

Statistik og journalistik

 Af [Hans Degn](#) 14. januar 2010 • I kategorien [Videnskabsteori](#) •

Den 4. januar 2010 vågnede jeg op om morgenen og var sprogforsker. Det kunne jeg læse i aviserne, at jeg var, og så måtte det jo det være sandt, eller hvad?

Det hele begyndte en dag i begyndelsen af december. Jeg sad ved computeren og forsøgte at gøre min fremtidsroman om Medea Olsen færdig, og som så ofte kunne jeg ikke rigtig finde på noget at skrive. Tankerne vandrede.

De foregående dage havde jeg lagt mærke til, at forskellige mennesker, der blev udspurgt om det ene eller det andet i fjernsynet, alle sagde, at de var *optaget af* noget. Hvis jeg anstrænger mig, kan jeg huske helt tilbage til tiden omkring afslutningen af den spanske borgerkrig, og jeg mener bestemt ikke, at man sagde "optaget af" i hver anden sætning dengang. Jeg mindes ikke, at jeg har hørt Stauning sige det.

Jeg forsøgte så at bruge Google til at finde ud af, hvornår folk var begyndt at sige "optaget af" i tide og utide, men jeg opdagede, at Google har en meget kortere hukommelse, end jeg selv har. Google er ikke velegnet til at følge ændringer i sprogbrugen over længere tid, fordi hovedparten af de prøver på almindeligt dagligdags sprog, man kan finde, højst er nogle få år gamle. Der findes selvfølgelig ældre tekster, men de er som regel i et for højt stilleje til at kunne bruges til undersøgelser af almindelig sprogbrug.

Så faldt det mig ind, at Google måtte egne sig udmærket til at sammenligne sprogbrugen i nutidige tekster. Det kunne jeg hurtigt bekræfte, og så gik jeg i gang med at undersøge hvor hyppigt udtrykket "optaget af" forekommer i forskellige tekster. Mit slutopgør med Medea Olsen blev udsat, og jeg producerede i stedet indlægget "[Klicheer på alle hylder](#)", som ved en tilfældighed havnede i Sproguseet.

En naturvidenskabelig bagtanke

Det har været en fornøjelse at læse de mange [kommentarer](#) mit indlæg har udløst i Sproguseets hjemmeside og mange andre steder. Særligt har det mere mig, at forskellige kommentatorer har ment, at jeg havde en skjult dagsorden, og den skulle være at nedgøre Socialdemokraterne eller glorificere Dansk Folkeparti. De har ramt helt forkert.

Når sandheden skal frem, så havde jeg en bagtanke med min undersøgelse, og det var at demonstrere, at sproget kan betragtes som et biologisk fænomen og undersøges med en naturvidenskabelig metode. Mit fag er biokemi. Jeg har ikke det ringeste begreb om sprogforskning, men jeg ved, hvordan man griber en naturvidenskabelig undersøgelse an.

Naturvidenskabelig forskning beskæftiger sig med det, der kan tælles, måles og vejes. Man kan for eksempel tælle, hvor mange rødspætter, der er højrevendte og hvor mange, der er venstrevendte. Hvis en fangst er på 735 rødspætter og tre af dem viser sig at være venstrevendte så udgør de venstrevendte 0,4% af fangsten og sikkert også af den bestand som befinder sig på havets bund. Før man offentliggør et sådant resultat overvejer man, om der kan være fejlkilder i undersøgelsen, der bevirker at resultatet kan være forkert.

Rødspætten er en usymmetrisk fisk, så det er umuligt at tage fejl af, om den er højrevendt eller venstrevendt, det beror ikke på et skøn. Hvis man tæller fiskene et par gange, kan man også være sikker på, at der ikke sniger sig nogen fejl ind der. Lidt større problemer kan det volde, at afgøre om en fisk faktisk er en rødspætte eller, om den er en krydsning med en beslægtet art, men her vil en trænet biolog ikke tage fejl. Endelig kunne man forestille sig, at de højrevendte rødspætter lettere lader sig fange end de venstrevendte, og fangsten derfor ikke er repræsentativ for bestanden, men det er vist for langt ude. Når man har overvejet de mulige fejlkilder, kan man med god samvittighed fremlægge sit resultat ledsaget af en beskrivelse af metoden, man har brugt og en diskussion af de mulige fejlkilder.

Nu vil de læsere, der habituel leder efter skjulte dagsordner, allerede være overbevist om, at den pædagogiske historie om optælling af højre- og venstrevendte rødspætter i virkeligheden er ment som en fabel, der skal sige noget om højre- og venstreorienterede mennesker, men de tager fejl igen. Historien skal blot lede læseren ind på den

Der er forskel på hvor flittige politikerne er til at bruge klicheer. Se

© DR

Skrevet af: [Randi Isager](#)

Sprogforsker: Socialdemokrater bruger flest klicheer

04. jan. 2010 10.22 Politik Opdat.: 04. jan. 2010 11.36

Når politikerne åbner munden, kommer der tit klicheer ud. Og det sker oftest, når det er en politiker fra Socialdemokratiet, der taler.

Det skriver Morgenavisen Jyllands-Posten på grund af en undersøgelse som sprogforsker og mag.scient. Hans Degn, har lavet og lagt ud på hjemmesiden sprogmuseet.dk.

Udsnit af Randi Isagers artikel på dr.dk, 4. jan

Ny undersøgelse sætter Socialdemokraterne grundigt til vægs: Næsten samtlige udtryk på partiets hjemmeside er sproglige klicheer, men hvilket parti bruger mindst?

 Af [Søren Schauser](#)
 Mandag den 4. januar 2010, 14:32

Ak, ja... Man kan måle på meget. Senest har sprogforskeren Hans Degn brugt nogle decemberdage på en usædvanlig undersøgelse: Han tjekkede de politiske partiers sprogbrug for klicheer – defineret som en sproglig vending, der er blevet trivial og pinlig ved gentagen brug. Var alle partier lige slemme? Eller ville der være mærkbare forskelle?

Udsnit af Søren Schausers artikel på berlingske.dk, 4. jan

Seneste sprognyheder

19/3	Evolution i sprog: Dovenskab har Åndret ord videnskab.dk
30/10	Hvem taler: Kan du hÅre forskel på menneske og maskine? www.dr.dk
6/3	DansklÅrer: Tosprogede bÅrn bliver sprogligt forsÅmt www.bt.dk
28/2	Flere og flere ordblind Åsker pÅ universitetet - Magisterbladet magisterbladet.dk
22/2	Keeper eller mÅlmand Å? hvilket ord er bedst? Dansk SprognÅvn dsn.dk
28/3	Indigenous languages come from just one common ancestor, researchers say - ABC News (Australian Broadcasting Corporation) www.abc.net.au
27/3	UVic news - University of Victoria www.uvic.ca
26/3	New York Nonprofit Media: Native American nonprofit aims to boost community health by teaching Lakota language nynmedia.com
22/3	Teenage business owner aims to revitalize Blackfoot language globalnews.ca
19/3	Language revival preserving history www.smh.com.au

FLERE NYHEDER >>>

Verdens sprog på Sproguseet.dk på et større kort

Nye kommentarer

Artikler om

Arkiv

tankegang, at der ikke er nogen principiel forskel på at tælle rødspætter og tælle ord. Ord kan lige så vel være objekt for en naturvidenskabelig undersøgelse, som rødspætter kan, og præsentationen af resultatet af undersøgelsen skal også her være ledsaget af en metodebeskrivelse og en diskussion af mulige fejlkilder. Metodebeskrivelsen skal gøre det muligt for andre at eftergøre undersøgelsen og kontrollere resultatet.

En offentliggjort undersøgelse regnes i naturvidenskaben først for helt troværdig, når den er blevet eftergjort af flere andre. Derfor appellerede jeg i mit indlæg flere gange til læserne om selv at prøve metoden, som jeg havde beskrevet i detaljer. Det er jo ikke så svært.

Blandt kommentatorerne er der to, der fortæller at de selv har afprøvet metoden på andre hjemmesider, end dem jeg havde valgt, men ingen fortæller, at de har tjekket mine optællinger. Derimod er der flere, der synes at være overbevist om, at hvis undersøgelsen blev gjort på den efter deres mening rigtige måde, så ville den give et helt andet resultat. Til dem sige jeg, som man gør det i naturvidenskaben: Det kan vel være, men bevis det.

S sprøjter klichéer ud
 Af MATHIAS STIGSGAARD
 Offentliggjort 04.01.10 kl. 08:59

"En bred vifte", "laveste fællesnævner" og "tvingende nødvendigt". Socialdemokratiet bruger flest sproglige klichéer, viser en undersøgelse.

Klichéerne:
 på sigt, en bred vifte, vi som politikere, konkrete tiltag.

Når politikerne åbner munden, kommer der tit klichéer ud. Og når en politiker fra Socialdemokratiet taler, sker det endnu hyppigere.

Udsnit af Mathias Stigsgaards artikel på jp.dk, 4. jan

Fejlkilder

Der især min klichéliste, der er blevet kritiseret. Den kritik havde jeg forudset. Jeg skrev udtrykkeligt i mit indlæg, at min klichéliste var et subjektivt input til undersøgelsen og, at enhver var velkommen til selv at prøve med en klichéliste efter eget valg, men ingen fortæller, at de har gjort det. Kritikerne har valgt den nemme vej, at nøjes med at forudsige resultaterne af sådanne undersøgelser. Hvis de prøvede at udføre undersøgelserne på alternative klichélistes, ville de muligvis opleve, at virkeligheden kan være genstridig overfor forudfattede konklusioner.

Jeg har også i mit indlæg antydnet, at jeg ikke aner, hvordan Google virker. Jeg kan derfor have overset metodiske fejlkilder. De, som har indsigt i Googles virkemåde og særheder, kan muligvis forbedre undersøgelsen, og jeg vil indtrængende opfordre dem til at gøre det.

Når jeg så energisk forsøger at lokke andre til at gentage og forbedre min undersøgelse, så er det fordi, den drejer sig om noget vigtigt. De politiske partier har brug for et objektivt portræt af deres sprog. Det kan give anledning til gavnlige korrektioner. Mit indlæg ville nok ikke have vakt postyr i dagspressen, hvis emnet havde været helt uinteressant.

De politiske partiers villighed til at få ryddet op i deres klichéer, når de får dem stukket op under næsen, kan naturligvis ikke betragtes som givet. Således spørger kulturordfører Mogens Jensen (S) BT's journalist, hvordan man kan sige "samfundets svageste" eller "ufinansieret" på en anden måde. Dog mente han, at man nok skal underkaste sproget på sin hjemmeside et servicetjek.

"servicetjek" – Ak ja. Spildte Guds ord på Balle-Lars.

Socialdemokraterne er det danske politiske parti, der bruger flest klichéer, viser ny sprogforskning. Indsat i taleboblen: eksempler på de mest brugte klichéer ifølge undersøgelsen.

Se hvordan partierne er rangeret

S er den absolut største klichéfabrik

SØREN MORTENSEN - Tirsdag den 5. januar 2010, 21:07

Sprogforsker: S er 93 pct. kliché, DF kun 20 pct. Siden Dansk Folkepartis stiftelse, har partiet har været dygtigt til at stjæle vælgere fra Socialdemokraterne. Og det kan der være en ganske simpel grund til, viser ny sprogforskning.

50%

Udsnit af Søren Mortensens artikel på bt.dk, 5. jan

Procentregning

Det er en almindelig erfaring, at når man læser i aviserne om noget, som man selv har et godt kendskab til, så opdager man, hvor overfladiske og upålidelige avishistorier kan være. Den erfaring har jeg gjort endnu en gang.

I samtlige tilfælde, hvor mit indlæg har været journalistisk behandlet, er jeg blevet kaldt sprogforsker. Den opmærksomme journalist burde måske have bemærket, at jeg har angivet min titel som mag. scient. og derfor have undersøgt, om det kan passe, at en person med en naturvidenskabelig uddannelse er sprogforsker. Det er jeg som allerede nævnt ikke, medmindre et enkelt indlæg i en blog har gjort mig fortjent til denne titel.

Meget værre er journalisternes omgang med procentregning. Hvordan kan nogen finde på at skrive noget så åbenlyst urimeligt som: "S er 93 pct. kliché" på grundlag af en undersøgelse, hvor det flere gange er fremhævet, at de fremlagte tal er relative?

Hvis jeg havde gjort min liste over almindelige udtryk længere ville summerne af forekomster af almindelige udtryk være blevet større, og alle tallene for klichéer pr tekstenhed ville være blevet tilsvarende mindre, men forholdet imellem dem ville have været uændret.

Hvis hvert partis tal for klichéer pr

BETA 180GRADER

SPIL

FORSIDE SAMFUND UNDERHOLDNING LIVSSTIL SKRIV AF

Erfverv og finans | IT | Krimi | Økonomi | Politik | Udland | Videnskab |

Socialdemokraternes hjemmeside er 93% klichéer

Udgivet for 10 dage siden i Politik
 Tilføjet 02/01 07:20 til Boblere af Hans Henrik Juhl
 Kilde: sprogmuseet.dk

Mag. scient. Hans Degn har ovre på Sprogseets blog stillet sig selv spørgsmålet: "Er alle politiske partier lige slomme til at bruge klichéer?".

Han søger svaret i korpuslingvistikens metoder og kommer blandt andet frem til, at Socialdemokraterne er de flittigste til at bruge klichéer på deres hjemmeside med hele 93%, mens Dansk Folkeparti bruger færrest med kun 20%.

93%

tekstenhed var blevet ti gange mindre, ville tallet for S have været 0.093 og for DF 0.02, men forholdet mellem dem ville stadig have været ca. 5.

Hvis journalisten så havde brugt den samme omsætningsmetode fra relative tal til procenter ville han have skrevet "S er 9,3 pct. kliché", hvis han overhovedet havde skrevet noget.

Det er ikke kun én journalist, der er skyldig. De bruger alle mine relative tal, som om de var procenter.

Nu sidst har ekstern lektor i sprogs psykologi Charlotte Marie Bisgaard Nielsen skrevet et [indlæg](#). Det begynder med en henvisning til mit indlæg og slutter med at udtrykke skepsis overfor "lingvisternes brug af skemaer og computerprogrammer". "Inden for mit studieområde taler vi fortrinsvis med mennesker om deres sprog" slutter hun.

Det, der står mellem indledningen og slutningen, er sort tale for mig. Jeg er ikke sikker på, at jeg har forstået, hvad det handler om, men jeg har da i det mindste lært et nyt ord: "sprogsyn". Det ord kan sikkert ligesom "menneskesyn" bruges til at fordømme andre med.

Jeg troede nu, at jeg havde skrevet slutreplikken, men så fik jeg den tanke, at der kunne være opfundet andre sammensatte ord, der ender på -syn og har fordømmelsespotentialer. Der var gevinst i første søgning på Google: "[klimasyn](#)".

*Hans Degn, mag. scient.
biokemiker*

Læs også:

1. [Ord med fordømmelsespotentialer](#) I et tidligere indlæg her i Sproguseet har jeg antydnet, at ordet sprogsyn ligesom ordet menneskesyn har fordømmelsespotentialer. Betegnelsen opfandt jeg til lejligheden og den er vist selvforklarende. Jeg overvejede...
2. [At tage hånd om sproget](#) At tage hånd om sproget var det glade julebudskab fra MF, kulturordfører for Socialdemokraterne Mogens Jensen, i hans elektroniske nyhedsbrev Kulturnyt den 22. december 2009. I hvilken anledning er Mogens...
3. [Klichéer på alle hylde](#) I kvantitativ lingvistik analyserer man tekster ved at optælle, hvor hyppigt udvalgte ord og vendinger forekommer. Af de indsamlede tal forsøger man så at drage konklusioner ved brug af mere...
4. [StandUpResearch](#) Tema: (Næsten) alt om sprog – Kommer Peter til festen i morgen? – Det er ikke godt at vide. Hvorfor er det ikke godt at vide når afsender formentlig synes...

Tagget med: aviserne, Berlingske Tidende, biokemi, biologisk fænomen, BT, computerprogram, dagligdags sprog, dagspressen, Dansk Folkeparti, diskussion, DR, fejkilder, fjernsynet, fordømme, Formidling, forskning, Google, hukommelse, højreorienteret, input, journalistik, Jyllands-Posten, kliché, klimasyn, korpus, korrektioner, kulturordfører, lingvist, maskinoversættelse, menneskesyn, Mogens Jensen, naturvidenskabelig metode, naturvidenskabelig undersøgelse, objektiv, politiske partier, procentregning, sandsynlighed, skema, Socialdemokraterne, sprogforsker, sprogs psykologi, sprogsyn, statistik, tekst

4 kommentarer



Hans Henrik Juhl

14. januar 2010 • 19:10

Jeg skal gerne indrømme, at den overskrift jeg valgte da jeg delte dit indlæg på 180grader var intet mindre end overmåde sensationel – Det var skam også med fuldt overlæg og rettet mod en bestemt målgruppe. Jeg havde aldrig regnet med at det ville sprede sig til andre medier, hvor journalisterne gjorde med teksten som det passede dem.

Jeg er nu heller ikke journalist eller noget i den dur, men blot en simpel student. Og som sagt valgte jeg at udlægge tallene på den mest sensationelle måde, velvidende, at det var urimeligt overfor Liste A. Urimegheden til trods, ændrer det dog ikke så meget på, hvordan fordelingen af klichéer er de forskellige partier imellem.

Jeg vil dog påpege, at det af min tekst fremgår, at du er naturvidenskaber, som benytter sig af en metode, som er kendt fra korpuslingvistikken (som uanset, hvad Fr. Bisgaard Nielsen måtte mene, er en fremragende metode til at undersøge sprogbrug).

Svar

Ruben Schachtenhaufen

20. januar 2010 • 11:40



Jeg synes du er en udmærket sprogforsker af en biokemiker at være. Tænk hvor let man kan skifte kurs hvis man bare har de statistiske værktøjer i orden.

Jeg fik i hvert fald mere ud af at læse dine indlæg end Marie Bisgaards, som, når nu vi snakker om titler, er ekstern lektor i sprogpsykologi. *Det er mig uvist hvad det egentlig er sprogpsykologer tror de undersøger ved at tale med mennesker om deres sprog.*

Svar



Ole Stig Andersen

23. januar 2010 • 06:20

@ Hans Henrik

Når du vidste at din opsummering af Hans Degns artikel var løsagtig, fejlagtig, vildledende, hvorfor gjorde du det så?

Er det en arbejdsform du vil anbefale andre? Én du ofte selv bruger? Bør man læse alt hvad du eller 180 grader skriver, med ekstra agtpågivenhed?

Svar



Hans Henrik Juhl

19. februar 2010 • 06:08

Man bør altid tage indhold som man finder på et brugerdrevent site med et gran salt, for man kan sagtens finde indhold, som er lagt op af andre årsager end blot at informere den generelle befolkning. For mit vedkommende var det et spørgsmål om at finde ud af om jeg kunne spinne artiklen til at få en høj ranking. Dét kunne jeg. Det var intet mindre end et eksperiment udi sociale medier.

Svar

Skriv en kommentar

Navn (kræves)

E-mail (kræves)

Hjemmeside

Send mig en e-mail når der kommer flere kommentarer.

