



Gak til korpus og bliv viis

Om korpora, sprognormer og tysk

Farø, Ken Joensen

Published in:

Rette ord - Festskrift til Sabine Kirchmeier-Andersen i anledning af 60-årsdagen

Publication date:

2015

Document version

Også kaldet Forlagets PDF

Document license:

[Ikke-specificeret](#)

Citation for published version (APA):

Farø, K. J. (2015). Gak til korpus og bliv viis: Om korpora, sprognormer og tysk. I E. Skafte Jensen, D. Duncker, & O. Ravnholt (red.), *Rette ord - Festskrift til Sabine Kirchmeier-Andersen i anledning af 60-årsdagen: Dansk Sprognævns skrifter 46* (s. 109-124). Dansk Sprognævn. Dansk Sprognævns skrifter, Nr. 46

Rette ord

**Festskrift til
Sabine Kirchmeier-Andersen
i anledning af 60-årsdagen**

Redigeret af
Dorthe Duncker, Eva Skafte Jensen og Ole Ravnholt

Dansk Sprognævns skrifter 46

© Dansk Sprognævn og forfatterne 2015

www.dsn.dk

Tryk: GP-tryk A/S

Dtp: Jannerup A/S

Layout: Anna Falcon

ISBN: 978-87-89410-64-7

ISSN: 0415-0155 (Dansk Sprognævns skrifter)

Indholdsfortegnelse

- 13 Gratulationsliste
- 21 Særskrivningsfejl i de nye medier og i skolen
Af Margrethe Heidemann Andersen
- 31 'Jeg kan bare godt lide dansk'. Dansk som fremmedsprog på udenlandske universiteter – holdninger og udfordringer
Af Nina Møller Andersen
- 49 Ord og ordforbindelser, som ”hører nøje sammen med en efterfølgende ledsætning”. Om startkommaets placering i henhold til Retskrivningsordbogens § 50.1 og 50.3
Af Christian Becker-Christensen & Jørgen Schack
- 73 Den terminologiske værktøjskasse set ud fra forskellige målgruppers behov. Eksemplificeret ved begrebernes udtryksside og valensmønstre som søgemønstre
Af Lotte Weilgaard Christensen
- 87 Hvad blev der af adjektanalysen?
Af Niels Davidsen-Nielsen
- 97 I ulige linje fra *Linje* til *linje*. Og andre nedslag i Retskrivningsordbøger gennem historien
Af Philip Diderichsen, Anna Sofie Hartling, Anne Kjærgaard & Anna Kristiansen
- 109 Gak til korpus og bliv viis: om korpora, sprognormer og tysk
Af Ken Farø
- 125 Hvad er forudsætningerne for en (u)mulig retskrivningsreform?
Af Frans Gregersen

- 141** **Neksus – en ubestemmelig kategori**
Af Hans Götzsche
- 155** **Beholder og indhold. En semantisk anekdote**
Af Erik Hansen
- 159** **Talebob – den tålmodige transnordiske udtaletræner**
Af Peter Juel Henriksen & Auður Hauksdóttir
- 171** **Hvad vi dog slås om. Sprogkamp(e) i Danmark i 350 år**
Af Henrik Galberg Jacobsen
- 187** **Videnskab, normativitet og formidling**
Af Per Anker Jensen
- 205** **Forståelige og imødekommende tekster til borgere og kunder – er der forskningsmæssigt belæg for bestræbelsen?**
Af Anne Kjærgaard
- 219** **Nordisk sprogsamarbejde gennem 40 år. Uformelle iagttagelser om det nordiske sprogsamarbejde med særligt henblik på de nordiske sprognævn i tiden efter 1974**
Af Jørn Lund
- 227** **Fra tekst til struktureret viden**
Af Bodil Nistrup Madsen & Hanne Erdman Thomsen
- 247** **Eksperimenter med et skalérbart betydningsinventar til semantisk opmærkning af dansk**
Af Bolette S. Pedersen, Sanni Nimb & Sussi Olsen
- 261** **Det typiske drab i ridderviserne. Et folkevisord i kontekst**
Af Hanne Ruus
- 273** **Valens og konstruktion**
Af Lene Schøsler

- 285** Dansk Sprognævn og Danmarks Radio. Mastodonter i normeringen af
”det danske sprog”
Af Jacob Thøgersen
- 301** En drøm om ideologisk valens
Af Ole Togeby
- 315** Redaktionel note

Gak til korpus og bliv viis: om korpora, sprognormer og tysk

Af Ken Farø

Dette bidrag demonstrerer og diskuterer et par muligheder for udnyttelse af korpora inden for tysk som fremmedsprog.

Fremmedsprog og korpuserkompetence

At kunne anvende et elektronisk tekstkorpus bør ikke indgå i kompetencemålene i fremmedsprogsundervisningen i grundskolen. Selv gymnasiet er næppe tidspunktet, hvor man som lærer skal kunne anvende et korpus: Der er rigeligt med basisproblemer at se til. Men på universitetet bør tekstkorpora høre til sprogstuderendes standardarbejdsredskaber, på linje med grammatikker, ordbøger, bibliografier og andre hjælpemidler. Det skyldes, at studerende ideelt og forhåbentlig også i praksis netop trænes til selvstændighed og kritisk omgang med viden og formidlingen af den, herunder med kilder og faglige autoriteter. Studerende skal desuden vænnes til at arbejde med naturligt sprog, til at overskue større datamængder og til at stille undersøgelses spørgsmål til sprogmateriale og indrette korpussøgninger strategisk derefter.

Et elektronisk tekstkorpus er en ressource, der kan anvendes til på systematisk, præcis, overskuelig og kontrollerbar vis at afgøre tvivlsspørgsmål, at undersøge betydning ud fra brugskontekster, at vurdere ords, udtryks og konstruktioners udbredelse, at fastslå hvilke komponenter i en ordforbindelse, der er faste, og hvilke, der er variable, at gennemskue syntaktiske mønstre o.m.a.

I forhold til de gængse fremmedsprogskompetencer, nemlig de 4 basale sprogfærdigheder (lytte- og læseforståelse, skrive- og talefærdighed), så er det i første omgang skrivefærdigheden (= den skriftlige produktionsevne), der kan få gavn af korpusanvendelse. Skrivefærdigheden er kendetegnet ved ofte at være asynkron, dvs. skriveudførelsen sker ikke samtidig med plan-

lægningen af skriveakten. Det gør det overhovedet meningsfuldt at konsultere et korpus her sammenlignet med fx aktiviteter inden for den mundtlige sprogfærdighed, som typisk er spontan (undtagen i mediant mundtlige, men konceptionelt skriftlige kontekster som taler og manuskriptbaserede forelæsninger). De fleste korpora er desuden primært skriftsprogsbaserede og -orienterede og dermed ikke specielt gavnlige for den mundtlige sprogfærdighed. Herudover er det især inden for de mere sprogbevidsthedsmæssige, analytiske og operationelle dele af fremmedsprogsindlæringen, at korpora kan yde hjælp. Det vil konkret sige sprogbeskrivelse, sprogsammenligning og oversættelse.

Om brug af dansk korpus i tysk

Vi lægger ud med det tilsyneladende mindst oplagte: Man kan sagtens få brug for et dansk korpus i tyskfaget, fx i forbindelse med oversættelse fra tysk. Det skyldes, at der netop i denne disciplin ofte opstår tvivl om, hvad der i grunden er grammatisk og idiomatisk dansk¹. Det hænger igen sammen med den naturlige interferens, der er en næsten uundgåelig bestanddel af arbejdet med fremmedsprog. Har man ikke et meget effektivt filter, eller har man ikke til stadighed et stærkt metafokus på interferens, mister man paradoksalt nok som tysk-beskæftiget i nogle tilfælde en fornemmelse for, hvad der er gangbart dansk sprog. Det er derfor en kompetence, der skal trænes og holdes ved lige i undervisningen. Navnlig i oversættelse, hvor man gang på gang kommer ud i situationer, hvor sprogene og deres forskellige inventarer er inkongruente. Her ser man ofte studerende aflevere tekstprodukter, der er helt ude af trit med selv deres egen sprogfornemmelse, sproglige logik, ja endda med den danske basisgrammatik. Det behøver dog ikke at være så grelle eksempler. Herudover, og måske navnlig til de mere subtile spørgsmål, kan korpora yde hjælp, hvis fx ordbøgerne svigter, hvad de ofte gør, når det bliver subtilt:

I en oversættelsesopgave forekom overskriften *Berlin macht Lust auf mehr*. En studerende havde, ikke helt skævt synes jeg, oversat originaltitlen med *Berlin giver blod på tanden*. Alligevel var der noget ved den sætning, jeg syntes ikke fungerede. Men jeg kunne i situationen ikke se, hvad det var. En senere søgning i KorpusDK tilvejebragte de manglende oplysninger. Ved hjælp af en konkordans over korpus' i alt 112 realiseringer af frasemet med kontekst

er det muligt at analysere sig frem til fællesnævneren for alle disse frasemto-
kens:

har givet Bent Christensen og hans kolleger i Envision **blod på tanden**. Næste arrangement, der skal have en på markedsførings-hatten, kører hver lørdag har I midlertid givet de tre venner **blod på tanden** og samtidig også penge på bankkontoen til at realisere ønsket også en Grammy som årets sangskriver. Det har givet ham **blod på tanden**, så han er igang med nye numre i hjemmestudiet i bag sig. "Allerede min første rolle i Frændeløs gav mig **blod på tanden**. Og jeg ved nu, at jeg vil være skuespiller ", fastslår mest for sjov, men det her har virkelig givet mig **blod på tanden** ", siger en af deltagerne, 24-årige Per Hagerup, der nu overvejer talt givet den serbisk dominerede forbundshaar **blod på tanden**. Store troppestyrker, kampvogne og artilleri blev i aftes rykket med omgang. DA eventyret i Manchuriet gik godt, fik japanerne **blod på tanden**. I slutningen af trediveerne var store dele af Kina under arenaen, da jeg kom til, og de havde sandelig fået **blod på tanden**. Som et led i arrangementet får læreren udleveret en løsladmap kr. Det positive resultat har givet redningskorpset **blod på tanden** til at etablere fast udrykningshold til skibsbrande. For at gøre Vi fik succes med de første ting, hvilket gav ekstra **blod på tanden**. Er bankfolk mere alvorlige end folk er flest? "Vi er se så mange gode billeder. Så nu har Olympus fået **blod på tanden**. Til 1990's Internationale fotokonkurrence er der tre kategorier med mange voksne, som ikke lærte at svømme i deres skoletid, **blod på tanden** til at kaste sig i det våde element og få i Paris-Tours har atter givet Rolf Sørensen **blod på tanden**. Tror på ny succes i dagens World Cup-løb i Norditalien. for juniorer, og det har for alvor givet hende **blod på tanden**. "Hendes store drøm er at få sin egen træner, lige hemmelig. Det lørd spændende og gav hende endnu mere **blod på tanden**. Han fulgte hende ud til bilen. Solen var ved at et halvt eller et helt minut. Men det giver nok **blod på tanden** at konkurrere lidt, med andre og med sig selv. Jeg danske koloni. De tre år i Bruxelles gav Doris Mulvad **blod på tanden** til at rejse til udlandet igen, og hun kan sagtens derfra med et smil på læberne. Nu har jeg fået **blod på tanden**. Jeg kommer i tanke om en tv-udsendelse, hvor en fantastisk udvikling i al almindelighed. Så vi fik altså " **blod på tanden** " for princdpperne i den slags arrangementer, da en af mine talsmand. Han startede skolebladet, og det gav ham vel **blod på tanden**, så nu er han blevet lokalredaktør ved den rigtige avis. succesen både i- 86 og- 87. Nu har Claus fået **blod på tanden**. Han vil ikke længere nøjes med DM-titlerne. Han vil for er sådan noget, de laver, for at man skal få **blod på tanden** og købe bladet igen i næste uge. Den tredje og med den, der foregik i Aktivbanken, der giver spekulanter **blod på tanden**. Også i Hope er der spekulation i et indløsningsstilbud, og ungerne stod i kø. Derefter havde Regner Grasten fået **blod på tanden** til selv at leje film ud, og han startede med "

Fig. 1: Søgning på frasemet *blod på tanden* på KorpusDK.

Gennemgående gælder for disse kontekster: Der er et sem [+ udfordring] til stede. Der er altså mht. *blod på tanden* ikke bare tale om lyst til noget, men om *motivation for at tage en udfordring op*. En sådan udfordring udgør et rejsemål som Berlin næppe i sig selv; det ville kræve, at byen indgik som et senere punkt i en serie af (minimum 2) udfordringer, hvoraf den sidste var den største.

Med sådan en undersøgelse i baghånden, der ikke direkte modsiger hypotesen, men tværtimod ser ud til at underbygge den, er man på mere sikker grund, end hvis man alene stoler på sin fornemmelse. Det er ikke det samme, som at man har fundet Sandheden, for den finder vi ikke, som Popper (1959) understregede, i hvert fald ikke i sprogforskningen. Men mindre kan også gøre det.

Når den officielle normbeskrivelse ikke slår til

Ifølge retskrivningsordbogen fra Dudenverlag (Duden 2013), hvis autoritet for tysks vedkommende er det tætteste vi kommer Sprognævnets, kan man frit vælge, om man ved præpositionelle syntagmer som *im Haus* og *im Jahr* tilføjer et dativ-*e* eller ikke (jf. Farø 2013). Men under *Jahr* i *Duden Universalwörterbuch* (Duden 2014) finder man, godt gemt, *im Jahr(e)* i form af et eksempel: ”im Jahre 1000 nach Christi [Geburt]”, dvs. uden den variation, samme forlags retskrivningsordbog sanktionerer. Hvordan skal vi som brugere forholde os til disse dels tilsyneladende liberale rammer, dels modstridende oplysninger?

Til alt held har vi bl.a. et lettilgængeligt tysk redskab som DWDS-korpus (forkortelsen står for *Digitales Wörterbuch der Deutschen Sprache*, som korpus er hovedkilden til). Det er, ligesom KorpusDK, balanceret mht. tekstgenrer (jf. Asmussen 2004), og det er, for kernekorpussets vedkommende, omtrent dobbelt så stort som det danske, målt i antal ord. Det skulle vel være muligt at foretage en søgning på disse syntagmer i korpus, og så finde ud af, hvad der er op og ned i de lidt modstridende oplysninger. Men det er faktisk ikke så lige til. For DWDS-Korpus dækker hele det 20. århundrede og tekstgrundlaget er indrettet tilsvarende. Dvs. når vi undersøger et fænomen vha. korpus, får vi hele periodens billede af det, inkl. den udvikling der har fundet sted. Og netop dativ-*e* er et fænomen med en markant diakroni (se nedenfor). Korpus er m.a.o. velegnet til at beskrive det 20. århundredes tyske sprog. Men det er ikke så velegnet som referencekorpus mhp. *aktuel produktion* på tysk. Det er det dilemma, store dokumentationsordbøger står over for², når de både vil beskrive sprogbrug empirisk over et længere diakrontidsspænd og samtidig vejlede i synkron sprogbrug. Det er ikke teknisk umuligt at integrere disse funktioner (altså dokumentation og hjælp til produktion) i samme opslagsværk eller -system, men det er vanskeligt, fordi det kræver et differentieret kildegrundlag. Et korpus, der skal tjene som reference for tekstproduktion, navnlig for fremmedsproglørnere, er nødt til at være langt mere fokuseret og kronologisk punktuel end et, der enten skal hjælpe med sprogbeskrivelse, tekstforståelse eller dokumentation.

Til alt held indeholder DWDS-pakken også et helt nutidigt avis-korpus, bl.a. delkorpuset *Die Zeit-Korpus*. Det er langt mere anvendeligt til vores

formål, også fordi et seriøst avissprog ikke er nogen dårlig stilistisk varietet at styre efter som (skriftlig) default-sprogform for en L2-lørner. Det såkaldte litteratursprog er en mere idiolektal affære med risiko for eksperimenter, der nok er interessante og æstetiske. Men de er ikke nødvendigvis den gyldne middelvej, en lørner gør sikrest i at følge, hvis han eller hun ikke vil vække unødigt opsigst, men blot *fungere* i den fremmede kode. Det er netop hele filosofien bag korpuslingvistik, at man sorterer det ekstreme fra og retter fokus mod det typiske, hyppige, almindelige, normale. Eller programmet gør det (delvist) for en.

Og hvad er så det typiske for disse præpositionelle fraser, der ifølge den officielle norm godt ”må” have *-e*? Lad os starte med *im Hause(e)*: Vi nøjes med at se på belæg fra 2014 for at få den helt dugfriske tendens i (*Zeit-*)tysk. Der er 111 træffere i alt. Heraf er kun 16 med den lange *-e*-variant. Vi kunne stoppe her og konstatere, at den korte variant er langt den mest udbredte, den udgør nemlig 86 % af alle realiseringerne, og at det gør udslaget for beskrivelsen af syntagmet eller vejledningen af lørnerne: ”Brug selv den korte, men vær samtidig forberedt på den lange, når du læser!”. Men så snyder vi os selv for en dybere indsigt. For hvad er det for kontekster, *dativ-e*-varianterne optræder i? Er der et distributionelt mønster? Ja, det er der.

Der kan af materialet udskilles:

1. Betydningen ’i den (betydningsfulde) familie’ med formalismen: *im Hause* [EFTERNAVN]
(belæg i korpus: *im Hause Slomka/Heidegger/Wagner/Wulff/Wahnfried/Falk*)
2. Betydningen ’i firmaet, organisationen o.l.’ med den simple formalisme *im Hause*:
(belæg i korpus: *im Hause des Wirtschaftsministers (= i ministeriet); Künstler stellen ihre Werke im Hause aus (= in der Herberge); der Stellvertreter ist nicht im Hause (= in der Firma)*)
Begge disse betydninger er idiomatiske, idet *Haus(e)* her ikke længere har sin konkrete betydning.
3. Resten af belæggene kan enten henføres til direkte citater fra ældre,

dvs. egentlig irrelevante tekster, der burde sorteres fra, eller til litterære eksempler, der også delvis aspirerer til en arkaiserende stil.

Så for dette specifikke syntagmes vedkommende har vi en klar tendens, hvor afvigerne ydermere er semantisk eller stilistisk forklarlige, så der ikke er nogen ren variation tilbage: Der er med andre ord *ikke* valgfrihed her. Valget afhænger helt og holdent af, hvad man ønsker at udtrykke.

Med det andet syntagme forholder det sig lidt anderledes: 38 ud af 1365 syntagmer fra 2014 er med dativ-*e*. Det er lidt under 3 %. Altså ikke noget at råbe højt om i korpuslingvistisk sammenhæng, og dermed heller ikke indlæringsmæssigt. Dvs. den førhen omtalte officielle norm, der i retskrivningsordbogen overlader valget til brugeren i stedet for at oplyse om sprogbrugen eller direkte vejlede i det sikre valg, er ikke i trit med sprogbrugen. Det er imidlertid mindre påfaldende end den nævnte oplysning i *Duden Universalwörterbuch*, der alene nævner den nu sjældne form i sit eksempel, så man som ubefæstet bruger må udlede, at dét er modellen. Og det er jo ikke forkert, jf. den officielle normbeskrivelse, men det er på den anden side heller ikke typisk.

Havde vi nu udnyttet DWDS-*Kernkorpus*, så havde vi fået en ganske anderledes ramme at arbejde ud fra: Sætter man programmets præsentationsmodus til ”Zufällig”, så får man belæg fra hele det 20. århundrede. Gennemløber man de 100 første resultater for ”im Jahr”, viser det sig, at kun 10 ud af 100 er *uden -e*, altså et mønster stik modsat det, vi før nåede frem til. Dvs. at syntagmet på et århundrede har skiftet norm fra langform til kortform. Blander man de to resultater, får man en blød mellemvej, men det er jo ikke hensigtsmæssigt. Det afgørende er 1. den sproglige udvikling, der har fundet sted, og 2. de semantiske brugsforskelle mellem de forskellige realiseringer af syntagmet. Valgfriheden er derfor uhensigtsmæssig. Selve det sproglige eksempel er naturligvis en detalje, men det er velegnet til at demonstrere muligheder og begrænsninger i de foreliggende korpusressourcer i forhold til opgaven.

Parallellkorpora

Vi har nu set på et par anvendelsesmuligheder af KorpusDK og DWDS-Korpus i tyskfacet. Her har de fungeret monolingvalt. Næste spørgsmål er så,

om de to korpora kan anvendes som parallelkorpora, dvs. kan de danne udgangspunkt for egentlige kontrastive undersøgelser? Som lakmusprøve kan vi forsøge at falsificere en relativt velkendt hypotese om dansk vs. tysk. Ifølge den er dansk kendetegnet ved en udstrakt brug af takkeformler, hvorved det adskiller sig fra tysk. Vi har at gøre med en særlig dansk pragmatisk markør, der i øvrigt sætter den traditionelle forestilling om tysk som et særligt ”høfligt” sprog i relief: *Sie*- vs. *du*-distinktionen er i hvert fald ikke nogen særlig god basis for denne opfattelse, for brugen af *Sie/du* reguleres snarere af situationelle faktorer som offentlig/privat, formel/uformel samt absolut/relativ alder hos de talende. Der er ikke noget specielt høfligt ved *Sie*, eftersom det simpelthen er udgangspunktet for offentlig kommunikation blandt voksne.³ Men der kan fx være noget distanceret i at fastholde *Sie* i nogle situationer, hvor der principielt er mulighed for *du*. Her er fastholderen ikke høfligere end *du*-forsøgeren. Set i det lys er det danske takkesystem faktisk langt høfligere, for her er der ikke tale om et valg mellem to former, der er temmelig stramt reguleret, men om at man kan vælge at takke for noget eller vælge helt at lade være.

Men hvad viser korpus os om den sag? Kan vi vha. de to korpora få bekræftet denne hypotese om dansk takke-høflighed vs. en manglende tysk ditto?

Vi søger på frasen *tak for* hhv. *danke für* og får flg. antal træffere:

tak (etc.)⁴ for: 3542

danke (etc.) für: 1006

Allerede forskellen i omfanget af belæg er påfaldende: Med en basis på hhv. 56 mio. (KorpusDK) og 100 mio. ord (DWDS-Kernkorpus) skulle man alt andet lige forvente en dobbelt så stor forekomst af takkefraser på tysk som på dansk. Men forholdet er lige omvendt: Det *danske* korpus har trods sin halve størrelse 3 ½ gang så mange takkefraser som det tyske. Sammenlagt kan vi altså tale om en faktor 7 for dansks vedkommende for takkefraser i forhold til tysk.

Men der er en forskel mere, denne gang kvalitativt: De danske belæg er langt hen ad vejen kendetegnet ved to forhold:

han ville være med til fødslen, og han sagde mig **tak for**, at jeg ville føde vores barn. Da jeg skulle hjem mig i armen: -De er mit vidne, Andersen. Det er **takken for**, at man har gået og pløjet og passet hende i en lille raslende lyd. Hun var lige nået til at **takke forsynet for**, at hun da i det mindste så ordentlig ud i Så kunne hun alligevel bedre klare, dem der sagde: " **Tak for** nul -Så vidste man, hvad man havde at rette sig efter hende: "Det kan du tro- " råbte hun tilbage. **Tak for** din hilsen, det var rart at høre fra dig. Ja, hvilt i os selv [...] - og værret helvedes lykkelige? **Tak for** opringen inat og alle ordene- jeg husker vist de fleste, forbrydelsen klitrer på alt. Der er intet at grine ad **tak dig selv for** at takke og blive til grin. Missilerne står klar og husk, kan du huske hvordan du kunne glemme dig selv. **Tak dig selv for** at tro, sekunderne preller af på din bøn og opfyldelsen på alt. Der er intet at græde over du kan **takke dig selv for** at takke og blive til gråd. Du kan tro på hvert fald til. Og du har kun dig selv at **takke for** din tro. Farven er grøn, transparent. Tonen elektrisk og horisontal. kan se lidt igen. Farmor beder endnu. Nu mumler hun: **Tak for** hvad? Sveden er væmmelig, men nakken gør lidt mindre ondt, Loewenthal en offentlig ydmygelse. Men denne " **tak for** sidst " er desværre posthum, og den til hævn og retfærdigheden fire år, og nu var det mor, som sagde nej **tak. Klogt af hende. For** nok var jeg lelfødet, men hvordan skulle jeg nogensinde fået et stipendium til et amerikansk universitet som **tak for** en forelæsning Palme havde holdt, blev han dybt såret. -De med tekster på spansk, engelsk, tyrkisk og græsk. " **Tak for** alt ". "Vi elskede dig. Den smukkeste og mest lavmælte hyldet cellistinde, havde en vidt fremskreden tuberkulose at **takke for** sin skælven og sine idelige besvimelser. For ikke at udsætte fjernt ophav disse spættede kvinder kunne have at **takke for** deres stakkels skæbner. Tværs henover hvor mange generationers barbar han. "Ved samme lejlighed kan jeg også få sagt **tak for** sidst ". "Ja, ja, men det blir uden mig ", svarede Ole. mand, råbte min farmor, der var lidt tungbar. -Vi kan **takke socialdemokratiet for** alt, og du sidder og snakker om en stærk mand. folks børn! Jeg mødte mange af dem senere, og jeg **takkede forsynet for**, at jeg ikke var blevet optaget der. Efter gymnasiet læste hendes mor havde forlangt det. Hun bøjede hovedet som **tak for** det, der ikke kunne være anderledes, og hun lovede sig hyrevogne og kareter bare for at trykke deres hænder og **takke for** storartede aftner på teatret. I Pedrins konditori trængtes gæsterne ved slig fri. Betjentene kunne deres håndværk. De kunne **takke frikorpsen for** den rutine, de havde fået i de seneste tre uger.

Fig. 2: Screenshot af udvalgt af søgningen tak for i KorpusDK.

l-Walter zur Linden der Universitätsbibliothek geschenkt hat , als **Dank für** die Verleihung der Ehrendoktorwürde .
 59. Dem einzigen Zuhörer schenkte der Ortsbürgermeister aus **Dank für** dessen Erscheinen eine Flasche Doornkaat .
 Den **Dank für** eine derart treudeutsche Militärbegeisterung der Sozialdemol
 Vielen **Dank für** den Einkauf .
 Vielen **Dank für** alles !
 » Äh ... vielen **Dank für** die Rettung !
 iz , dessen riesige Kirche nach dem Ersten Weltkrieg als später **Dank für** die Errettung aus der Theißflut errichtet wurde .
 pfund Kokain wieder aus meinen Postkasten entfernt hat , vielen **Dank für** ihr Verständnis und mit freundlichen Grüßen , Hinark Husen .
 ürbenden Grübeleien über die richtigen Geschenke , kein fauler **Dank für** die falschen , keine bekloppten Geburtstagsständchen und ke
 Titus in seiner Großmut nimmt ihr alle Ängste : Er **dankt für** ihre Offenheit und erklärt sich bereit , auf sie zu verzichten (A)
 Zunächst einmal sei Constanze Stelzenmüller **gedankt für** den die Hintergründe des Disputes um die neue Nato-Dok
 Herzlichen **Dank für** einen bewegenden Bericht .
 Rainer Degethoff , Grebenstein Dem Autor sei **Dank für** seinen engagierten Beitrag über die Unsäglichkeiten im Vorfe
 Horst Bohlen , Hamburg **Dank für** den deutlichen Leitartikel zu den Vorgängen um die Wehrmac
 Wir schulden unseren Vorfahren den tief empfundenen **Dank für** ihren großartigen Durchhaltewillen , für ihre Opfer , ihre Entbet

Fig. 3: Screenshot af udvalgt af søgningen dank(-) für i GWDS-Korpus.

- 1) De danske fraser er almensproglige, tenderende mod uformel sprogbrug;
- 2) de er i højere grad end på tysk standardiserede, både mht. inventar og variation (styrelsen er ofte fast, og udtrykkene er pragmatiserede). Jf. følgende selektive eksempler:

”Kære Poul Nyrup! **Tak for de pæne ord** om os i din nytårstale”
 ”Med det muntre Flintstone-tema kunne vi pænt sige **tak for denne gang**”
 ”**Tak for hjælpen**, kammerater!”

De tyske realiseringer går i en anden retning: 1) De tyske takkefraser tilhører modsat dansk i langt højere grad *formel* sprogbrug; 2) desuden er de fordelt på en meget stor syntaktisk variation, jf. fx:

”Dem Autor **sei Dank für** seinen eingagierten Beitrag”
 ”**Vielen Dank für** den Einkauf!”
 ”**Wir schulden** unseren Vorfahren **Dank**”
 ”**Meinen Dank für** die Karte”
 ”Herr Baumgart, **wir danken Ihnen für dieses Gespräch!**”

tak for-formler:	danke/(vielen) Dank für-formler:
Tak for alt/besøget/brevet/de pæne ord/denne gang/Deres brev af [DATO]/det/en dejlig aften/god ro og orden/mad/skænken/tilliden/tippet/turen	Herr/Frau [EFTERNAVN] wir danken Ihnen für dieses Gespräch! Vielen/Herzlichen/Tausend Dank für Deinen Brief vom [DATO]

Fig. 4: Eksempler på danske og tyske takkeformler.

Selv om det altså ikke er sådan, at man slet ikke finder takkeformler i tysk – ovenfor er der samlet et udvalg fra korpussøgningen – så spiller de både kvanti- og kvalitativt en anden rolle i tysk end i dansk. Der er færre af dem, de er mindre standardiserede, og de bruges snarere i formelle kontekster på tysk.

I den ordbog, som er udviklet hovedsageligt på grundlag af forgængerer for KorpusDK, *Den Danske Ordbog*, er artiklen *tak* set på denne baggrund efter min mening faldet for kortfattet ud: Der mangler simpelthen et bredt udvalg af takkeformler. Det er klart, at en stor del af ovenstående danske takkeformler er semantisk transparente. Men for det første er dét fænomen altid

relativt, for det andet er der også det pragmatiske element at tage hensyn til: Hvornår og hvordan bruges takkeformlerne hver især? Herudover er kollokationer, som det typisk drejer sig om her, også nyttige for en del brugere og situationer, transparente eller ej. Jeg er godt klar over, at tanken nok har været, at *tak for* + styrelse er så produktiv en forbindelse, at det kan forekomme næsten uoverkommeligt og også arbitrært at vælge nogle relevante kollokationer ud, eftersom det er et kontinuum fra fri dannelse til fast frase. Ovenstående er ikke desto mindre et bud på et sådant udvalg.

Hvad der mangler – og som er på vej

Tilbage til spørgsmålet om sammenlignelighed: Nej, KorpusDK og DWDS-Korpus i forening kan *ikke* anvendes som parallelkorpus, for de er ganske enkelt ikke parallelle nok. Det ville have krævet, at KorpusDK havde haft en lignende longitudinal dækning over hele det 20. århundrede, som DWDS har. Der mangler altså et egentligt parallelkorpus, der er bygget op af danske og tyske paralleltekster med det formål at gennemføre kontrastive undersøgelser. Med paralleltekster forstås i denne sammenhæng tekster, der har samme pragmatiske grundfunktion, og som dermed tilhører samme tekstgenre, fx *vittigheder, nytårstaler, reklamer, indlægssedler, opskrifter, kamppreferater, dødsannoncer, politiske tekster* etc. Udformningen af teksterne, såvel grammatisk som leksikalsk, vil helt sikkert være forskellig, men tertium comparationis er klart, også selv om genrerne kan variere mht. kulturel betydning, udbredelse, konnotationer m.m.

Fordelen ved sådan et korpus er, at man kan anvende det både til almene kontrastive undersøgelser på et omfangsmæssigt og tekstgenremæssigt balanceret grundlag og til tekstgenrespecifikke undersøgelser. Det er efterhånden veldokumenteret, at både grammatiske og leksikalske fænomener ikke er autonome størrelser, men også interagerer med parametre som tekstgenerer (jf. Farø 2009). Indsamlingen af sådanne paralleltekster og diverse pilotstudier gennem de senere år ved undertegnede har efterhånden resulteret i følgende materiale: Et dansk-tysk tekstkorpus⁵ p.t. bestående af i alt knap 500.000 ord for dansk og tilsvarende for tysk, fordelt på tekster fra ca. 30 forskellige tekstgenerer og emner, som hver især er repræsenteret af typisk 25 teksteksemplarer. Hvert sprogspecifikt delkorpus, fx *tyske boligannoncer*,

danske lederartikler, tyske artikler om det amerikanske valg, danske børnesange, har et omfang på gennemsnitligt ca. 10.000 ord. I enkelte tilfælde, som fx ved genren *anmeldelser*, er der eksperimenteret med forskellige former for subdifferentieringer af genren, eksempelvis *samme anmelder, samme værk, samme spil* vs. *forskellige anmeldere* etc., dvs. en ret minutios udskiftning af de forskellige parametre for at muliggøre undersøgelser af deres respektive betydning for sprogbrugen.

Et andet delkorpus, der er blevet særligt omfattende, er *politiske tekster* hhv. *tekster om politik*, hvad der pragmatisk er to meget forskellige fænomener, sådan at man må tale om hver sin grundlæggende *teksttype*: Den ene tilhører gruppen af direktive tekster, mens den anden hører til de assertive. Dette emnebestemte delkorpus indeholder politiske interviews, taler, kommentarer, portrætter, valgprogrammer og principprogrammer. For blot at antyde nogle af mulighederne i sådan et korpus vil jeg demonstrere nogle kontrastive pilotiagttagelser af kansler Merkels og statsminister Thorning's 3 nytårstaler fra 2012-14. Analysen baserer sig udelukkende på et frekvensbaseret udvalg af de hyppigst brugte politiske nøgleord i talerne (jf. Carius/Schröter 2013). Som datalingvistisk software er anvendt *Wordsmith*, vs. 6.

De summerede ordlister⁶ over de 24 hyppigste ord ser sådan ud:

	Thorning 2012-14	Merkel 2012-14
1.	vi	wir
2.	vores	unsere
3.	Danmark	Deutschland
4.	ældre	Zukunft
5.	børn	Familie
6.	fællesskab	Krise
7.	arbejde	gemeinsam
8.	uddannelse	Freiheit
9.	hinanden	Sicherheit
10.	generation	Gesellschaft
11.	krisen	Jungen

12.	unge	Kraft
13.	arbejdspladser	Mut
14.	ansvar	Zusammenhalt
15.	fælles	Polizisten
16.	virksomheder	Wohlstand
17.	offentlige	Afghanistan
18.	penge	Herausforderungen
19.	Afghanistan	Finanzen
20.	frihed	Mauer
21.	indsats	Schule
22.	forandring	Bildung
23.	hjælp	Gottes
24.	fremgang	Gesundheit

Fig. 5: Politiske nøgleord hos hhv. Thorning og Merkel.

Det første, man lægger mærke til, er de tre øverste pladser, der er besat med enten dansk-tyske ækvivalenter (1 og 2) eller en pendant (3). Indholdsmæssigt drejer det sig om det nationale fællesskab, både afsender og modtager er en del af, og som i begge tilfælde ikke overraskende er meget positivt konnoteret. En stor del af selve talens funktion handler jo om at bekræfte og vedligeholde dette fællesskab, og om muligt udnytte det til at appellere til modtagerne om at handle i en bestemt retning (Schröter/Carius 2013). Ser vi lidt længere nede på listen, så er der både ligheder og forskelle i prioriteringen af nøgleord. Først de fælles: *krise, frihed, fællesskab, uddannelse, Afghanistan*, om end *frihed* prioriteres lavere hos Thorning end hos Merkel. Der er dog også andre, endnu klarere forskelle, hvor fx Merkel alene inddrager *Gud*, mens Thorning bl.a. nævner det *offentlige* og *arbejdspladser*, som Merkel ikke tematiserer. Thorning får oftere talt om *børn, ældre* og *unge*, mens Merkel nøjes med at omtale *familien*, der slet ikke optræder i Thornings højfrekvente taleordforråd. Hos Merkel er der også plads til både *Polizisten* og *Sicherheit*, og også til appellen til den enkelte: *Kraft* og *Mut*. Ingen af disse

værdibegreber optræder hos Thorning i det udsnit, vi her fokuserer på. Derimod klumper tre begreber sig sammen til sidst i hendes liste, som peger i en bestemt retning: *forandring*, *hjælp* og *fremgang*, dvs. nøgleord, der antyder en aktivistisk-reformistisk politik. Det er også interessant, at Thorning taler om *penge*, mens Merkel taler om *Finanzen*.

Selv om der naturligvis er tale om en helt kontekstløs og overfladisk analyse, så kunne det tyde på et mønster, lighederne ufortalt. Samlet set danner der sig alene ud fra hyppigheden af ordvalget nogle tyngdepunkter, som med stor sandsynlighed kan tilskrives politisk-biografiske forskelle.

Korpus, normer og viden

Korpora er gennemgående ikke så stærke normsættere som fx ordbøger, grammatikker, lærebøger, lærere og native speakere. Det skyldes, at korpora grundlæggende er videnskabelige værktøjer, mens de øvrige nævnte kildetyper snarere er defineret som rettesnore. Rettesnore skal ikke fortolkes, men korpora er netop kendetegnet ved, at de *skal* udlægges. Når studerende således af og til kommer til at skrive, at ”der står i KorpusDK, at... (fx *laufen* er langt hyppigere anvendt end *rennen*)”, så er det en formulering, der vækker mistanke om, at meningen med hele øvelsen måske ikke helt er faldet på plads. Der *står* ikke noget i et korpus i den forstand, det her er ment, dvs. at der er fremsat et udsagn om sprog. Korpora *er* sprog, gjort let(tere) tilgængeligt. Man kan ikke spørge korpus om noget (selv om det tyske *abfragen* antyder det), men man kan vha. forskellige manipulationer af det udnytte dets samling af tekster på en sådan måde, at man får et output, der gør det muligt at opstille kvalificerede hypoteser om sproglige elementers adfærd i *real text*. En korpus*søgning* er videnskabsteoretisk noget ganske andet end et *opslag*: Et opslag søger viden og information, dvs. allerede bearbejdet og godkendt stof, mens korpus-*Abfragen* er det samme som dataanalyse, dvs. selvstændig og typisk undersøgende analyse af større data. En akademisk opgave mao.

Som det fremgår, er der et (velbegrundet) potentielt nært venskab mellem korpuslingvistik og fremmedsprogsstudier. Faktisk er venskabet mere oplagt her end mange andre steder. Men det er også afgørende, at metoder, kilder og spørgsmål harmonerer. Vi har langt flere metodiske muligheder end nogensinde før, og et næsten uendeligt datamateriale. Nu gælder det om at få det

til at spille sammen og udnytte mulighederne for helt nye kontrastive undersøgelsestyper med fokus på tekstgenrens og andre parametres betydning for sprogbrug og grammatik.

FORFATTER

Ken Farø, ph.d.
Lektor i tysk sprog
Institut for Engelsk, Germansk og Romansk
Københavns Universitet
kenfaroe@hum.ku.dk

NOTER

- 1 Der er i denne såvel som i andre discipliner altid nogle, der ikke tvivler på egne bedømmelsesevne. Det gør dem ikke nødvendigvis til bedre analytikere eller oversættere. Se Bergenholtz (1988).
- 2 Jf. også *Den Danske Ordbog*, om end problemet her er mindre, fordi kilderne i hovedsagen er fra 1983-92, jf. værkets videnskabelige indledning i bd. 1, s. 54.
- 3 Et af de bedste eksempler på *Sie* anvendt i en uhøflig kontekst er den i tysk politolingvistik ofte citerede udtalelse af den yngre Joschka Fischer i Forbundsdagen: "Mit Verlaub, Herr Präsident, Sie sind ein arschloch!" (jf. Klaus Geyer på SMIK-konferencen på SDU, 26.03.15).
- 4 I begge søgninger inkluderes verbale former som (da.) *jeg vil gerne takke for.../jeg takker dig for... ; (ty.) [dativ-obj.] sei Dank für [akk.-obj.]/jmd dankt für [akk.-obj.]/ich danke el. wir danken für [akk.-obj.]*. Disse udgør en større andel af materialet på tysk end på dansk.
- 5 Korpus er endnu upubliceret, men det forventes offentliggjort i løbet af 2015.
- 6 Forinden har jeg foretaget en frasortering af de første mange funktionsord fra den automatisk genererede frekvensliste; ikke fordi funktionsord som artikler, hjælpeverber og præpositioner pr. definition er politisk irrelevante (jf. Carius/Schröter 2013), men deres illustrative værdi i denne sammenhæng er begrænset.

LITTERATUR OG ANDRE KILDER

- Asmussen, J. (2004): Korpus 2000 – til hvilken nytte? Muligheder og grænser for empiriske sprogundersøgelser, i Duncker, D. (red.): *Studier i Nordisk 2002-2003*. København: Selskab for Nordisk Filologi, 17-19.
- Duden (2013): *Rechtschreibung*. Mannheim: Bibliographisches Institut.
- Duden (2014): *Universalwörterbuch*. Mannheim (app-version).

- DWDS-Kern-Korpus: <http://www.dwds.de/ressourcen/kernkorpus/>
- Farø, Ken 2009: Zur Grammatik der Instruktionen. I: Lenk/Lüger (red.): *Kontrastive Medienlinguistik*. Landau: Landauer Schriften zur Hochschulpädagogik.
- Farø, Ken 2013: Dativ-e revisited. I: *Ny dansk forskning i grammatik*. Odense: Odense Universitetsforlag.
- KorpusDK: <http://ordnet.dk/korpusdk>
- Popper, Karl (1959/2005): *The Logic of Scientific Discovery*. London: Routledge.
- Schröter, Melani/Carius, Björn (2009): *Vom politischen Gebrauch der Sprache. Wort, Text, Diskurs. Eine Einführung* (Leipziger Skripten 5). Frankfurt am Main: Peter Lang.
- Zeit-Korpus: <http://www.dwds.de/ressourcen/zeitkorpus/>

Redaktionel note

Følgende bidrag er fagfællebedømt: Margrethe Heidemann Andersen, Nina Møller Andersen, Christian Becker-Christensen & Jørgen Schack, Lotte Weilgaard Christensen, Niels Davidsen-Nielsen, Philip Diderichsen, Anna Sofie Hartling, Anne Kjærgaard & Anna Kristiansen, Ken Farø, Frans Gregersen, Hans Götzsche, Peter Juel Henriksen & Auður Hauksdóttir, Henrik Galberg Jacobsen, Per Anker Jensen, Anne Kjærgaard, Bodil Nistrup Madsen & Hanne Erdman Thomsen, Bolette S. Pedersen, Sanni Nimb & Sussi Olsen, Hanne Ruus, Lene Schøsler, Jacob Thøgersen, Ole Togeby.