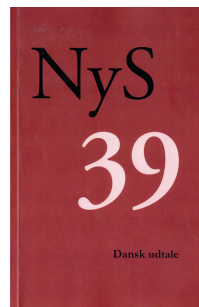


# NyS

Titel:	K-a-t-t-e-p-i-n-er Om komplekse bogstav-lyd-forbindelser i danske ord
Forfatter:	Holger Juul
Kilde:	<i>NyS – Nydanske Sprogstudier 39. Dansk udtale,</i> 2010, s. 10-32
Udgivet af:	NyS i samarbejde med Dansk Sprognævn
URL:	<a href="http://www.nys.dk">www.nys.dk</a>



© NyS og artiklens forfatter

## 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 NyS-numre (NyS 1-36) 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.

# K-a-t-t-e-p-i-n-e-r

## Om komplekse bogstav-lyd-forbindelser i danske ord

HOLGER JUUL

Da Dansk Sprognævn i 1986 tilbød danskerne stavemåden *majonæse* som alternativ til *mayonnaise*, var modtagelsen som bekendt kølig (Jacobsen 2010: 228 ff.). Den nye stavemåde vinder kun langsomt frem og giver for tiden 16.300 google-hits (dansksprogede netsider), mens *mayonnaise* giver 53.900 hits. Dette selv om *majonæse* på flere måder er langt enklere: For det første får tre af lydene deres mest almindelige stavemåde, nemlig *j*-lyden, *n*-lyden og *æ*-lyden – som jo netop omtales sådan og ikke som „*y*-lyden“, „*nn*-lyden“ og „*ai*-lyden“. For det andet bliver relationen mellem bogstaver og lyde bragt i overensstemmelse med det alfabetiske princip, sådan at der til en enkelt lyd i ordet blot svarer et enkelt bogstav.

Den første forenkling betyder at det bliver lettere at forudsige hvordan de enkelte bogstaver skal udtales, og hvordan de enkelte lyde skal staves: Bogstav-lyd-forbindelserne bliver mere konsistente. Den anden forenkling betyder at det bliver nemmere at finde ud af hvilke bogstaver og lyde der faktisk hører sammen i bogstav-lyd-forbindelser – eller med et mere teknisk udtryk: hvordan bogstaver og lyde skal aligneres. Det er oplagt at komplekse bogstavenheder – som når *ai* i *mayonnaise* kun repræsenterer en enkelt lyd, og ikke to lyde som i *aids* og *naiw* – kan udgøre en vanskelighed når man skal læse og skrive ord, især hvis man ikke kender ordet på forhånd. Ligeledes kan komplekse lydenheder være en vanskelighed – som når [dj] i ordet *job* [djʌb] modsvares af et enkelt bogstav, og ikke af to.

Nu tyder modtagelsen af formen *majonæse* jo på at danskerne generelt har det fint med en ortografi der byder på visse uregelmæssigheder. På

den anden side er det en almindelig antagelse at uregelmæssighederne vitterligt udgør en vanskelighed, især for børn der skal lære at læse og stave. Man kan endda møde den opfattelse at afvigelserne fra det alfabetiske princip i dansk efterhånden er så mange at det knap nok giver mening at lære børn om bogstav-lyd-forbindelser.

Når man taler om uregelmæssighederne i dansk ortografi, er det næsten altid bogstav-lyd-forbindelsernes begrænsede konsistens der bliver fremhævet som det store problem. Imidlertid kan man mene at det lidt mere oversete problem med komplekse enheder faktisk er det mest grundlæggende. Først når man ved hvilke bogstaver der tilsammen udgør en enhed, kan man begynde at se på hvor mange forskellige udtaler denne enhed kan have. Og der er faktisk en del forskning der tyder på at børn i de små klasser har ganske væsentlige problemer med komplekse enheder (Elbro 2004, Juul & Sigurdsson 2005, Juul m.fl. 2007, Reece & Treiman 2001, Spencer under udg.).

I denne artikel vil jeg se nærmere på bogstav-lyd-forbindelserne i dansk og søge svar på hvor mange forskellige forbindelser der egentlig findes, hvor mange af dem der involverer komplekse bogstav- eller lydenheder, og endelig på hvilke enheder og forbindelser der anvendes mest. Artiklen handler også om de principielle vanskeligheder som selve aligneringen af bogstaver og lyde stiller én over for.

Inden jeg går i gang, vil jeg forklare et par termer jeg har dannet til lejligheden. Jeg taler i artiklen om *bogstavenheder* og skelner her mellem *simple* enheder (der udgøres af et enkelt bogstav) og komplekse enheder (der udgøres af to eller flere bogstaver). Tilsvarende taler jeg om simple og komplekse *lydenheder* (som udgøres af en enkelt lyd, hhv. en flerhed af lyde).

Den traditionelle betegnelse for en bogstavenhed er et *grafem*, hvor man også kan tale om komplekse grafemer og eventuelt om *digrafer*, *trigrafer* osv. (dvs. grafemer med to, hhv. tre bogstaver). Det er nogle udmærkede termer, men der mangler, så vidt jeg ved, tilsvarende veletablerede termer om lydligge enheder. Den oplagte parallel ville være at tale om

*fonemer* – men dette ord refererer (i hvert fald i den sprogvidenskabelige tradition) til abstrakte fonologiske enheder, hvis psykologiske realitet for sprogbrugerne i mange tilfælde er uklar (Basbøll 2005, Grønnum 2005). Da min hensigt her er at undersøge relationen mellem bogstaver og de *ikke-abstrakte* enkeltlyde som gengives i lydskrivningen af ord i en udtaleordbog, kunne jeg vælge at tale om *foner* i stedet for fonemer. Tilfælde som [dj] i *job* kunne så omtales som en *kompleks fon* eller som en *difon*.

Jeg foretrækker imidlertid at danne nye sammensætninger med de betegnelser som alle kender: bogstaver og lyde. Jeg taler således om bogstavenheder og lydenheder, og forbindelser mellem dem omtaler jeg som *bogstav-lyd-forbindelser*. Hvis én af enhederne i forbindelsen er kompleks, taler jeg om en kompleks bogstav-lyd-forbindelse, og hvis begge enheder er komplekse, taler jeg om en *dobbelt kompleks* bogstav-lyd-forbindelse.

## SYV PRINCIPPER FOR ALIGNERING

Når antallet af bogstaver og lyde i et ord svarer til hinanden, er alig-neringen normalt uproblematisk, hvis man tager udgangspunkt i en situation hvor både stavemåden og udtalen er kendt på forhånd. Ord-formen *kat* [kad] har eksempelvis tre bogstaver og tre lyde, som det er ganske ligetil at alignere så resultatet bliver tre simple bogstav-lyd-forbindelser:  $k=[k]$   $a=[a]$   $t=[d]$ .

Men hvad gør man når antallet af bogstaver og lyde ikke matcher hinanden? Lad os som et eksempel se på flertalsformen *katte* ['kadə] med fem bogstaver, men kun fire lyde. Med mindre man vil lade et af bogstaverne være stumt (se nedenfor om denne mulighed), må en af de fire lyde forbindes med to bogstaver. Der er fire muligheder<sup>1</sup>:

$$(1) \quad ka=[k] \quad t=[a] \quad t=[d] \quad e=[ə]$$

$$(2) \quad k=[k] \quad at=[a] \quad t=[d] \quad e=[ə]$$

$$(3) \quad k=[k] \quad a=[a] \quad tt=[d] \quad e=[ə]$$

(4)  $k=[k]$   $a=[a]$   $t=[d]$   $te=[ə]$

Disse fire løsninger er bestemt ikke lige gode. De fleste vil uden tvivl være enige i at løsning (3) med  $tt=[d]$  klart er at foretrække – og i at løsning (1), hvor konsonantbogstavet  $t$  aligneres med vokallyden  $[a]$ , er helt uacceptabel. Som bruger af skriftsproget danner man sig tydeligvis en opfattelse af hvordan bogstaver og lyde kan høre sammen, og af hvordan de *ikke* kan høre sammen.

Men hvad er det egentlig der gør (3) til den mest attraktive løsning? Hvad er der fx i vejen med mulighed (2)? Her kan forbindelsen  $at=[a]$  virke ejendommelig – men der er faktisk ord hvor den bliver nødvendig (ord som *nougat*  $[ˈnuga]$  og *muscat*  $[myˈsga]$ ). Så hvorfor ikke også bruge den i ordet *katte*? I håb om at kunne give konsekvente svar på den slags spørgsmål har jeg opstillet syv principper for alignering. Principperne står i gæld til diskussion af segmenteringsprincipper i Carney 1994: 32 ff.

1. *Bogstaver og lyde i et ord skal aligneres restløst.*
2. *Aligneringen skal give bogstav-lyd-forbindelser der er så simple som muligt.*
3. *Bogstaver og lyde skal aligneres uden overlap, opsplitninger eller ombytninger.*
4. *Bogstaver og lyde skal forbindes direkte.*
5. *Bogstav-lyd-forbindelser med begrænset anvendelse i andre ordformer skal undgås.*
6. *Komplekse bogstav-lyd-forbindelser skal have lighed med simple bogstav-lyd-forbindelser.*
7. *Aligneringen skal respektere morfemgrænser.*

De første fire principper definerer hvilke former for bogstav-lyd-forbindelser der overhovedet tillades, mens de sidste tre er retningslinjer der skal støtte valget mellem alternative løsninger. I det følgende motiveres og diskuteres hvert princip kort. I næste afsnit ser jeg lidt nærmere på nogle vanskeligheder i anvendelsen.

#### *1. Bogstaver og lyde i et ord skal aligneres restløst*

Dette princip betyder at bogstaver ikke må være stumme, ligesom lyde ikke må være uden ortografisk repræsentation. Man kan indvende at

princippet strider mod den gængse beskrivelse af ortografien, ifølge hvilken eksempelvis ordet *hund* [hun<sup>?</sup>] ender på et stumt *d*. Denne beskrivelse er imidlertid problematisk hvis man vil lave konsistensberegninger i staveretningen (fra udtale til ortografi) – for så må man sammenholde antallet af tilfælde hvor ingenting repræsenteres af *d* (eller et andet stumt bogstav), med tilfælde hvor ingenting repræsenteres af ingenting. Og det giver naturligvis ikke mening.

Princip nr. 1 betyder at *hund* må aligneres med en kompleks bogstav-enhed finalt: *b*=[h] *u*=[u] *nd*=[n<sup>?</sup>]. Derved bliver konsistensberegninger i læseretningen og staveretningen sammenlignelige, fordi de kan tage udgangspunkt i de samme bogstav-lyd-forbindelser.

En anden fordel er at man undgår at skulle tage stilling til hvornår stumme bogstaver er at foretrække for komplekse bogstavenheder. I tilfældet *katte* kan man nok vælge at sige at det ene af de to *t*'er er stumt – men at betragte *tt* som en kompleks bogstavenhed fremhæver bedre den regelmæssighed i dansk at korte trykstærke vokaler ofte følges af konsonantfordoblinger. Endvidere er det ikke altid indlysende hvilket bogstav i en fordobling der skal anses for at være stumt. Princippet betyder at man helt undgår inkonsekvente eller tilfældige valg mellem stumme bogstaver og komplekse bogstavenheder. I staveretningen betyder princippet at fx ordet *taxa* ['tagsa] må aligneres med forbindelsen *x*=[gs] frem for med et „ustavet“ [g] eller [s].

Hermed er ikke sagt at det *generelt* er u hensigtsmæssigt at tale om stumme bogstaver, men blot at tilladelsen af stumme bogstaver giver en række analyseproblemer, som jeg gerne vil undgå her.

## *2. Aligeringen skal give bogstav-lyd-forbindelser der er så simple som muligt*

Når antallet af bogstaver og lyde ikke matcher, er komplekse enheder som netop vist en nødvendighed, hvis man har fravalgt stumme bogstaver og „ustavede“ lyde. Princip nr. 2 siger at komplekse enheder skal undgås så langt det overhovedet er muligt.

Dette princip er motiveret af et ønske om at minimere det samlede antal regler der forbinder ortografiske og fonetiske former. Hvis man tillader komplekse enheder der kan opdeles yderligere, kan man ultimativt lige så godt betragte hvert eneste ord som et komplekst, uanalyseret hele: *kat*=[kad] og *katte*=['kadə]. Dermed tilsøres den

centrale egenskab ved alfabetiske skriftsystemer at en opdeling af ord i et relativt begrænset antal bogstav-lyd-forbindelser er en mulighed. Og spørgsmålet her er jo netop hvor få og hvor simple bogstav-lyd-forbindelser man kan klare sig med i dansk.

Det er naturligvis en anden sag i hvilken grad ortografibrugere faktisk udnytter de minimale bogstav-lyd-forbindelser der opstilles. Princippet skal ikke forstås som en påstand om at ortografibrugere *aldrig* benytter sig af komplekse enheder der kan opdeles yderligere – jf. den såkaldte grain size-hypotese i læseforskningen (Ziegler & Goswami 2005).

### 3. Bogstaver og lyde skal aligneres uden overlap, opsplittninger eller ombytninger

En anden central egenskab ved alfabetiske skriftsystemer er at bogstav-enheder *én ad gangen* kan omkodes til lydenheder (og vice versa). Princip nr. 3 skal sikre at aligeringen respekterer denne egenskab. Ovenfor anførte jeg fire mulige aligninger af ordet *katte*. Men antallet af muligheder bliver langt større hvis ikke det kræves at bogstaver og lyde aligneres i indbyrdes klart adskilte forbindelser. Princippet siger at muligheder hvor enhederne overlapper, opsplittes eller byttes om, bør undgås – også i tilfælde hvor der kan argumenteres for dem.

I tilfældet *katte* kan man mene at det er fordoblingen *tt* der angiver at vokalen [a] er kort. Men forbuddet mod overlap forhindrer at [a] forbindes med dobbeltkonsonanten (fx *tt*=[ad]) hvis vokalen samtidig er forbundet med en anden bogstavenhed (*a*=[a]). At tydeliggøre forbindelsen mellem vokallængde og dobbeltkonsonant ved at tillade en mere kompleks forbindelse *att*=[ad] kunne være en løsning – men den er allerede forbudt af princip nr. 2. Dette forbud mod overlappende forbindelser betyder også at [ŋ] i ordet *tango* ['tango] ikke kan forbindes med en kompleks bogstavenhed *ng*, hvis *g*'et samtidig skal forbindes med [g]. Lyden [ŋ] må derfor i dette tilfælde aligneres med *n* alene.

Forbuddet mod at splitte enheder op betyder at *e*'et til sidst i engelske låneord som *website* ['web,sajd] ikke kan kombineres med *i*'et til en splittet bogstavenhed *i\_e* med udtalen [aj]. Kessler & Treiman 2001 og Carney 1994 tillader faktisk sådanne opsplittede grafemer i deres beskrivelser af engelsk, hvor et finalt *e* ofte fungerer som markør af vokalers udtale (sammenlign fx udtalen af engelsk *sit* og *site*). I en beskrivelse af dansk ortografi forekommer det uhensigtsmæssigt at

tillade en så afvigende type bogstavenhed alene for at gøre rede for nogle få engelske låneord. I stedet må endelsen af ordet aligneres således:  $i=[aj]$ ,  $te=[d]$ .

Forbuddet mod at bytte om på enheder betyder blandt andet at der ikke må byttes om på de to midterste bogstaver/lyde i ordet *vignet* [vin'jød]. Det er ellers fristende for at undgå to højst usædvanlige bogstav-lyd-forbindelser:  $g=[n]$  og  $n=[j]$ . Her ville en ombytning af enheder (dvs. en forbindelse på kryds) give de mere velkendte forbindelser  $n=[n]$  og  $g=[j]$ . Men det er næppe ønskeligt at åbne for forbindelser på kryds og tværs. I stedet mener jeg at man her må give køb på kravet om at bogstav-lyd-forbindelser skal være så simple som muligt (princip nr. 2) og tillade den dobbelt komplekse forbindelse  $gn=[nj]$  ud fra den betragtning at  $g=[n]$  alligevel kun ville forekomme før  $n=[j]$  og vice versa.

#### 4. Bogstaver og lyde skal forbindes direkte

Dette princip betyder at bogstaver og lyde ikke kan forbindes via et mere eller mindre abstrakt mellemed. Det kunne måske ellers være fristende at forbinde eksempelvis *give* med udtalen [gi<sup>7</sup>] via en mere bogstavnær udtale ['gi:və]. Men mellemed er naturligvis ikke hensigtsmæssige hvis formålet er at undersøge den direkte relation mellem skrift og udtale, og det er netop sådan en undersøgelse jeg mener man må begynde med.

Princip nr. 4 betyder at der heller ikke må inddrages abstrakte lydenheder, som når Becker-Christensen 1988 i sin beskrivelse af dansk ortografi ikke forbinder bogstavenheden *sj* direkte med lyden [ɕ], men i stedet forbinder *s* med [s] og *j* med [j]. Hos Becker-Christensen er [ɕ] resultatet af en efterfølgende fusion af [s] og [j], sådan at forbindelsen mellem *sj* og [ɕ] er indirekte, opdelt i to faser med [s] og [j] som mellemed.

Indirekte forbindelser kan i nogle tilfælde forenkle beskrivelsen af bogstavernes udtale betragteligt – som i det nævnte tilfælde hvor den komplekse bogstavenhed *sj* opdeles i to simple enheder (se også beskrivelsen af bogstav-lyd-relationer i dansk i Basbøll 2005). Prisen for forenklingen er imidlertid ganske høj, da de abstrakte enheders (fonemers) psykologiske realitet for sprogbrugerne som nævnt er uklar



(Basbøll 2005, Grønnum 2005). Ikke mindst er det uklart i hvilket omfang kendskab til ortografien præger opfattelsen af lydlig enheder

5. *Bogstav-lyd-forbindelser med begrænset anvendelse i andre ordformer skal undgås*

Dette princip bidrager til valget mellem de fire mulige aligninger af *katte* ['kadə], idet det peger på en opdeling der undgår sjældent anvendte forbindelser. Opdelingsmulighed (3) er klart attraktiv fordi den komplekse forbindelse *tt*=[d] også kan benyttes i opdelingen af en lang række andre ord med intervokalisk *tt*. Anvendeligheden af de øvrige mulige komplekse forbindelser er derimod meget mere begrænset.

Også dette princip er motiveret af et ønske om regeløkonomi. Hvis ikke man prioriterer de mest anvendte forbindelser, kan man ende med langt flere forbindelsesregler end man strengt taget har brug for. Med valget af *tt*=[d] kan forbindelserne *ka*=[k] og *te*=[ə] undgås helt, mens forbindelsen *at*=[a] kun er nødvendig i analysen af et par enkelte ord (*muscat* og *nongat*).

Princip nr. 5 er vigtigt hvis aligneringen skal afspejle ords reelle sværhedsgrad. Man må formode at et ord bliver lettere at læse og stave når de indgående bogstav-lyd-forbindelser i forvejen kendes fra andre ord, og aligneringen bør så vidt muligt tilgodesee dette. En vanskelighed ved princippet er imidlertid at det inddrager aligneringssmuligheders konsekvenser for den samlede analyse. Med mindre man faktisk gennemfører alternative analyser og sammenligner resultaterne, må forbindelsernes samlede anvendelighed bero på et skøn.

#### 6. *Komplekse bogstav-lyd-forbindelser skal have lighed med simple bogstav-lyd-forbindelser*

Dette princip skal sikre at komplekse bogstav-lyd-forbindelser så vidt muligt fremstår som gennemskuelige modifikationer af simple bogstav-lyd-forbindelser.

I tilfældet *katte* ['kadə] er den komplekse forbindelse *te*=[ə] i løsning (4) anderledes end de øvrige derved at bogstavet længst til venstre (*t*'et) ikke indgår i en simpel bogstav-lyd-forbindelse med samme lyd-værdi. (Det gør derimod bogstavet til højre (*e*'et) idet *e*=[ə] er en meget anvendt forbindelse.) Eftersom man læser og skriver fra venstre mod højre, forekommer det hensigtsmæssigt at ligheden med en simpel

bogstav-lyd-forbindelse findes på venstrekanten, sådan at man i læsning og skrivning af et ord begynder med den velkendte del af en kompleks forbindelse. For princip nr. 6 kan der derfor opstilles et underprincip der prioriterer venstrelighed højere end højrelighed. Det betyder at  $ka=[k]$  og  $at=[a]$  i løsning (1) og (2) er bedre end  $te=[ə]$  i løsning (4). Men  $tt=[d]$  i løsning (3) er endnu bedre fordi der både er venstre- og højrelighed, hvilket øger forbindelsens gennemskuelighed. Princip nr. 6 støtter således løsning (3).

Prioriteringen af venstrelighed betyder at  $u'$ et i *guitar*  $[ˈgi,taʔ]$  bør danne enhed med  $g'$ et frem for med  $u'$ et, altså  $gu=[g]$  og ikke  $ui=[i]$ . Men undtagelser må naturligvis accepteres hvor problembogstavet eller problemlyden er placeret først i ordet og derfor kun kan gå til højre – således *hjem*  $[jemʔ]$ , der giver  $hj=[j]$ , og *job*  $[dʒab]$ , der giver  $j=[dʒ]$ , begge med højrelighed.

En anden form for lighed som kan være fristende at inddrage, er *fonetisk* lighed (jf. begrebet *phonetic transparency*, der benyttes af Carney 1994). I tilfældet *katte*  $[ˈkadə]$  opleves forbindelsen  $t=[a]$  i løsning (1) formentlig ikke kun som problematisk fordi den ikke kendes fra andre ord, men også fordi  $t$ , der normalt repræsenterer en konsonant, her forbindes med en vokallyd. Men det er vanskeligt at operationalisere begrebet fonetisk lighed, da lighederne kan gælde mange forskellige forhold: syllabicitet, artikulationssted, artikulationsmåde, stemthed etc.

### 7. Aligneringen skal respektere morfemgrænser

Dette princip skal sikre at aligneringen så vidt muligt svarer til den almindelige opfattelse af ords struktur. Ud fra de tidligere opstillede principper kan *dt* i ordet *kridttid*  $[ˈkʁid,tiðʔ]$  udmærket fordeles på forbindelserne  $d=[d]$  og  $tt=[t]$  – idet begge forbindelser kendes fra andre ord, fx *ditto*  $[ˈdito]$ . Men denne opdeling er kontraintuitiv fordi  $tt=[t]$  både dækker det sidste bogstav i førsteleddet *kridt* og det første bogstav i andetleddet *tid*. Princippet peger på en alignering så morfemgrænsen også er en grænse mellem bogstav-lyd-forbindelser:  $dt=[d]$  og  $t=[t]$ .

Princip nr. 7 betyder også at *ndt* i *sandt*  $[sanʔd]$  hellere skal aligneres i overensstemmelse med ordets morfologiske struktur *sand*+*t* med  $nd=[nʔ]$  og  $t=[d]$  frem for  $n=[nʔ]$  og  $dt=[d]$ . En fastholdelse af prin-

cippet kan imidlertid give en konflikt med princippet om at undgå sjældne forbindelser. Hvis den morfologiske struktur i ordet *sød+t* [sød] skal respekteres, bliver forbindelsen *ød*=[ø] nødvendig, og denne forbindelse kendes (til forskel fra *nd* i sandt) ikke fra den ubøjede form. Her foretrækker jeg at opstille *dt* som en kompleks bogstavenhed *dt*=[d] (med anvendelse i mange ord) på trods af morfemgrænsen.

En anden vanskelighed ved princippet er at det tit er tvivlsomt om sprogbrugerne faktisk oplever en morfemgrænse. Opleves ordet *tillid* ['te?'] fx som en sammensætning *til+lid*? I så fald bør dobbeltkonsonanten *ll* ifølge princip nr. 7 fordeles på to enheder: *il*=[e] og *l*=[l]. I sådanne tilfælde lægger jeg større vægt på princip nr. 5, og foretrækker den hyppige forbindelse *ll*=[l] frem for den morfemgrænserespekterende forbindelse *il*=[e] – som dog stadig er nødvendig i et ord som *tilfreds* [te'fɛs]. Jeg vender tilbage til den slags principkonflikter i næste afsnit.

## HVOR MANGE BOGSTAV-LYD-FORBINDELSER HAR VI I DANSK?

Med de ovenfor anførte principper som grundlag vil jeg nu give et svar på hvor mange forskellige bogstav-lyd-forbindelser der er brug for i en dækkende beskrivelse af dansk ortografi. Svaret er 565 – heraf 186 simple forbindelser (33 %) og 379 komplekse forbindelser (67 %) – hvor en kompleks forbindelse som tidligere nævnt omfatter en kompleks bogstavenhed eller en kompleks lydenhed (eller begge dele).

Dette svar bygger selvfølgelig ikke kun på de syv principper for aligering, men også på en række metodemæssige forudsætninger, som jeg i det følgende vil gøre kort rede for. Svaret rejser en række nye spørgsmål som jeg vender tilbage til i næste afsnit.

### Metode

Tallet 565 bygger på en aligering af bogstaver og lyde i godt 47.000 forskellige ordformer, hvor en ordform er defineret som en unik kombination af en udtale og en stavemåde. Listen med ordformer<sup>2</sup> omfattede (næsten) alle en- og tostavellesord som kan afledes ved

almindelig bøjning af opslagsordene i *Den Danske Ordbog* (Hjorth & Kristensen, 2003-2005, herefter forkortet DDO). Hvor DDO anfører flere mulige stavemåder eller udtaler, blev den mest gængse form lagt til grund<sup>3</sup>. Opslagsordene blev velvilligt stillet til rådighed i elektronisk form af ordbogens redaktion, mens bøjningsformer og manglende udtaleangivelser blev tilføjet af lektor Peter Juel Henriksen, CBS. Bøjningsformerne udgør 34 % af ordene på den endelige liste.

Der blev set bort fra ord med flere stavelser for at gøre undersøgelsen overkommelig, men ordlisten har en ganske god dækningsgrad, idet omkring tre fjerdedele af ordene i Korpus 2000 (ifølge en optælling ved Peter Juel Henriksen) udgøres af en- og tostavelsesord. Dog må det bemærkes at DDO ikke medtager *proprier*, og at jeg i min undersøgelse så bort fra visse opslagsord. Det gjaldt bogstavord (eksempelvis *cd* og *wc*), ord der ikke indeholder mindst én vokal (eksempelvis interjektionerne *hm* og *ssh*), samt ord der indeholder ikke-bogstaver såsom tal, bindestreg eller ordmellemrum (således er *A5*, *c-mol* og *au pair* ikke taget med).

Aligneringen af bogstaver og lyde blev i praksis gennemført ved hjælp af et computerprogram udarbejdet af lektor Peter Juel Henriksen. Det ville have været interessant at basere programmet på en formalisering af de ovennævnte syv principper, men det lod sig desværre ikke gøre inden for projektets rammer. I stedet blev programmet baseret på applikationen af en regelliste der specificerede de tilladelige bogstav-lyd-forbindelser. Udgangspunktet for denne regelliste var oversigten over bogstav-lyd-forbindelser i Becker-Christensen 1988, som jeg modificerede og supplerede indtil samtlige ord på ordlisten kunne aligneres restløst.

For en del ords vedkommende gav regellisten mulighed for at alignere på mere end én måde. Eksempelvis blev løsningerne (2) og (3) ovenfor begge foreslået for ordet *katte* ['kade]. I sådanne tilfælde blev de uønskede løsninger undertrykt ved hjælp af ad hoc-indførte kontekstbegrænsninger, fx sådan at forbindelsen *at*=[a] kun blev tilladt ordfinalt og før *s* (og derfor ikke i *katte*).

### *Nogle vanskelige tilfælde*

I visse tilfælde var det ikke indlysende hvilke bogstav-lyd-forbindelser der skulle tillades, og hvilke aligneringsprincipper der skulle gives højest prioritet. Antallet af bogstaver og lyde er det samme i *vignet* [vin'jed], men som nævnt valgte jeg alligevel at opstille en dobbelt kompleks forbindelse  $gn=[nj]$  for at undgå to helt undtagelsesvisse forbindelser som alligevel kun ville forekomme i hinandens selskab. Tilsvarende overvejede jeg forbindelsen  $\mathfrak{z}\mathfrak{z}=[ds]$  i *pi $\mathfrak{z}\mathfrak{z}$ u* ['pidsa] (i lighed med behandlingen af ordet i Carney 1994). Dermed kunne man undgå en undtagelsesvis forbindelse  $\mathfrak{z}=[d]$ , der kun ville forekomme for  $\mathfrak{z}=[s]$ . Men da  $\mathfrak{z}=[s]$  har en langt videre anvendelse og ikke kun forekommer efter  $\mathfrak{z}=[d]$ , valgte jeg at alignere på minimalt niveau  $\mathfrak{z}=[d]$  og  $\mathfrak{z}=[s]$ .

Derimod tillod jeg en dobbelt kompleks forbindelse  $ge=[d\mathfrak{c}]$  i *image* ['imide] ud fra den betragtning at  $g=[d]$  og  $e=[\mathfrak{c}]$  kun ville forekomme i hinandens selskab. Dette var imidlertid ikke helt korrekt, da forbindelsen  $g=[d]$  faktisk blev nødvendig i *magyar* [ma'dja<sup>?</sup>]. Imidlertid forekom forbindelsen af vokalbogstav og konsonantlyd i  $e=[\mathfrak{c}]$  så fonetisk afvigende at jeg stadig foretrak den dobbelt komplekse forbindelse. Forbindelsen burde måske have været den endnu mere komplekse  $age=[id\mathfrak{c}]$ , da også  $a=[i]$  er en enestående forbindelse, men her spillede det ind at *a* generelt kan forbindes med mange forskellige vokallyde (en fonetisk lighed). Også i ordet *loggia* ['ladça] kunne man have argumenteret for en dobbelt kompleks forbindelse  $ggi=[d\mathfrak{c}]$ , men her foretrak jeg igen de simplest mulige forbindelser  $g=[d]$  (der dermed fik endnu en forekomst) og  $gi=[\mathfrak{c}]$  (hvis lighed med  $g=[\mathfrak{c}]$  i ord som *geni* [çe'ni<sup>?</sup>] ville blive tilsløret af en dobbelt kompleks forbindelse).

Ordene *chat* [tjad], *ciao* [tjaw] og *tchader* ['tjadΛ] (en person fra Tchad) satte mig i en lignende katepine. I det første ord valgte jeg at tillade en dobbelt kompleks forbindelse  $ch=[tj]$  for at undgå de to sære forbindelser  $c=[t]$  og  $h=[j]$ , og fordi *h* er et velkendt sidstebogstav i bogstavenheder som *ch* i *chablis*, *sh* i *sheik*, *th* i *thriller* og *i* i interjektioner som *ih* og *åh*. Forbindelsen  $c=[t]$  blev dog alligevel nødvendig i et enkelt ord, nemlig *ciao*, hvor  $i=[j]$  ikke var til at komme uden om. Ganske vist er der her tale om en forbindelse af vokalbogstav og konsonantlyd, men der er en

oplagt fonetisk lighed med  $i=[i]$ , og  $i=[j]$  forekommer desuden i flere andre ord som *radio* og *bankier*. I ordet *tchader* var forbindelsen  $t=[t]$  ikke til at komme uden om – hvilket så nødvendiggjorde forbindelsen  $ch=[j]$ . Med tilladelsen af  $ch=[j]$  kunne man så genoverveje afvisningen af  $b=[j]$  i ord som *chat* (Becker-Christensen 1988 opstiller faktisk denne forbindelse), men jeg fastholdt analysen med  $ch=[tj]$ .

Morfemgrænser (princip nr. 7) kan som nævnt ikke altid respekteres hvis man vil undgå sjældent anvendte enheder (princip nr. 5). I ord som *bore* ['bo:Λ] og *gære* ['gæ:Λ] valgte jeg den finale forbindelse  $re=[Λ]$  der kunne bruges i begge ord (og mange andre), frem for at indføre de marginale forbindelser  $or=[o:]$  og  $ær=[ε:]$ , selv om disse ville respektere morfemgrænserne (*bor+e* og *gær+e*). På lignende vis indførte jeg  $rde=[Λ]$  for at redegøre for ord som *borde* ['bo:Λ] og *gærde* ['gæ:Λ], frem for at tillade  $ord=[o:]$  og  $ærd=[ε:]$ . Man kunne måske så forvente at ord som *såre* ['sø:ɐ] og *gårde* ['gø:ɐ] ville få en tilsvarende behandling med indførelse af reglerne  $re=[ɐ]$  og  $rde=[ɐ]$ . Men her var reglerne  $år=[ɐ:]$  og  $ård=[ɐ:]$  i forvejen nødvendige i sammensætninger som *sårbar* ['sø:baʔ] og *gårdbo* ['gø:boʔ], og morfemgrænserne i *sår+e* og *gård+e* kunne derfor respekteres:  $s=[s]$   $år=[ɐ:]$   $e=[ɐ]$  og  $g=[g]$   $ård=[ɐ:]$   $e=[ɐ]$ . Det er ikke ganske ligetil at alignere bogstaver og lyde i danske ord konsekvent! Som det er fremgået, er det dog til dels nogle meget lavfrekvente låneord (som *magyar* og *loggia*) der volder problemer.

## DE HYPPIGSTE ENHEDER OG FORBINDELSER

Som nævnt skulle der 565 forskellige bogstav-lyd-forbindelsesregler til for at samtlige ord på min ordliste kunne aligneres restløst. At to tredjedele af forbindelserne involverede komplekse enheder, bekræfter umiddelbart den almindelige opfattelse at dansk ortografi er præget af mange afvigelser fra det alfabetiske princip. De komplekse enheder findes dog næsten udelukkende på bogstavsidens (fx *nn* og *ai* i *mayonnaise*, der kun svarer til en enkelt lyd hver). Der var kun få eksempler på komplekse lydenheder (typen  $j=[dj]$  i *job*).

I alt blev der opstillet 234 forskellige bogstavenheder. Af disse 234 var der kun 33 simple bogstavenheder (14 %). (De 33 svarer til alfabetets 29 små bogstaver plus versalerne *D* og *I* (hvis forekomst var begrænset til de personlige pronominer *De*, *Dem*, *Deres* og *I*) samt to bogstaver med accenter (*é* og *î*)).

På lydsiden var der 109 forskellige enheder, altså langt færre end på bogstavside, og blandt disse var der hele 85 simple enheder (78 %). Det kan desuden bemærkes at hvis man ikke skelner mellem varianter med og uden stød, reduceres antallet af forskellige lydenheder til 84, heraf 63 simple enheder (75 %). Hvis man heller ikke skelner mellem lange og korte vokaler, og altså alene fokuserer på kvalitativt forskellige lydenheder, reduceres antallet yderligere til 59, heraf 43 simple enheder (73 %).

Dette betyder at komplekse enheder primært er et problem når man læser og skal finde ud af hvilke bogstaver der hører sammen, og hvilke bogstaver der udgør selvstændige enheder. Når man staver, har man derimod ikke et tilsvarende segmenteringsproblem. Man kan næsten altid gå ud fra at hver eneste enkeltlyd udgør en selvstændig enhed – hvis stavemåde det dog så på grund af inkonsistente forbindelser stadig kan være en vanskelig opgave at finde frem til!

Hvor stort et problem de komplekse bogstavenheder *reelt* udgør, må naturligvis afhænge af hvor tit man faktisk møder dem når man læser danske ord. Nok var det nødvendigt at opstille en lang række komplekse bogstavenheder, men de fleste af dem forekom kun i forholdsvist få ord. De enheder der forekommer i virkelig mange ord, er alle simple bogstavenheder.

Dette fremgår af tabel 1, som viser de mest anvendte bogstav- og lydenheder i dansk ortografi. Det skal understreges at det der er opgjort her, ikke er enhedernes hyppighed i løbende tekst, men deres hyppighed på min ordliste, altså forbindelsernes leksikonhyppighed. At topscorerne er bogstavenheden *s* med en andel på 13,2 % af alle bogstavenheder og lydenheden [s] med en andel på 13,7 % af alle lydenheder, skyldes givetvis at ordlisten inkluderer bøjningsformer og dermed et stort antal genitiv- og passivformer på *s*/[s].

TABEL 1. DE HYPPIGSTE BOGSTAV- OG LYDENHEDER

Rang	Simple bogstavenheder	Hyppighed *	Komplekse bogstavenheder	Hyppighed *	Simple lydenheder	Hyppighed **
1	s	13,2	er	2,1	[s]	13,7
2	e	9,1	nd	1,0	[d]	6,4
3	t	7,7	ng	0,8	[n]	5,8
4	n	5,8	or	0,5	[ə]	5,6
5	r	5,5	ar	0,5	[l]	5,5
6	l	5,4	ld	0,5	[g]	4,7
7	a	4,7	ig	0,4	[ʌ]	4,7
8	k	4,4	tt	0,3	[b]	3,9
9	i	3,8	kk	0,3	[ɐ]	3,9
10	o	2,9	ll	0,2	[ð]	2,7
11	d	2,9	pp	0,2	[a]	2,4
12	g	2,5	mm	0,2	[m]	2,3
13	u	2,5	ss	0,2	[e]	2,3
14	m	2,5	re	0,2	[f]	2,2
15	b	2,5	dt	0,2	[k]	2,0
16	v	2,4	år	0,1	[a]	1,9
17	p	2,3	ds	0,1	[t]	1,8
18	f	2,3	ch	0,1	[e]	1,8
19	æ	1,8	rd	0,1	[i]	1,8
20	ø	1,5	hv	0,1	[v]	1,8
21	y	1,2	dd	0,1	[ɛ]	1,3
22	h	1,1	ug	0,1	[u]	1,2
23	j	0,9	ff	0,1	[p]	1,1
24	å	0,6	nn	0,1	[h]	1,1
25	c	0,2	gg	0,1	[w]	1,1
26	w	0,1	rr	0,1	[ɔ]	1,0
27	z	0,1	hj	0,1	[ŋ]	0,8
28	x	<0,1	bb	0,1	[nʲ]	0,8
29	é	<0,1	lv	0,1	[lʲ]	0,7
30	q	<0,1	ck	0,1	[-j] #	0,6
31	D	<0,1	sj	0,1	[j-] #	0,6
32	ü	<0,1	ag	0,1	[ø]	0,6
33	I	<0,1	sh	0,1	[y]	0,5

\* Procent af alle bogstavenheder. \*\* Procent af alle lydenheder. # [-j] og [j-] repræsenterer hhv. postvokalisk og prævokalisk j-lyd.



Tabellen viser top-33 for fuldstændighedens skyld: Samtlige simple bogstavenheder er taget med. Nogle af disse er helt marginale såsom versalbogstaverne D og I (der som nævnt kun forekom i nogle få personlige pronominer). Men samlet set er de simple bogstavenheder meget anvendte. Deres andel af forekomsterne udgør tilsammen 90 % af bogstavenhederne på ordlisten.

Det betyder så at de komplekse bogstavenheder kun udgør 10 % – og blandt de hyppigst forekommende er der, som tabel 1 viser, kun 15 der tegner sig for mere end 0,1 % af samtlige bogstavenheder. De komplekse bogstavenheders top-3 er *er*, *nd* og *ng*. Topscoreren *er* forekommer i en lang række nutids- og flertalsformer – men forekomsttallet for *er* rækker stadig kun til en samlet 19.-plads (mellem de simple bogstavenheder *f* og *a*). Selv om komplekse bogstavenheder er et velkendt fænomen i dansk, så er den typiske situation stadig at bogstavenhederne i danske ord er simple – dvs. at de udgøres af enkeltbogstaver. Endvidere kan man formode at læsernes identifikation af de vigtigste komplekse bogstavenheder støttes af en række fællestræk: Tolv af dem er konsonantfordoblinger. Fem andre involverer et *d* (*nd*, *ld*, *rd*, *dt*, *ds*). Fire består af vokal + *r* (*er*, *or*, *ar*, *år*), og andre tre af vokal + *g* (*ig*, *ug*, *ag*).

Ser man på de hyppigste lydenheder, er det påfaldende at listen domineres af konsonantlyde.

Den hyppigste vokal er ikke overraskende svagtryksvokalen [ə], mens den næsthypigste er [ʌ] der i DDO's lydskrift både kan være en fuldvokal og en svagtryksvokal – jf. første hhv. sidste stavelse i ordet *hopper* [ˈhʌbəl]. Alle vokalerne på listen er korte og uden stød – stød forekommer kun i konsonanterne [nʔ] og [lʔ], der står ret langt nede på listen. Komplekse lydenheder var så sjældne at de er ladet ude af betragtning i tabellen. Den hyppigste var [dj] med en andel på 0,1 % af alle lydenheder.

Hvordan ser billedet så ud når bogstav- og lydenheder kobles sammen i bogstav-lyd-forbindelser? Her er svaret – ikke overraskende i lyset af det foregående – at de hyppigst forekommende forbindelser er simple forbindelser. Dette fremgår af tabel 2, hvor de eneste tre komplekse

forbindelser er  $er=[\Lambda]$ ,  $ng=[\eta]$  og  $nd=[n]$  – de to sidste er endda kun lige akkurat med på listen.

Ser man på bogstavernes lydværdier, er der ganske mange forekomster af såkaldt betingede udtaler (se Elbro 2006) både når det gælder konsonanter og vokaler (fx  $t=[d]$ ,  $k=[g]$ ,  $p=[b]$ ,  $\text{t}=[\delta]$ ,  $\text{d}=[\delta]$ ;  $i=[e]$ ,  $u=[\text{ɔ}]$ ). Der er dog også mange „børnevenlige“ forbindelser hvor bogstaverne har deres standardlyd (fx  $t=[t]$ ,  $k=[k]$ ,  $p=[p]$ ,  $s=[s]$ ,  $m=[m]$ ;  $i=[i]$ ,  $u=[u]$ ). Det er desuden bemærkelsesværdigt at fænomenerne vokallængde og stød slet ikke forekommer i listens forbindelser.

Listens 33 forbindelser udgør kun små 6 % af de 565 forskellige bogstav-lyd-forbindelser der blev opstillet for at dække hele ordlisten. Men ser man på antal forekomster, udgør de 33 forbindelser hele 77 % af alle ordlistens bogstav-lyd-forbindelser. Det siger vel næsten sig selv at begynderlæsere og -staver kan have et enormt udbytte af at lære disse forholdsvis få forbindelser at kende.

TABEL 2. DE HYPPIGSTE BOGSTAV-LYD-FORBINDELSER

Rang	Bogstavenhed	Lydenhed	Forbindelsens hyppighed *
1	s	[s]	13,2
2	e	[ə]	5,6
3	n	[n]	5,2
4	l	[l]	5,0
5	t	[d]	4,7
6	r	[ʁ]	3,9
7	k	[g]	2,6
8	b	[b]	2,5
9	f	[f]	2,2
10	m	[m]	2,1
11	er	[ʌ]	2,1
12	a	[a]	1,9
13	k	[k]	1,8
14	a	[a]	1,8
15	t	[t]	1,8
16	g	[g]	1,6
17	v	[v]	1,6
18	i	[i]	1,5
19	d	[ð]	1,4
20	o	[ʌ]	1,3
21	t	[ð]	1,3
22	d	[d]	1,3
23	i	[e]	1,3
24	p	[b]	1,2
25	u	[u]	1,1
26	r	[ʁ]	1,1
27	e	[e]	1,1
28	p	[p]	1,1
29	h	[h]	1,1
30	æ	[e]	0,9
31	u	[ə]	0,8
32	ng	[ŋ]	0,6
33	nd	[n]	0,5

\* Procent af alle bogstav-lyd-forbindelser i den undersøgte ordliste.

## DISKUSSION

For at forbinde bogstaver og lyde i danske ord restløst med hinanden må man ifølge den undersøgelse der er rapporteret her, opstille 565 forskellige bogstav-lyd-forbindelser, hvoraf to tredjedele er „komplekse forbindelser“ idet de involverer det jeg har kaldt „komplekse enheder“. Tallene viser at komplekse bogstav-lyd-forbindelser ikke er et problem man kommer uden om i dansk – og som nævnt i indledningen er der empiriske belæg for at de faktisk udgør et problem for begynderlæsere og -staver. Det er derfor vigtigt at dansklærere kender til problemet og er opmærksomme på de vanskeligheder det kan give at finde ud af hvilke bogstaver og lyde der hører sammen.

Undersøgelsen har dog også vist at komplekse forbindelser gennemgående har en ret begrænset „arbejdsbyrde“. Af samtlige forekomster på den undersøgte ordliste med godt 47.000 en- og tostavelsesord var kun hver tiende bogstav-lyd-forbindelse en kompleks forbindelse.

Hovedkonklusionen må derfor være at princippet om én lyd til ét bogstav stadig er bærende i dansk ortografi. Dette kan understøttes af en nærmere analyse af de enkelte ordformer på min ordliste: Halvdelen af dem (50,8 %) består udelukkende af simple bogstav-lyd-forbindelser. Og kun 1,3 % af listens ord har *meget* komplekse forbindelser, forstået som forbindelser der omfatter enheder med flere end to bogstaver/lyde (fx *age*=[æ<sup>ɹ</sup>] og *ager*=[a<sup>ɹ</sup>] i ordene *tage* og *tager*) samt de omtalte „dobbelt komplekse“ forbindelser (som *gn*=[nj] i *vignet*).

Et problem jeg ikke er kommet ind på i denne artikel, er bogstav-lyd-forbindelsernes konsistens. Også på dette område har min undersøgelse sigtet på at sætte tal på, nemlig ved at beregne såkaldte konsistenskoeficienter som et mål på for hvor svært det er at gætte bestemte bogstavenheders udtalemuligheder, hhv. lydenheders stave- muligheder (Juul 2008). Sammen med en kollega er jeg for tiden ved at udvikle et netsted der skal formidle resultaterne fra denne del af min undersøgelse. Hold øje med adressen [www.bogstavlyd.ku.dk](http://www.bogstavlyd.ku.dk).

Det portræt der her er tegnet af bogstav-lyd-forbindelserne i dansk, er selvfølgelig præget af undersøgelsens metodemæssige forudsætninger. Hvis den undersøgte ordliste havde omfattet ord med flere end to stavelser, var der uden tvivl blevet brug for at opstille endnu flere bogstav-lyd-forbindelser – og hvis ordlisten havde omfattet *proprier*, ville der blive brug for *langt* flere, eftersom *propriers* stavemåder er et kapitel helt for sig. Det har også haft betydning at Den Danske Ordbog, som undersøgelsen bygger på, gengiver en forholdsvis konservativ og distinkt udtale (Kristensen 2003). DDO har fx ordet *bræt* med udtalen [ˈbræt] skønt den hos mange er faldet sammen med *brat* [ˈbræt]. Og i lydskriften bruger DDO af hensyn til læsbarheden ikke syllabiske konsonanter – ordet *cykel* lydskrives således [ˈsygəl] og ikke [ˈsyg!]. Syllabiske konsonanter ville have gjort komplekse bogstavenheder langt hyppigere.

Med hensyn til aligeringen af bogstaver og lyde var det måske mest afgørende valg at jeg ikke tillod stumme bogstaver, hvilket var stærkt medvirkende til at øge antallet af bogstav-lyd-forbindelser. Formuleringen og anvendelsen af faste principper viste sig at være en vanskelig sag, hvor mange valg givetvis kan diskuteres. Det jeg har præsenteret her, er blot de fremgangsmåder jeg har fundet rigtigst (eller mest gennemførlige) i forhold til de spørgsmål jeg gerne ville svare på.

Hvis man er uenig i mine valg, håber jeg ikke desto mindre at have skabt opmærksomhed om de spørgsmål man må forholde sig til i en kvantitativ beskrivelse af dansk ortografi. Min vigtigste pointe er nemlig at sådanne beskrivelser er et nyttigt grundlag – både hvis man vil forstå skoleelevers vanskeligheder med at læse og stave ord med komplekse bogstavenheder, og hvis man vil argumentere for (eller eventuelt imod) forenklede stavemåder som *majonæse*.

Holger Juul

Institut for Nordiske Studier og Sprogvidenskab

Københavns Universitet

juul@hum.ku.dk

## NOTER

- 1 Jeg ser i aligneringen bort fra tryk, som er en egenskab ved stavelser snarere end en egenskab ved enkelte lyde. Det samme kan man mene om længde og stød (jf. Grønnum 2005), men dette er mindre indlysende eftersom det er langt mere udbredt at tale om længde og stød som en egenskab ved vokaler (og for støds vedkommende nogle gange konsonanter) end som en egenskab ved stavelser. Jeg ser derfor *ikke* bort fra stød og længde.
- 2 Ordlisten blev opbygget i forbindelse med et postdoc-projekt finansieret af FKK (Det Frie Forskningsråd, Kultur og Kommunikation).
- 3 Det skal bemærkes at den trykte udgave af DDO bruger lydskriften Dania, mens jeg her i artiklen har valgt at følge internetversionen af DDO, der benytter en IPA-baseret lydskrift. I IPA-lydskriften repræsenteres både præ- og postvokalisk *j*-lyd (der skelnes i Dania) af symbolet [j] – men skellet mellem de to varianter er opretholdt i de analyser der rapporteres her, jf. tabel 1.

## LITTERATUR

- Basbøll, H. (2005) *The phonology of Danish*. Oxford: Oxford University Press.
- Becker-Christensen, C. (1988). *Bogstav og Lyd*. Dansk retskrivning og rigsmålsudtale. København: Gyldendal.
- Carney, E. (1994) *A Survey of English Spelling*. London: Routledge.
- Elbro, C. (2004) "Literacy acquisition in Danish: A deep orthography in cross-linguistic light". R.M. Joshi & P.G. Aaron (red.) *Handbook of Orthography and Literacy*. Mahwah, New Jersey: Lawrence Erlbaum. 31-45.
- Elbro, C. (2006) *Læsning og læseundervisning* (2. udg.). København: Gyldendal.
- Grønnum, N. (2005) *Fonetik og fonologi*. Almen og dansk (3. udg.). København: Akademisk Forlag.
- Hjorth & Kristensen (red.) (2003-2005) *Den Danske Ordbog*. København: Det Danske Sprog- og Litteraturselskab, Gyldendal.
- Jacobsen, H.G. (2010) *Ret og Skrift. Officiel dansk retskrivning 1739-2005. Bind 1*. Odense: Syddansk Universitetsforlag.
- Juul, H. (2008) "Er danske bogstaver til at stole på?" *Magasinet Humaniora*, 23. årgang, nr. 4. 22-27.
- Juul, H., Elbro, C., Aro, M., Forsman, C., Lund, R., Sigurðsson, B., Þráinsdóttir, A., Seymour, P.H.K., Duncan, L.G. (2007) *Early reading development in three Nordic countries and the UK*. Report to the Joint Committee for Nordic Research Councils for the Humanities and the Social Sciences. 32 sider.
- Juul, H. & B. Sigurðsson (2005) "Orthography as a handicap? A direct comparison of spelling acquisition in Danish and Icelandic". *Scandinavian Journal of Psychology* 46. 263-272.
- Kessler, B. & Treiman, R. (2001) "Relationships between sounds and letters in English monosyllables". *Journal of Memory and Language* 44. 592-617.
- Kristensen, K. (2003) "En lille test af lydskrift." H.G. Jacobsen, D. Bleses, T.O. Madsen & P. Thomsen (red.) *Take Danish - for instance. Linguistic studies in honour of Hans Basbøll presented on the occasion of his 60th birthday 12 July 2003*. Odense: University Press of Southern Denmark. 181-187.
- Reece, C. & Treiman, R. (2001) "Children's spelling of syllabic /ɾ/ and of letter-name vowels: Broadening the study of spelling development". *Applied Psycholinguistics* 22. 139-165.
- Spencer, K.A. (under udg.) "Predicting children's word-reading difficulty for common English words: The effect of complexity and transparency". *British Journal of Psychology*.

Ziegler, J., & Goswami, U. (2005) "Reading Acquisition, Developmental Dyslexia, and Skilled Reading Across Languages: A Psycholinguistic Grain Size Theory". *Psychological Bulletin* 131(1). 3-29.