The image features two large, bold, black letters, 'A' and 'B', set against a white background. The letter 'A' is on the left and the letter 'B' is on the right. Inside the upper triangular part of the 'A', the text 'Er danske bogstaver' is written in a white, sans-serif font. Inside the upper loop of the 'B', the text 'til at stole på?' is written in the same white, sans-serif font. The letters are thick and have a slightly irregular, hand-drawn appearance. There are also some faint, curved black lines scattered around the letters, possibly representing motion or design elements.

Er danske bogstaver

til at stole på?

## Hvor meget betyder afstanden mellem skrift og tale, når børn skal lære skriftens kode? Og hvor stor er afstanden egentlig?

Af Holger Juul

Vi kan næsten ikke få bogstaver nok. Man ser det i navngivningen af nye fænomener og firmaer. FCK, USB, BMI, MGP, PMS, FKK. SMC-biler. BJ Autodele. KB Motorservice. Og når firmaerne får udformet logoer, er bogstaverne næsten altid det grafiske grundelement. Det ligner en regulær forelskelse i alfabetet.

Det er ikke så mærkeligt. Det er fascinerende og fantastisk at vores skriftlige kommunikation i det store og hele består af kombinationer af blot en god snes minimale symboler. Når man tænker på hvad bogstaverne gør for os, er det ganske forståeligt at vi er glade for dem.

Selv om 'bogstavord' som FCK og USB er populære, så er bogstavernes funktion som byggeklodser i nye navne og betegnelser dog kun en slags bijobber. Bogstavernes hovedfunktion er at repræsentere de ord der allerede findes i talesproget.

Hvilke kombinationer der bruges til at repræsentere hvilke ord, er bestemt af det alfabetiske princip. Ideen er som bekendt at hver enkelt lyd i et ords udtale repræsenteres af et bestemt bogstav. Det er enkelt og smart.

Men når det kommer til stykket, er det alligevel ikke helt så enkelt. Dansk ortografi er mere kompliceret end som så ...

**Komplekse og inkonsistente bogstav-lyd-relationer** Det er let at finde ord hvor forholdet mellem bogstaver og lyde afviger fra det alfabetiske princip. For det første er bogstav-lyd-relationerne ikke altid relationer mellem et enkelt bogstav og en enkelt lyd. Ordet *farlig* har eksempelvis fire lyde, men seks bogstaver. Bogstavparrene *ar* og *ig* i *farlig* er nemlig såkaldt komplekse grafemer, der hver kun repræsenterer en enkelt lyd. Omvendt repræsenterer *x*'et i et ord som *taxa* to lyde, og ikke blot en enkelt. Forekomsten af komplekse bogstav-lyd-relationer betyder at man ikke kan være sikker på at hvert enkelt bogstav i et ord svarer til en enkelt lyd.

For det andet kan et bogstav ofte svare til forskellige lyde i forskellige ord. Bogstavet *g* har for eksempel fire forskellige lydverdier i ordene *genial*, *geni*, *krage* og *mage*. På samme måde kan en lyd være repræsenteret af forskellige bogstaver i forskellige ord. En åben *a*-lyd repræsenteres normalt af bogstavet *a* som i *haj* og *ham*. Men der er også ord hvor denne lyd repræsenteres af *ar* (som i *farlig* og *var*) eller af *e* (fx *jeg* og *hej*) eller af *i* (fx *mig*). Bogstav-lyd-relationerne kan med andre ord være inkonsistente. Inkonsistens i læseretningen betyder at man ikke kan være sikker på hvordan et ord skal udtales, og inkonsistens i staveretningen betyder at man ikke kan være sikker på hvordan et ord skal staves.

Det er oplagt at begge de nævnte afvigelser – kompleksitet og inkonsistens – kan have negative konsekvenser for brugerne af skriftsproget, fordi relationen mellem skrift og tale bliver sværere at gennemskue. Det gælder naturligvis ikke mindst for børn der skal lære at beherske skriftens kode.

**Virker det alfabetiske princip overhovedet?** Det er et ofte diskuteret spørgsmål hvad man skal stille op over for afvigelser som de nævnte. En teoretisk mulighed er at vi begynder at udtale ordene som de staves. Vi kunne for eksempel begynde at udtale ordet *jeg* med *e*-lyd og *g*-lyd, sådan at det kom til at rime på *slik*. Ordet *slik* ville dog så skulle udtales med en *i*-lyd som i *gik*, sådan at det alligevel ikke ville rime på *jeg*. Det er overordentlig usandsynligt at danske sprogbrugere skulle kunne enes om at ændre deres udtalevaner. Alligevel er der af og til læserbrevsskribenter der opfordrer dem til at gøre det.

En anden mulighed er at ændre på retskrivningen. Det stilles der også regelmæssigt forslag om. I de senere år har blandt andet professor Frans Gregersen fra Køben-

havns Universitet slået til lyd for en ortografisk reform. Med samme regelmæssighed afvises reformforslagene med henvisning til at stabiliteten i skriftsprognormen bør veje tungere end hensynet til udtaleoverensstemmelse. Blandt andet formanden for Dansk Sprognævn, Niels Davidsen-Nielsen, har lagt sig på denne linje.

En tredje mulighed er at acceptere afvigelserne – og så tage konsekvenserne af dem i undervisningen i læsning og stavning. Men hvilke konsekvenser?

Læsepædagogisk er der lærere der har ment at undervisningen af begynderlæsere slet ikke skal fokusere så meget på det alfabetiske princip. Princippet er jo alligevel ikke helt til at stole på, og det interessante ved ord er i sidste ende slet ikke ordenes lyde og bogstaver, men derimod deres betydning. Begge dele er rigtigt nok. Men det er en forhatet konklusion at det alfabetiske princip slet ikke virker. Mange undersøgelser viser nemlig at selv erfarne læsere støtter sig til forbindelserne

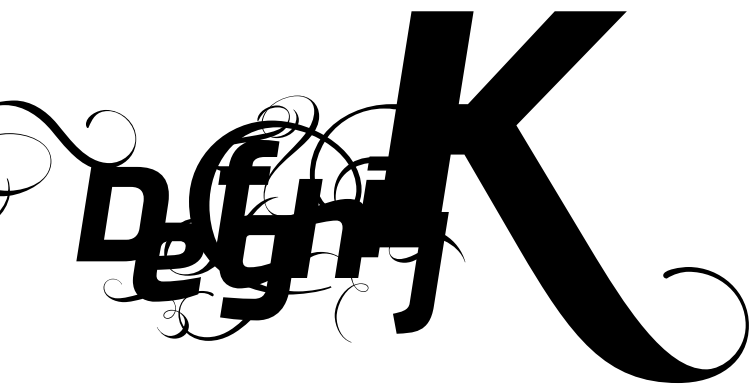
mellem bogstaver og lyde når de læser.

Der er da også mange lærere der har draget mere moderate konsekvenser. Eksempelvis er det blevet almindeligt at kategorisere de ord der bruges i begynderundervisningen efter deres ortografiske egenskaber. Man taler blandt andet om 'lydrette' ord, som er ord hvor alle bogstaverne har deres mest almindelige udtale, og som derfor kan være

det spørgsmål, kan det være en god ide at skelne mellem flere grader af regelmæssighed. Traditionelt har man opereret med enten-eller-kategorier som 'regelmæssige vs. uregelmæssige' ord og 'lydrette vs. ikke-lydrette' ord. Dermed ser man stort på at uregelmæssigheden i et uregelmæssigt ord normalt kun gælder en del af ordet – måske kun en enkelt bogstav-lyd-relation som fx *b'et* der

ner end man finder i dansk og engelsk, sådan at det er lettere for finske børn at 'knække koden'.

Der er udbredt enighed om at finsk er en meget regelmæssig ortografi, mens svensk og tysk er forholdsvis regelmæssige, hvori- mod engelsk er en meget uregelmæssig ortografi. Dansk karakteriseres gerne som en uregelmæssig ortografi, men dog ikke helt så



mere begyndervenlige end andre ord.

Hvad der imidlertid forbliver uklart i diskussionerne, er hvor stort et problem der egentlig er tale om. Hvor meget betyder afstanden mellem skrift og tale for tilegnelsen af læsning og stavning i praksis? Og hvor stor er afstanden egentlig?

**Regularitetseffekten** Det er veldokumenteret at ortografiske uregelmæssigheder faktisk har en indflydelse på sprogbrugerens læsning og stavning. 'Uregelmæssige' ord rammes oftere end 'regelmæssige' ord af stavefejl, og man har vist at selv erfarne læsere læser dem langsommere end regelmæssige ord. At man kan iagttage en effekt af ortografisk regelmæssighed, den såkaldte 'regularitetseffekt', betyder dog ikke nødvendigvis at afvigelser fra det alfabetiske princip udgør noget alvorligt problem i praksis. Selv om 'uregelmæssige' ord som *seksten* og *peber* er relativt svære, så er det jo langt fra umuligt at lære dem. Sprogbrugerne ville nok få det lettere hvis man kunne enes om at gå over til mere lydrette stavemåder som *sajsten* og *pever*. Men hvor meget lettere?

Hvis man skal sige noget mere præcist om

udtales som en *w*-lyd i *peber*. Ligeledes ser man stort på at ikke alle uregelmæssigheder er lige uregelmæssige. Både *b*-udtalen i *peber* og *k*-udtalen i *seksten* er klare afvigelser fra bogstavernes normaludtale – men det er ikke sikkert at afvigelserne er lige store, hvis man gør et mere detaljeret regnskab op. Det er derfor heller ikke indlysende at ordene er lige svære i praksis, selv om begge hører til i kassen med uregelmæssige ord.

Hvis man vil se nærmere på forholdet mellem ortografisk regelmæssighed og ords sværhedsgrad i praksis, må man desuden have med i billedet at også andre faktorer kan have betydning – eksempelvis ordenes længde og hyppighed.

**En tværsproglig regularitetseffekt?** Også tværsprogligt har man iagttaget noget der kan ligne en regularitetseffekt. Man kan nemlig ikke bare karakterisere ord som mere eller mindre regelmæssige, men også hele ortografier. Flere undersøgelser har vist at finske børn kommer hurtigere i gang med at læse end børn i Danmark og Storbritannien. Og en forklaring på denne forskel kan være at finsk har færre uregelmæssige bogstav-lyd-relatio-

uregelmæssig som engelsk. Der er imidlertid her tale om meget overordnede sammenligninger af ortografierne. Hidtil har det således stået hen i det uvisse om engelsk faktisk er en mere uregelmæssig ortografi end dansk hvis man sammenligner mere detaljeret og bruger ensartede målemetoder.

Der er naturligvis også andre forhold end ortografisk regelmæssighed der kan have betydning for hvor hurtigt børn kommer i gang med at læse i forskellige lande. Tværsproglige sammenligninger kompliceres eksempelvis af at der også er væsentlige forskelle mellem lande når det gælder uddannelsessystemets indretning (fx skolestartalderen og lærernes uddannelsesniveau).

**En kvantitativ beskrivelse af dansk ortografi** Hvis man vil mere i dybden med regularitetseffekten, vil det som nævnt være en fordel at etablere nogle mere fintmærkende mål på ortografisk regelmæssighed.

Når det gælder dansk, foreligger der udmærkede kvalitative beskrivelser af relationen mellem bogstaver og lyde. At beskrivelserne er kvalitative vil sige at de ikke sætter tal på hyppigheden af komplekse grafemer

og på graden af konsistens. Mulighederne for at sætte tal på, og dermed komme et skridt videre, er imidlertid forbedret markant i de senere år med computerens indtog i sprogvidenskaben. Det satte jeg mig for at udnytte.

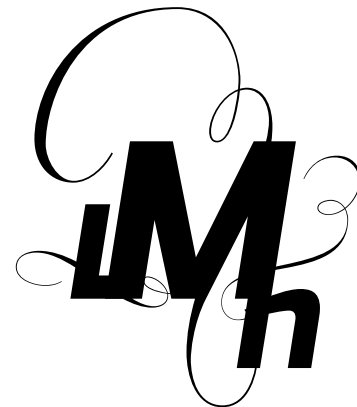
Redaktionen af Den Danske Ordbog (2003-2005) var så venlig at give mig adgang til en liste med samtlige opslagsord i værket i ortografisk form og for de fleste ords

I mange tilfælde var der flere muligheder når ordene skulle opdeles i grafemer. I *købmand* kan man eksempelvis mene at *b*'et lige så godt kan høre sammen med det følgende *m* som med det foregående *ø*. Når løsningen med *øb* blev foretrukket, skyldtes det blandt andet at sammensætningsdelene *køb* og *mand* dermed blev holdt adskilt.

I visse ord giver en konsekvent opdeling i

blandt andet i *ch* og *ck* (som i *check*).

For *b*'s vedkommende er det så godt som problemfrit at afgøre om bogstavet skal betragtes som et selvstændigt grafem eller som en del af et komplekst grafem. Hele 95% af *b*-forekomsterne på ordlisten var nemlig tilfælde hvor *b* optræder som simpelt grafem, dvs. repræsenterer en lyd 'på egen hånd'. Af de sidste 5 % udgjorde den let



vedkommende tillige i en computerlæsbar lydskrift. Denne liste dannede grundlag for en undersøgelse af bogstav-lyd-relationen i (næsten) alle danske en- og tostavelsesord.

Inden jeg kunne begynde at lave statistiske opgørelser, måtte der dog et større forarbejde til. Opslagsordene blev med uvurderlig hjælp fra en datalingvist suppleret med alle gængse bøjningsformer, og manglende udtalesformer blev tilføjet med automatiske lydskrivningsmetoder. Ordlisten blev derefter tjekket for fejl og mangler. Resultatet blev en liste med 47.148 ordformer, hvor hver ordform var en unik kombination af en ortografisk og en fonetisk form.

Endelig blev ordene opdelt i bogstav-lyd-relationer. Dette er uproblematisk i ord som *kat* og *hus* hvor bogstaver og lyde kan matches en-til-en. Men hvad med et ord som *købmand* som i den gængse udtale har stumt *b* og stumt *d*? Hvis man kategoriserer bogstaver som stumme, sker der det at de forsvinder ud af konsistensregnskabet i staveretningen. Det fører til kunstigt høje konsistenser og ødelægger samtidig muligheden for at sammenligne med konsistensen i læseretningen. Derfor valgte jeg i stedet at betragte *øb* og *nd* i *købmand* som komplekse grafemer der repræsenterede hhv. en *ø*- og en *n*-lyd.

minimale bogstav-lyd-relationer mærkværdige resultater. Ordet *vignet* kan i princippet opdeles sådan at *g*'et repræsenterer en *n*-lyd og *n*'et en *j*-lyd. Disse usædvanlige forbindelser forekommer imidlertid kun i hinandens selskab: *g* kan kun sige [n] før et *n* der siger [j], og vice versa. I dette og nogle få andre tilfælde forekom det mere meningsfuldt at tillade ikke-minimale bogstav-lyd-forbindelser og sige at *gn* som helhed repræsenterer lyd-følgen [nj].

**A, B og C: Simple eller komplekse grafemer?** I det følgende vil jeg fortælle om nogle af undersøgelsens resultater. Og hvad er mere nærliggende end at begynde med alfabetets første tre bogstaver?

Af samtlige bogstaver i de undersøgte ordformer var 4,9 % *a*'er, 2,4 % var *b*'er, mens kun 0,4 % var *c*'er. Nogle bogstaver spiller således en langt vigtigere rolle end andre.

*A*, *b* og *c* forekommer både som simple grafemer og som en del af komplekse grafemer. *A* indgår blandt andet i de komplekse grafemer *ar* (som i *farlig*) og *ag* (som i *goddag*). *B* indgår blandt andet i *bb* (som i *skubbe*) og *øb* (som i *købmand*). Og *c* indgår

identificerbare dobbeltkonsonant *bb* 4,9 % af tilfældene. Tilfælde hvor *b* indgår i andre komplekse grafemer såsom *øb* i *købmand*, udgør altså kun 0,1 % af *b*-forekomsterne. Og så skal det nævnes at det jo ikke er forbudt at udtale *b*'et i ordet *købmand*, hvad mange sprogbrugere faktisk gør. Det er bare ikke ordets mest almindelige udtale (ifølge Den Danske Ordbog), som er den der er lagt til grund i undersøgelsen her.

Situationen er også forholdsvis klar når det gælder *a*, der har status af simpelt grafem for 86,5 % af forekomsternes vedkommende. Det komplekse grafem *ar* tegner sig for yderligere 9,9 % af forekomsterne, sådan at i alt 97,4 % af *a*-forekomsterne udgøres af tilfælde hvor *a* enten er et selvstændigt grafem eller en del af grafemet *ar*.

Derimod fordeler forekomsterne af bogstavet *c* sig mere ligeligt på simple grafemer (48,7 %) og tilfælde hvor *c* indgår i et komplekst grafem (51,3 %). Ser man på samtlige bogstaver i de undersøgte ordformer, repræsenterede 81,6 % af forekomsterne simple grafemer. *C* er altså ikke bare et sjældent bogstav, men også et bogstav som usædvanlig ofte indgår i komplekse grafemer.

Tabel 1. C-udtaler og deres sandsynlighed.

s-lyd	cykel	269	0,435
k-lyd	cola	251	0,406
g-lyd	action	91	0,147
sj-lyd	cello	4	0,006
d-lyd	boccia	2	0,003
t-lyd	ciao	1	0,002
I alt		618	1,000

**Er A, B og C til at stole på?** Hvis man ser snævert på de tilfælde hvor *a*, *b* og *c* fungerer som simple grafemer, hvor svært er det så at finde frem til deres udtale? Som et mål på dét kan man beregne hvor konsistent de tre bogstaver er forbundet med de lyde de repræsenterer. Grafemet *c*, der kan svare til seks forskellige udtaler, kan tjene som eksempel.

Som det ses i Tabel 1, er de to hyppigste *c*-udtaler omtrent lige hyppige – og dermed nogenlunde lige sandsynlige. Sandsynligheden for at et *c* udtales som en *s*-lyd er 0,435, mens sandsynligheden for en *k*-lyd er 0,406. Som mål på *c*-grafemets konsistens kunne man vælge at tage gennemsnittet af sandsynlighederne for hver af de seks *c*-udtaler – dvs. 0,167. Dermed ville sandsynligheden for den helt afvigende *t*-lyd imidlertid veje lige så tungt som sandsynlighederne for *s*-lyd og *k*-lyd. For at undgå sådan en skævvridning kan man i stedet bruge det vægtede gennemsnit af sandsynlighederne – dvs. 0,376. (Her indgår hver sandsynlighed med en vægt der svarer til andelen af forekomster: 269/618 for *s*-lyd, 251/618 for *k*-lyd osv.)

I Tabel 2 vises konsistenscoefficients for både *a*, *b* og *c*. Første kolonne viser ubetingede konsistenser, dvs. konsistenscoefficients beregnet som netop vist for *c*. Det ses her at *a* med en konsistens på 0,324 er end-

nu mere inkonsistent end *c*. Derimod ligger *b* med 0,994 meget tæt på den maksimale konsistenscoeffcient (som er 1), hvilket vil sige at udtalen er næsten helt forudsigelig.

Nu er det sådan at bogstavets udtale ofte afhænger af sammenhængen. Eksempelvis udtales *c* næsten altid som en *s*-lyd før *y*, hvorimod udtalen før *o* næsten altid er en *k*-lyd. Hvis man beregner sandsynlighederne for de forskellige udtaler under hensyntagen til enten foregående eller følgende grafem og igen tager det vægtede gennemsnit, får man såkaldt 'betingede' konsistensmål. De betingede konsistenser for *a*, *b* og *c* er vist i tabelens to sidste kolonner.

Som man kan se, er de betingede konsistenser for *a* og *c* markant højere end de ubetingede konsistenser. Det gælder især for *c*, og det gælder især når det er det følgende grafem der tages hensyn til.

Selv om nogle grafemer umiddelbart er meget inkonsistente, kan situationen altså være langt bedre hvis der tages hensyn til sammenhængen. For læseren kan det med andre ord være en fordel at se på grafemernes omgivelser. I tilfældet *c* kan det betale sig at se på om det følgende grafem er en vokal og i så fald hvilken. I tilfældet *a* bliver eksempelvis en åben *a*-udtale mere sandsynlig når der følger et *m*, men mindre sandsynlig når der følger et *t* – sammenlign *a*-

udtalerne i *ham* og *hat*. Andre grafemer, som *b*, har en høj grad af konsistens uanset om man inddrager sammenhængen.

**Læsning og stavning** For at få et samlet indtryk af forholdene i dansk ortografi har jeg beregnet den gennemsnitlige konsistens for samtlige grafemer (både simple og komplekse). Igen er grafemerne vægtet efter deres hyppighed. Tilsvarende beregninger er lavet i staveretningen, fra lyde til grafemer. Resultatet ses i Tabel 3.

Som det fremgår, er graden af konsistens samlet set højere i staveretningen end i læseretningen. Forskellen skyldes især at vokaler er mindre konsistente i læseretningen end i staveretningen. Vokalerne er dog mindre konsistente end konsonanterne i begge retninger.

Både i læse- og staveretningen er de betingede konsistenser klart højere end de ubetingede. Interessant nok er den samlede fremgang cirka lige stor når hhv. foregående og følgende grafem/lyd inddrages. Men det er ikke ligegyldigt om der er tale om vokaler eller konsonanter. For vokaler ses den største fremgang når man inddrager følgende grafem/lyd, mens det for konsonanter er foregående grafem/lyd der giver den største fremgang. Der er ikke desværre ikke plads til at gå dybere ind i årsagerne til dette her.

Når vokalgrafemerne har så lav en konsistensgrad, skyldes det blandt andet at de fleste vokallyde kommer i tre udgaver: De kan både være korte og lange og have stød. Vokalbogstavet *u* repræsenterer for eksempel en kort *u*-lyd i *bus*, en lang *u*-lyd i *huse* og en *u*-lyd med stød i *hus*. At finde ud af

Tabel 2. Konsistenscoefficients for *a*, *b* og *c*.

		Foregående grafem	Følgende grafem
<i>a</i>	0,324	0,396	0,583
<i>b</i>	0,994	0,994	0,994
<i>c</i>	0,376	0,601	0,842



Tabel 3. Konsistenskoefficienter for danske grafemer og lyde.

	Ubetinget	Betinget	
		Foregående grafem/lyd	Følgende grafem/lyd
<b>Læseretningen (alle grafemer)</b>	<b>0,604</b>	<b>0,743</b>	<b>0,754</b>
vokalgrafemer	0,378	0,479	0,600
konsonantgrafemer	0,713	0,874	0,829
<b>Staveretningen (alle lyde)</b>	<b>0,725</b>	<b>0,827</b>	<b>0,811</b>
vokallyde	0,672	0,768	0,807
konsonantlyde	0,750	0,875	0,813

hvilken slags *u*-lyd et givet *u*-bogstav repræsenterer, er selvfølgelig vanskeligere end blot at finde ud af at det repræsenterer en *u*-lyd.

Hvis man ser stort på fænomenerne længde og stød og kun registrerer vokalerne kvalitativt, så ser regnskabet for vokaler i læseretningen noget bedre ud: Den ubetingede konsistens går op til 0,487. Til gengæld går den ubetingede vokalkonsistens i staveretningen ned til 0,607. Der ses tilsvarende stigninger og fald i de betingede konsistenser.

Konklusionen bliver at mens konsonanter er *nogenlunde* til at stole på, så er vokalerne påfaldende inkonsistente, især i læseretningen. Situationen bedres dog noget hvis man ser på de betingede konsistenser, og hvis man ser bort fra længde og stød.

### Vokalkonsistens i dansk og engelsk

Hvordan placerer dette resultat så de danske vokaler i en international sammenhæng? Her er det blandt andet muligt at sammenligne med en undersøgelse af amerikansk engelsk. Sprogforskerne Brett Kessler og Rebecca Treiman fandt for engelske vokaler en ubetinget konsistens på 0,717 i læseretningen, hvilket jo langt overgår tallet 0,378 for danske vokaler. På dette område er den engelske ortografi altså *mere* regelmæssig end den danske. I staveretningen var det imidlertid omvendt. Her var konsistensen af engelske vokaler 0,529 – en hel del lavere end tallet 0,672 for danske vokaler.

Man bør dog bide mærke i at det ikke kun er graden af konsistens der adskiller dansk og engelsk, men også brugen af komplekse grafemer. Komplekse vokalgrafemer bruges langt oftere i engelsk end i dansk, jf. engelske

ord som *blue*, *green* og *good*. Og det skal bemærkes at Kessler og Treiman betragtede *e*'et til sidst i ord som *hate* og *love* som en del af todelte komplekse vokalgrafemer (hhv. *a-e* og *o-e*).

Kessler og Treiman undersøgte kun enstavelsesord, og de udelod alle regelmæssige bøjningsformer. De anførte danske tal ændrer sig imidlertid kun en smule når man begrænser grundlaget til et nogenlunde tilsvarende udsnit af danske enstavelsesord.

Når det er interessant at kende til specifikke ligheder og forskelle mellem dansk og engelsk ortografi, skyldes det ikke mindst at langt de fleste undersøgelser af læsning og stavning er gennemført i engelsktalende lande. Når man kender ords ortografiske egenskaber i detaljer, bliver det lettere at vurdere om forskningsresultater kan generaliseres fra den ene ortografi til den anden, og det bliver muligt at lave mere velkontrollerede tvær sproglige undersøgelser.

**Lette ord og svære ord** En pointe i min undersøgelse er at tallene først bliver rigtig interessante når man ser på deres betydning i praksis. Derfor undersøger jeg i projektets sidste (og endnu uafsluttede) del hvilken betydning andelen af komplekse grafemer og grafemernes gennemsnitlige konsistens

har for om ord er lette eller svære at læse og stave for børn i skolens yngste klasser.

Nu kan nogen måske få den mistanke at inkonsistensen i dansk er det sørgelige resultat af sjuksk med udtalen i de yngste generationer. Det er imidlertid ikke her forklaringen ligger. Den Danske Ordbog, som var udgangspunktet for undersøgelsen, anfører nemlig tydelige og ganske konservative udtaler. Afstanden mellem skrift og tale går i mange tilfælde langt tilbage i tiden og er måske nærmest at betragte som en del af den danske kulturarv. I værket *Det Danske Sprogs Historie* (Bind 1, 1944) mener Peter Skautrup at stavemåderne allerede i de ældste håndskrifter af Jyske Lov (fra omkring år 1300) afviger fra datidens udtaler af næsten hvert eneste ord. Det er tvivlsomt om Skautrup har ret i dette, men sikkert er det at afvigelser fra det alfabetiske princip ikke er noget nyt fænomen.

Afslutningsvis vil jeg minde om at min undersøgelse kun omfattede ord med en og to stavelser. En optælling i tekstsamlingen *Korpus 2000* (<http://korpus.dsl.dk>) viser dog at enstavelsesord udgør ikke mindre end halvdelen af ordene i danske tekster, mens tostavelsesord udgør yderligere en fjerdedel. Undersøgelsen dækker således en meget stor del af de ord der læses og skrives på dansk.



**HOLGER JUUL** er postdoc ved Center for Læseforskning, Institut for Nordiske Studier og Sprogvidenskab, KU. Projektet 'Lette ord og svære ord. Bogstav-lydrelationen i dansk og dens betydning for udviklingen af læse- og stavefærdighed' er støttet af Forskningsrådet for Kultur og Kommunikation.

# 4:08

DECEMBER 2008

23. ÅRGANG

FORSKNINGSRÅDET FOR

KULTUR OG KOMMUNIKATION

# magasinet **Humaniora**



**Japansk pop**  
**Indiana Jones på spil**  
**Er danske bogstaver til at stole på?**