansk Glossarium*. København: Trykt paa Hofboghandler Simon Poulsens Forlag.
- Nudansk Ordbog 1999 = *Politikens Nudansk Ordbog*. 17. udg. København: Politikens Forlag.
- Nudansk Ordbog 2001 = *Politikens Nudansk Ordbog*. 18. udg. København: Politikens Forlag.
- Nudansk Ordbog 2005 = *Politikens Nudansk Ordbog*. 19. udg. København: Politikens Forlagshus.
- Ordbog over det danske Sprog* 1919-1956. København: Gyldendalske Boghandel. Online: [www.ordnet.dk](http://www.ordnet.dk)
- Talemåder i dansk 1998 = Stig Toftgaard Andersen: *Talemåder i dansk*. København: Munksgaard.

Christian Becker-Christensen  
cand.mag.  
Børghlumvej 39  
DK-2720 Vanløse  
[cbcvanlose@gmail.com](mailto:cbcvanlose@gmail.com)