

NyS

Titel:	Komplekse relationelle enheder i dansk retskrivning
Forfatter:	Christian Becker-Christensen
Kilde:	<i>NyS – Nydanske Sprogstudier</i> 43, 2012, s. 10-42
Udgivet af:	NyS i samarbejde med Dansk Sprognævn
URL:	www.nys.dk



© 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.

Komplekse relationelle enheder i dansk retskrivning

CHRISTIAN BECKER-CHRISTENSEN

1 KOMPLEKSITETER

Lydrethed, forstået som enkel omsættelighed mellem lyd (fonem eller fon) og bogstav, er det bærende princip i alfabetiske skriftsystemer. Dansk retskrivning hører som bekendt her til de mere komplicerede af de europæiske ortografier med en "somewhat deep orthography" (Elbro 2006: 32), sammenlignelig med engelsk og forskellig fra mere lydrette ortografier som vore nordiske nabosprogs og fx tysk og spansk.

Et hjørne af kompleksiteten i retskrivningen vedrører relationer med komplekse enheder, nogle hjemlige som <hv> [v] i *hval* og <ds> [s] i *tudse* og andre fremmede som <ea> [i] i *team*, heriblandt komplekse lydenheder som [dj] <j> i *job*. Sådanne forbindelser kan være medvirkende til relationel uforudsigelighed, og de kan også i sig selv udgøre særlige vanskeligheder for begynderlæseren (Elbro 2002, 2006, Juul & Sigurdsson 2004). Derudover har de en rolle som beskrivelsesmæssigt problem for segmentering og inventaropstilling i relationelle beskrivelser, både når det drejer sig om formaliserede relateringer og om mere almene fremstillinger af dansk retskrivning.

Artiklen her følger tre spor i tilgangen til komplekse relationelle enheder: grafemisk afgrænsning af dem i ortografiske beskrivelser, deres funktion i retskrivningen og den status de har i retskrivningen, som denne fremgår af retskrivningsordbøgerne RO55 til RO12. Fremstillingen drejer sig hovedsagelig om komplekse bogstavenheder eller grafemenheder som de kaldes i de følgende afsnit. Grundlag for artiklen er korpusset og søgefunktionen Bogstavlyd (bogstavlyd.ku.dk) og Juuls artikel "Kat-te-pi-ner" i NyS 39 (Juul 2010) hvori han redegør for etableringen af bogstav-lyd-relationerne i Bogstavlyd og for komplekse enheders andel deri. Juuls relatering er baseret på restløs segmentering, dvs. på en segmenteringsprocedure som udelukker stumme bogstaver, og andelen af komplekse bogstavenheder er derfor høj, 201

digrafemer, trigrafemer m.m., set i forhold til det samlede antal af 234 bogstavenheder. Andre former for relateringer er sådanne som regner med nul-relationer og betingede regler og dermed mulighed for et mere begrænset inventar i sættet af komplekse enheder. Sådanne relateringer forekommer for engelsk hos bl.a. Carney (1994) hvis omfattende beskrivelse omfatter ”speech-to-text” og ”text-to-speech correspondences”, og i Venezkys (1970, 1999) morfofonemisk relaterede beskrivelse af skrift til lyd-forholdet. De to former for relatering, relatering baseret på restløs segmentering og relatering med inddragelse af nul-relationer, kan supplere hinanden som to sider af samme sag. Det er hensigtsmæssigt at kunne tale om at <hv> udtales [v], såvel som at <h>’et er stumt foran <v>. I forhold til en restløs segmentering kan inddragelse af nul-relatering imidlertid bidrage til mulighed for en mere generel beskrivelse og dermed også til forståelse af nogle mere overordnede træk i en ortografi.

Artiklen falder i disse hoveddele: I afsnit 3 omtaler jeg relateringen i Bogstavlyd. I afsnit 4 redegør jeg for forskellige typer af relationer og relationelle enheder. Afsnit 5 handler så om beskrivelser med andre former for relatering end i Bogstavlyd og om komplekse enheders position i ortografiske beskrivelser. Afsnit 6 omhandler komplekse enheders funktion. Og i afsnit 7 omtales regulering og tilkomst af komplekse grafemenheder i de senere retskrivningsordbøger.

Synsvinklen i artiklen er afgrænset til rent lydlig relatering (og ikke morfofonemisk relatering); der inddrages således ikke andre internt bestemmende forhold i retskrivningen såsom tegnkonstansprincippet eller det distinktive princip (jf. Ågerup Jervelund 2007 om disse og flere). Et kort indledende terminologisk overblik gives her i afsnit 2.

2 RELATIONELLE ENHEDER OG ANDEN TERMINOLOGI

De enheder der relateres til hinanden når man beskriver talesproget i forhold til bogstavskriften, kan man betegne som (simple og komplekse) *lydenheder* og (simple og komplekse) *bogstavenheder* (Juul 2010: 11) eller sammenfattende som *korrespondensenheder* (”units of correspondence”, Carney 1994: 32 ff.) eller relationelle enheder (Becker-Christensen 1988: 28 f.; jf. Venezky 1970: 50 ff., 1999: 78 ff., hos hvem ”relational units” dog kun bruges om enheder på grafemsiden). En

relationel enhed er således en enhed i ét plan som modsvarer en eller flere enheder (eller nul) i et andet plan – idet en enhed kan være et enkeltsegment, en dertil svarende sekvens eller en minimalt afgrænset sekvens der modsvarer en anden sekvens (se typerne 1-4 i afsnit 4.1). Selve betegnelsen *enbed* står for størrelser som har inventariel status i et relationelt system. Relationelle enheder aligneres (Juul 2010: 10), eller de korresponderer med eller relateres til hinanden. Enheder i ét plan repræsenteres ved enheder i det andet plan – eller blot: De udtales eller skrives som sådanne enheder. Betegnelsen *kompleks relationel enbed* veksler i artiklen med den mindre tunge betegnelse *kompleks enbed*. Efter den måde kombinationer som <hv> anskues på, angives de dels som komplekse enheder, dels som sekvenser.

Til betegnelse af forholdet mellem lyd og skrift er der termer som *fonografemik* (og *ortografi*) og *grafofonemik* (og *ortofoni*) (Allén 1965: 40-41; 1971: 8), men ikke nogen overordnet term for den gensidige beskrivelse af forholdet. Ortografi (og retskrivning) bruges jo imidlertid ikke blot om retstavning men også mere bredt i betydningen skriftsystem og kan i denne brug involvere både fonografemisk og grafofon(em)isk relatering. Lyd til skrift- og skrift til lyd-forholdet mellem relationelle enheder kan vises ved rækkefølgen, henholdsvis fx [v] <hv> og <hv> [v] (hos Carney (1994) med ”≡” imellem de to sider, hos Juul ”=”). I det følgende bruges begge rækkefølger dog også blot umarkeret med bidirektional betydning. Tydeliggørelse af unidirektional og bidirektional relation vises med ”=>”, ”<=” og ”<=>”.

Frem for betegnelserne bogstav og lyd bruges i det følgende hovedsagelig betegnelserne grafem¹ og fon (som står for lydskrifttegn i en fonetisk lydskrift) og fonem. Vokal og konsonant bruges om både foner (og fonemer) og grafemer (selv om de korrekte betegnelser på grafemsiden vil være autografem og syngrafem, jf. Diderichsen 1952: 10, Allén 1971: 29-30). I stedet for grafofon(em)isk skrives blot grafofonemisk som dækkende begge betydninger. For nul-korrespondenser skrives [Ø] i den fonetiske repræsentation. Lodret streg ”|” adskiller segmenterede størrelser i et ord, men står i citeret tekst også som skilletegn foran kontekstbetingelser i genskrivningsregler.

3 BOGSTAVLYD: ENHEDER I RESTLØS SEGMENTERING

Grundlæggende for mere formaliserede sammenstillinger af skrift og lyd er afgrænsningen af hvilke størrelser der relateres. Bestemmelsen heraf kan være influeret af beskrivelsesretningen, om den er grafofonemisk, fonografemisk eller bidirektionel.

Netstedet, eller som det benævnes på sitet ”søgefunktionen”, Bogstavlyd repræsenterer en gensidig relatering mellem skrift og lyd med benyttelse af samme sæt lydenheder og samme sæt grafemenheder i begge søgeretninger. Grundlaget for relateringen er en restløs, lineær segmentering, dvs. en segmentering uden relationer med nul og relationer der krydser hinanden. Korpuset består af et materiale på 47.000 ordformer (en- og tostavellesformer) baseret på udtræk af Den Danske Ordbog og med samme udtaleangivelser som i denne. Trods den stavelsesmæssige begrænsning giver Bogstavlyd ud fra de anlagte segmenteringsprincipper et repræsentativt billede af relationer og relationelle enheder i dansk ortografi (Juul 2010 nævner at tre fjerdedele af ordene i Det Danske Sprog- og Litteraturselskabs Korpus 2000 udgøres af en- og tostavellesord).

Juul (2010) beskriver principperne for segmenteringen som ligger til grund for Bogstavlyd. Denne er søgt gjort så enkel og konsistent som muligt. Det overordnede princip er at der ikke accepteres relationer med nul: ”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” (ibid.: princip 1). Således er fx <hv> en enhed i *hval*: ”hv:[v] a:[æʔ] l:[l]”. Princippet har sin begrundelse i beskrivelsens dobbeltsidighed, nemlig at ”konsistensberegninger i læseretningen og staveretningen [bliver] sammenlignelige, fordi de kan tage udgangspunkt i de samme bogstav-lyd-forbindelser” (ibid.: 14). En yderligere konsekvens af dette princip er at ”Bogstaver og lyde skal forbindes direkte” (princip 4), dvs. uden mellemed i relateringen.

Et andet forhold er at segmenteringen er lineært fremadskridende fra venstre mod højre og ikke må give relationer der krydser hinanden (princip 3). Således analyseres *case* [kejs] som ”c:[k] a:[ɛj] se:[s]” og ikke med [ɛj] <a..e> (Juul henviser til bl.a. Carney 1994 for en sådan opsplittet analyse). Princippet betyder også at <gn> i fx *vignet* [vin'jed] analyseres som en enhed og ikke med særsegmentalt krydsende relatio-

ner <n> [n] og <g> [j]; muligheden af en relatering <g> [n] og <n> [j] afvises ”ud fra den betragtning at g=[n] kun vil forekomme før n=[j] og vice versa”, hvilket også skal ses i sammenhæng med princip 5: ”Bogstav-lyd-forbindelser med begrænset anvendelse i andre ordformer skal undgås”.

Der er i Bogstavlyd 565 gensidige skrift-lyd-relationer med i alt 234 forskellige grafemenheder hvoraf 33 (inklusive D, I, é, ü) er simple, og 201 er komplekse (Juul 2010: 19, 23). De mange relationer må ud over princippet om restløs segmentering tilskrives omfanget af lydinventaret, 109 enheder, hvor korte vokaler, lange stødløse vokaler, vokaler med stød og konsonanter med og uden stød er regnet som selvstændige enheder. Hvis man ser bort fra stød og længde, reduceres antallet af lydenheder til 59 (ibid.: 23) og de 565 relationer til 386.

De mere hyppige af de 201 komplekse grafemenheder i Bogstavlyd, enheder med relationer til mere end 10 ordformer, fordeler sig i stigende orden sådan her:

36 komplekse enheder svarende til [V] med fra 11 til 5724 forekomster:

<ad, org, uv, od, au, eau, al, ed, et, ive, av, ey, iv, aw, ej, ai, ur, ård, rde, arr, ie, og, ee, yg, ou, oo, rer, ea, ag, ug, år, re, ig, ar, or, er>

50 komplekse enheder svarende til [C] med fra 11 til 2791 forekomster:

<qu, ne, pe, rg, ci, se, ssi, wh, zz, ke, sc, ce, th, ge, kh, nt, le, te, gh, me, si, sch, ti, nk, lg, sh, sj, ck, rr, lv, bb, hj, gg, nn, ff, dd, hv, rd, ch, ds, dt, ss, ll, mm, pp, kk, tt, ld, ng, nd>

Eksempler fra Bogstavlyd med [i], [u], [s] og [ç] relateret til disse enheder (der er set bort fra længde og stød):

[i] <ea> a|pp|ea|l, <ee> j|ee|p, <ie> b|r|ie|f|e, <ey> h|o|ck|ey,
<ig> p|ig|e, <iv> b|l|iv

[u] <oo> b|oo|k|e, <ou> c|l|ou, <ug> f|ug|l

[b] <bb> k|r|a|bb|e, <pe> t|a|pe, <pp> k|l|i|pp|e

[ç] <sch>sch|æ|f|er, <sh> sh|e|rr|y, <sj> sj|æ|l,

<si> v|i|si|o|n, <ssi> m|i|ssi|o|n, <ti> s|t|a|ti|o|n

Ordformerne i Bogstavlyd omfatter sammensætninger og bøjningsformer, hvilket bidrager til en forholdsvis høj score for nogle enheder. Fx består de 32 forekomster for <ej> [ɛʔ, ɛ] af ordet *vej*r med former som *vej*r, *regnvej*r, *vejrgud*. Noget lignende gælder relationer som fx <ag> [a] som i ugedagene plus *middag* sammen med genitivformer er oppe på 16 forekomster, mens de 128 forekomster af <ag> [æʔ] er repræsenteret af *bag*, *dag*, *fag*, *lag*, *sag*, *smag*, *svag*, *tag* og sammensætninger med disse som andetled samt bøjningsformer.

4 TYPER AF RELATIONELLE ENHEDER

4.1 Fire slags relationer

Beskrivelse af forholdet mellem lyd og skrift kan overordnet ses som fremstilling af hvordan enheder i ét system (eller medie) fremtræder i et andet system (eller medie). For lydrelaterende beskrivelser omfatter dette også sammenhængen mellem den auditivt eller akustisk baserede segmentering og den transskriberede fon (lydskrifttegnet) og mellem lydskrifttegn og fonemer. Forholdet mellem et system (A) og et system (B) kan betragtes på den måde at (B) er afbildende og tolkende, eller sådan at (B) i højere grad er regulerbar i forhold til (A) end omvendt. Som beskrivelse kan relationen ses begge veje: hvordan (A) repræsenteres i (B), hvordan (B) repræsenteres i (A), eller hvordan det gensidige forhold er mellem (A) og (B). Ved sammenstillingen af lyd (A) og skrift (B) er der tale om disse fire relationer mellem enkeltsegmenter og sekvenser, her og i det følgende med Bogstavlyds segmentering:

	(A)	(B)	
1	enkelt	<=> enkelt	[p] <p> [po:sə] pose
2	enkelt	<=> sekvens	[v] <hv>, [ɛ] <sɿ> [væʔl] hval, [ɛɛʔl] sjæl
3	sekvens	<=> enkelt	[gs] <x> [tagsa] taxa
4	sekvens	<=> sekvens	[nj] <gn> [kɒnʔjag] cognac

En relationel enhed kan indgå i forskellige typer. Fx hører [ɛ] til type 1 i *gene* men til type 2 i *sjæl*, og <x> hører til type 1 i *xylofon* men til type 3 i *taxa*, ligesom <ch> hører til type 2 i *chianti* men til type 4 i *chili*. Dertil kan det der i en sammenhæng er en kompleks relationel enhed,

i andre sammenhænge være en sekvens med særsegmentale relationer, fx <ea> i *jeans* og *seance* (henholdsvis [i:] og [ea]) og <ur> i *surfe* og *spurte* (henholdsvis [æ:] og [uɐ]). I nogle tilfælde går enhedens relation igen i en af sekvensens særsegmentale relationer, fx <ng> [ŋ] i *bange* og <n> [ŋg] i *tango*.

I Bogstavlyd udgør simple grafemenheder (type 1 og 3) 90 % af forekomsten af grafemenheder mens komplekse grafemenheder (type 2 og 4) udgør de 10 % (Juul 2010: 25).

Type 1 er defaultrelationen. Den omfatter alle alfabetets bogstaver og er enerådende i halvdelen af ordformerne i Bogstavlyd, dvs. i ord som *ø, kø, stå, bog, fest, soul, måne, nabo, menu, strejfe, haiku, advent, respons* (Juul 2010: 28). Enkelte foner er udelukkende eller i langt overvejende grad type 1-relaterede: [h] som kun korresponderer med <h> (*have*), og [a] og [ø] der korresponderer med henholdsvis <a, e, æ> (*mand, dreng, skræk*) og <y, ø> (*keys, love*) og blot i få ord med sekvenser (bl.a. i *nougat* og *monsieur*). Af de fremmede grafemer <c, q, w, x, z> er <q> i Bogstavlyd bundet til <u>, fx *quiz*, og forekommer således ikke som type 1-enhed (men RO01 har *Qatar* og Nudansk Ordbog (2010) de grønlandske byer *Qaanaaq* og *Qaqortoq*).

Færreste relationer har typerne 3 og 4 hvis forekomster er begrænset til fremmedord. Ud over [gs] <x> i fx *taxa* og [ds] <z> i *trapez* forekommer **type 3** i engelske lån med bl.a. [ej, aj, ɔw, ju, dj] <a, i, o, u, j>, fx *baby, site, dope, unit* og *jungle*. I Bogstavlyd indgår 10 forskellige grafemenheder i denne relationstype. I type 4 kan der etableres segmental forbindelse mellem (A)- og (B)-sekvenserne, men forholdet mellem dem er tolket sådan at det ikke med rimelighed kan opløses i relationer med enkeltsegmenter (Juuls princip 3 og 5, se afsnit 3). Der er kun enkelte relationer hver især med få forekomster af denne type, bl.a. [nj] <gn> i fx *cognac*, [əl] <le> i fx *hurdle* og [tj] <ch> i fx [tjili] *chili*. I Bogstavlyd er 5 grafemenheder involveret i denne relationstype.

Den større mellemgruppe, **type 2**, relation mellem (A) enkeltsegment og (B) sekvens af segmenter, omfatter i Bogstavlyd 186 grafemenheder. Fonen [ɕ] hører næsten udelukkende til i denne gruppe med korrespondenser som <sj, ch> (*sjæl, charme*, men også <g, j> i fx *gene, jaloux*). Omvendt hører også sekvenser som <sj, hv> (i monosyllabisk sammenhæng) kun til her. Det er hovedsagelig type 2-relationer som

fortsættelsen af denne artikel handler om. Det er i denne gruppe at man kan diskutere relevansen af nul-relateringer (afsnit 5), og det er enheder i denne gruppe som i særlig grad er genstand for ortografiske reguleringer (afsnit 7).

4.2 Endo- og exocentriske enheder

Type 2-enhederne <hv> i *hval* [væʔl] og <sj> i *sjæl* [œʔl] er forskellige ved at den første har et stumt bogstav, <h>, hvad den anden ikke har. Med lån fra grammatisk terminologi kan de to slags enheder betegnes som henholdsvis endocentriske og exocentriske (Carney 1994: 40-41, jf. fx Lyons 1968: 231 ff.).

I den endocentriske enhed kan et af segmenterne i sekvensen stå i stedet for enheden. Den exocentriske enhed kan ikke reduceres til ét af segmenterne i sekvensen uden at der sker ændring i enhedens (og hermed ordets) udtale. Exocentriske enheder med tre eller flere segmenter kan eventuelt reduceres til to segmenter, fx <oui> i *bouillon*, <ård> i *gård* og <sch> i *schæfer*, men ikke til ét.

Der er som bekendt to slags stumme bogstaver i dansk: stumt <h> i *hval* og stumt <d> i *kande*, hvor <d> i det sidste tilfælde ”ikke [er] helt stumt” men spiller sammen med ordets udtale som helhed (Hansen 1981: 57); *hval* kunne tænkes skrevet **val*, men *kande* vil ikke kunne skrives **kane*. De to typer stumme bogstaver betegnes hos Carney (1994) som henholdsvis ”empty letters” og ”auxiliary letters”, den sidste betegnelse bruges også om grafemer i exocentriske sekvenser.

Endocentriske enheder som <nd> der på denne måde fungerer eks-ternt som exocentriske, kan (i mangel af bedre) betegnes som endo-exocentriske. Nogle enheder kan således også fungere både rent endocentrisk og blandet endo-exocentrisk; det gælder særlig geminater henholdsvis foran og efter trykstærk vokal, fx <ll> i *ballade* (jf. *balance*) og *balle*.

	endocentrisk		
	internt	eksternt	
[v] <hv> <i>hval</i>	+	+	endocentrisk
[l] <ll> <i>ballade</i>	+	+	
[l] <ll> <i>balle</i>	+	-	endo-exocentrisk
[u:] <ou> <i>velour</i>	+	-	
[ɛ] <sj> <i>sjæl</i>	-	-	exocentrisk
[i:] <ea> <i>lease</i>	-	-	

Svarende til at en endocentrisk enhed kan substitueres med et af dens segmenter, kan en exocentrisk enhed i nogle tilfælde substitueres med en anden enhed, fx udskiftning af <eu> med <ø> (fx *milieu*, nu *miljø*).

Endocentriske konstruktioner underopdeles i sideordnings- og underordningsforbindelser ("co-ordinating" og "subordinating constructions" Lyons 1968: 233), og kategoriseringen i endocentriske og exocentriske konstruktioner kan dermed sidestilles med tredelingen i forholdsarterne paratagme og hypotagme over for neksus (eller kata-tagme) (Diderichsen 1957: 140 f.). I ortografisk sammenhæng er eksempler på de to slags endocentriske konstruktioner henholdsvis <ll> [l] i *ballade* og <hv> [v] i *hval*. Bortset fra *kraal*, *jamboree* og *shampoo* (med [aʷ], [ɛ] og [o]) er endocentrisk parataktiske sekvenser i dansk retskrivning kun konsonantiske og forekommer som rent endocentriske hovedsagelig i typer som *ballade* og *scene*.

Opdelingen i endocentriske, exocentriske og endo-exocentriske enheder gælder (B)-sidens type 2-enheder. Komplekse fonenheder (type 3 og 4) er exocentriske. Den basale inddeling har relevans for den videre grafofonemiske beskrivelse (afsnit 5) og for vurderingen af enhedernes funktionalitet i retskrivningen (afsnit 6).

5 ENHEDER I ANDRE FORMER FOR BESKRIVELSE

5.1 Relatering med nul-korrepondenser

Man kan tænke sig en skrift-lyd-relatering gennemført med udelukkende enkeltsegmentale relationer: type 1-relationer med tilføjelse af relationer med nul-segmenter i henholdsvis (A) og (B). I de segmenterede ord vil der så være tomme pladser, med visning som i Bogstavlyd

(og med Ø for nul) fx: s:[ɛ] j:[Ø] æ:[eʔ] l:[l] og t:[t] a:[a] x:[g] Ø:[s] a [a]. En sådan relatering vil segmenteringsmæssigt være uproblematisk (bortset fra at den ved en automatiseret matchning eller generering vil kræve flere kontekstafhængige regler end en segmentering med komplekse enheder). Men hvis en relatering også skal tjene til beskrivelse af en ortografis brugerfunktionalitet, vil en segmentering gennemført på denne måde ikke være optimal.

I forhold til modpolen, den restløse segmentering, er det dog formålstjenligt at en ortografisk beskrivelse i et vist omfang betjener sig af nul-relationer hvilket ud over begrænsning af inventaret af enheder kan være forudsætning for skildring af mere generelle træk. En omkostning i forhold til den restløse segmentering er at relationsreglerne bliver mere komplekse da en del af konteksten ikke længere er indeholdt i den relationelle enhed, fx <h> [Ø] foran <v> i stedet for blot <hv> [v], og <n> [ŋ] foran <g> og <g> [Ø] efter [ŋ] stavelsesfinalt i stedet for <ng> [ŋ] stavelsesfinalt. Dertil kommer at rammerne for hvornår der kan være tale om opløselige sekvenser i stedet for enheder, er vanskelige at sætte. Juul (2010: 14) anfører det da også som en fordel for den restløse segmentering ”at man undgår at skulle tage stilling til hvornår stumme bogstaver er at foretrække for komplekse bogstavenheder”. Nogle vanskeligheder beror også på at nul-relationer fungerer bedre i en beskrivelsesretning end i en anden. I det følgende diskuteres forskellige teknikaliteter ved indførelsen af nul-relationer i ortografiske beskrivelser, først (i dette afsnit og i 5.2) i grafofonemisk retning.

Ligesom i en fonologisk beskrivelse må man ved relatering af skrift og lyd med betingede regler søge at afveje inventariel økonomi over for regeløkonomi, dvs. forhold som antal regler og regelspecifikationer og også hvor dækkende en regel er. En forskel er ganske vist at i fonologiske beskrivelser fremanalyseres den ene side af relationen, (B), mens denne side er given i ortografiske beskrivelser.

Carney (1994: 35) har for sin beskrivelse det overordnede simplicitetskriterium, hans princip ”#S1”, at ”The number of different correspondences set up to describe the system should be as small as possible”. Ligesom Juuls princip 2, at ”Aligneringen skal give bogstav-lyd-forbindelser der er så simple som muligt”, gælder dette princip konsistent minimal afgrænsning af enheder og dermed minimering af

antallet af regler, men omfatter hos Carney også mulighed for korrespondenser med nul. Hos Venezky (1970: 50-51) er der tale om et generalitetsprincip som afgrænser hvad der kan betragtes som komplekse relationelle enheder: "only those strings whose morphophonemic correspondences cannot be predicted by general rules based upon smaller units contained in the string are classed as relational units. *ch* in *chair*, for example, is a relational unit since the morphophoneme {č} cannot be predicted from general rules based upon *c* and *h* separately."

Men principper for segmentering i en beskrivelse med nul-relationer og betingede regler er vanskelige at operationalisere. Et kriterium som "generelle regler" er selvsagt relativt anvendbart. I forhold til en fonologisk analyse kan man fx ikke vurdere regelkompleksitet trækanalytisk (medmindre regelbetingelserne først er omsat til fonologisk niveau, fx <c> [s] foran fortunevokal under forudsætningen at <i, e, æ, y, ø> er omsat til foner med markering som fortunevokaler). Brugen af begrebet *regel* er desuden ikke entydig (heller ikke i fonologiske beskrivelser). I Bogstavlyd er en regel en unik relation mellem en grafemenhed og en fonenhed. Men med inddragelse af kontekstbetingelser kan samme input og output være fordelt på flere regler når der er tale om forskellige kontekstbeskrivelser, som Carneys regel G.5 og regel G.6 der begge har "<g>≡Ø" men i forskellige kontekster.

Af de 86 opregnede Bogstavlydenheder i afsnit 3 vil de mest oplagte kandidater til beskrivelse med nul-relatering være de **endocentriske** sekvenser <gh, hj, hv kh, wh> hvor <h> i forbindelse med konsonant, bortset fra forbindelserne <ch, sh> (og enkelte forekomster af <th> [ð]), kan relateres til [Ø] (*ghetto*, *hjul*, *bval*, *kban*, *whisky*) hvorved disse sekvenser ikke vil figurere som inventarielle enheder. Mere begrænset kan <c> ses med nul-relation i <sc>, men reglen for dette ene tilfælde er mere kompleks: <c> [Ø] mellem <s> og <e, i> (*scene*, *fascist*). Leksikalsk betinget kan der være tale om <d> [Ø] i <rd> (*bord*) og i <ds> efter lang vokal (*stads*), <t> [Ø] i <et> (*filet* og i de mindre frekvente <it, ot>: *esprit*, *trikot*, desuden i <nt> i fx *restaurant*).

De **endo-exocentriske** forbindelser <ds, dt> efter kort vokal (*kryds*, *krudt*), <ld, nd> (*kulde*, *kande*) samt <lg, lv> (*salg*, *gulv*), fungerer grafofonemisk som geminater, og <d, g, v> kan relateres til nul efter at have trigget forudgående vokals kvantitet. Sekvenserne optræ-

der kun i hjemlige ord, og <d, g, v> er her i forskellig grad morfologisk latente. For <ds> gælder det kun i få tilfælde (*vidste* ~ *vide*), for <dt> i langt de fleste tilfælde (*godt* ~ *god*) og for <lg, lv> i tilfælde med distinkt udtale (*salg* ~ *sælge*, *gulv* ~ *gulve*).

Forbindelserne <ke, me, ne, pe, se, te> kan opløses med <e> som leksikalsk betinget markør for forudgående vokals udtale som diftong eller lang vokal (*joke, lime, deadline, tape, case, flute*), i <ce, ge> (*softice, rouge*) er <e> desuden markør for forudgående konsonantkorrespondens (disse kombinationer er her exocentriske). Der er tale om ord fra engelsk og fransk (ord med "silent <e>"). Både Carney og Venezky har nul-relatering for <e>, Carney med vokalregler som A.27 og O.37 for <a> og <o> i fx *trace* og *note* og reglen E.2 for <e>: "<e>≡Ø| __ #". Alternativet vil med en restløs segmentering være at anføre enhederne <ce, te> etc. i regelspecifikationerne for <a, o>. Hvis man med nul-relatering skal regne med præliminært komplekse enheder i disse ord, er det altså snarere <a..e>, <o..e> etc. som diskontinuerte enheder der drejer sig om.

For den **exocentriske** konsonantforbindelse <ng> (*tange*) kan opløsning motiveres af at <n> har samme betingede relation [ŋ] som når sekvensen optræder med særsegmentale relationer i andre omgivelser (*tango*). Carney har således ikke <ng> som relationel enhed, derimod reglerne N.1 for <n> [ŋ] og G.5 for <g>: "<g>≡Ø|<n> __ {#, <C>}". På en noget lignende måde kan der argumenteres for opløsning af <ar> (*arne*) hvor postvokalisk <r> betinger <a> [a:] ligesom også prævokalisk <r> gør det (*rane*). Derimod er <å> [v:] bundet til efterfølgende <r>, og en opløsning af <år> (*tårn*) kan bedst ske i analogi med opløsning af <ar> (eventuelt i forbindelse med inddragelse af grafotaktiske overvejelser).

De mere hardcore enhedskandidater blandt de nævnte enheder i afsnit 3 er <ch, ci, sch, sh, (s)si, sj, ti> [ɕ] (*charme, speciel, tusch, sherry, asie, mission, sjæl, station*) hvor <c, s, t> ikke har tilsvarende særsegmental korrespondens i nogen forbindelser (i modsætning til <n> i <ng>), og hvor der hvis forbindelserne skulle opløses, vil være tale om specielle regler for begge segmenter. Men opløsning vil eventuelt kunne begrundes analogisk med at <s> optræder i både <sj> og <sh>, og <h> i både <ch> og <sh>. Lignende forhold kan gælde for vokal-

konsonant-kombinationerne <al, aw> [v:] (*catwalk*, *traw*), men ikke for <ur> [œ:] (*burger*).

Vokalkombinationerne <ai, au, ea, eau, ee, oo> (*fair*, *chauffør*, *gear*, *bureau*, *speeder*, *boom*) må betragtes som udelelige enheder – også selv om et af segmenterne i sekvensen i andre men mere begrænsede tilfælde har samme relationer som sekvensen (som når <e> i <ea, ee> [i:] i *gear* og *jeep* kan siges at have samme relation som <e> [i:] i fx *default*). Den endocentriske kombination <ou> (*clou* og *tough*) kunne tænkes opløst, men så med noget specielle nul-regler for henholdsvis <o> og <u>; så går det bedre med <e> [Ø] i <ie> (*diesel*, *hippie*). Hverken Venezky eller Carney opløser vokalkombinationer med enkeltsegmental relation (bortset fra tilfælde hvor <e> med funktion som ”silent <e>” optræder i forbindelse med forudgående vokal som fx i *avenue*, *dye*), heller ikke hvor der er tale om endocentriske sekvenser som <ea> i *bread*, *leather* (jf. fx *bred*, *tether*).

Det forhold der taler mod opløsning af <sj, ea> etc.: at der her vil være specielle regler for begge segmenter, er det samme som det der taler mod særsegmental sammenstilling af fx <g> [n] og <n> [j] i *cognac*, og som dermed placerer <gn> som kompleks type 4-enhed (afsnit 4.1). Juul (2010) diskuterer segmenteringen af nogle sådanne forbindelser i afsnittet ”Nogle vanskelige tilfælde” hvor han bl.a. accepterer ”en dobbelt kompleks forbindelse ch=[tj] for at undgå de to sære forbindelser c=[t] og h=[j]”. Andre tilfælde som ud fra lignende overvejelser kunne medregnes som komplekse type 4-enheder, er kombinationerne <ea> [ej] (*steak*) og <oa> [ɔw] (*toast*) hvor Bogstavlyd har særsegmental opdeling (s|t|e|a|k, t|o|a|s|t) med de lidt specielle relationer <a> [j] og <a> [w] på tværs af grafemisk og fonemisk inddeling i vokaler og konsonanter.

Problemet ”sære forbindelser” er fælles for restløs segmentering og segmentering med nul-relateringer, men nul-relateringen åbner for flere tilfælde af problematisk karakter. For nul-relatering med betingede regler har desuden vurdering af reglernes kompleksitet betydning for segmenteringen. Enklest er opløsning af endocentriske sekvenser som <hv>, mens opløsning af exocentriske sekvenser som <ng> kræver mellemed i relateringen. I næste afsnit ses på fordeling af betingede regler på adskilte beskrivelsesniveauer.

5.2 Lægdelt beskrivelse

Relateringer der forbinder skrift og lyd, må forholde sig til et lydligt primært niveau, det samme som gælder for regler i en fonologisk analyse. Enheder på et sådant niveau vil så eventuelt igen kunne relateres til varianter på andre niveauer. Relateringen i Bogstavlyd forholder sig således til Den Danske Ordbogs ordudtaler hvoraf flere igen vil kunne relateres til andre udtaler på andre distinkthedsniveauer, fx *egern* <g> [j => Ø], *poststempel* <t> [d => Ø] og *hindbær* <nd> [n => m].

En mere abstrakt form for relatering er relatering til fonemisk niveau. For enkeltsegmenter som <i> [i, e, ε] i fx *lille*, *pille* og *drille* betyder det at de kan henføres til to leksikalsk betingede grupper *lille* og *pille*, *drille* med henholdsvis <i> /i/ og <i> /e/ og med efterfølgende kontekstbetinget /e/ => [ε] efter /ʁ/. For sekvenser som <ng> og <ar> betyder det at de vil kunne svare til /n/ + /g/ og /a/ + /r/ med fusion til [ŋ] og [a:] i en fonologisk analyse som den hos Grønnum (2005, 2007). Alt i alt kan en beskrivelse der søger at adskille grafofonemiske relationer fra fonologisk beskrivbare processer, give et klarere billede af forholdet mellem regelmæssigheder og uregelmæssigheder i det ortografiske system end en direkte fonetisk relatering.

Lidt mere abstrakt kan man operere med relatering til et præfonemisk niveau. Venezkys grafofonemiske beskrivelse er på denne måde en tolagsmodel hvor "the correct pronunciation can be derived by first mapping all spelling units onto a pre-phonemic level and then applying the rules for leveling non-English clusters to obtain the phonemic forms" (Venezky 1970: 45), og videre: "Geminate consonant clusters [...] are not single relational units since their morphophonemic forms can be obtained from rules based upon their separate constituents. (The leveling of clusters like {ff} to {f} can be accounted for by a general phonotactical rule)." (ibid.: 51). Med andre ord: Ved denne form for beskrivelse skal man ikke regne med to forskellige enheder <f> og <ff>, relateringen af <ff> i *kaffe* sker i første omgang ved en enkelt regel <f> /f/.

Med samme form for filter kan en forbindelse som <sc> beskrives med de regler der i øvrigt gælder for <c>, dvs. <c> /s/ i fx *scene* (i stedet for <sc> /s/) og derefter med reduktion af /ss/ til /s/ – hvilket vil betyde at <c> i *scene* og *score* vil kunne behandles ensartet ligesom <c>

i *celle* og *collie* i stedet for som <sc> i *sc|e|n|e* og <c> i *s|c|o|r|e*. Relateret af <g> i <ig, eg, æg, ag, yg, øg> i *pige, egern, læge, klage, fhygel, søge* vil kunne beskrives ved reglen <g> => /j/ med efterfølgende fonetisk forklarlig reduktion af /j/ til [Ø] i forskellige af forbindelserne; tilsvarende for <g> i <ug, og, åg, ag> i *ugle, snog, lüge, krage*.

Modsvarende et præfonemisk niveau kan man i en fonografemisk beskrivelse indføre et prægrafemisk niveau. En form for et sådant benyttes af Rollings (2004). Han beskriver hovedkorrespondenser ved "major rules" og modificerer derpå disse til andre korrespondenser ved "re-write rules", fx /k/ => <c> (*cal*) og derpå <c> => <k> foran <e, i, y> (*kill*), /ʃ/ => <sh> (*ship*) og <sh> => <ti> (*nation*).

5.3 Fonografemisk og bidirektional beskrivelse

Det omvendte af en grafonemisk beskrivelse med nul-regler er en fonografemisk beskrivelse med indsætningsregler, fx kan <h> indsættes i *hjul* og *hval* med den leksikalsk betingede regel [Ø] => <h> / # __ [j, v] (der også kan formuleres med den grafemiske specifikation <j, v>, dvs. efter [j, v] => <j, v>). Sekvensen <ng> i ord som *tunge* og *lang* kan genereres ved reglerne [Ø] => <g> / [ŋ] __ [ə, #] (med undtagelser af typen *beton, restaurant*) og [ŋ] => <n>.

Ligesom nul-relationer i den grafonemisk retning giver mulighed for mere generelle udsagn end enkeltrelateringer som <hj> [j] og <hv> [v], muliggør fonografemiske indsætningsregler sammenfatninger som at der er <h> foran [j, v] i de og de ord (frem for at [j] skrives <hj> og [v] <hv> i de og de ord), og at [ŋ] skrives <n> i både *tunge* og *tango* (i stedet for henholdsvis <ng> og <n>).

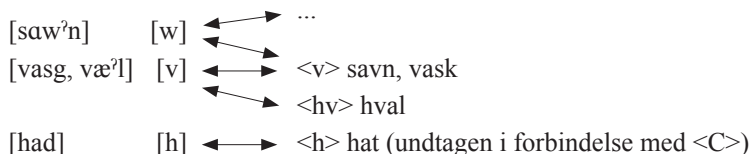
Alligevel kan det, især når man ser isoleret på relationer, synes mere indlysende at gå ud fra konkrete størrelser end at indsætte grafemer fra nul i en fonografemisk relateret. Svarende til nul-reglen E.2 for <e> i ord som *case* og *time* beskriver Carney (1994: 129, 164) under "Speech-to-text correspondences" vokaludtalerne i de diskontinuerte sekvenser <a..e> og <i..e> som "marked by an <e> after the consonant". Men han har såvel <kn> fx *knot* som <ng> i *king* som korresponderende enheder for henholdsvis /n/ og /ŋ/, fx "ŋ/ [...] is spelt with the digraph <ng> [...] unless /k/ or /g/ follows" (ibid.: 246-249), selv om <k> og

<g> i disse sekvenser under ”Text-to-speech correspondences” beskrives ved nul-reglerne K.1 og G.5. Der kan således være en vis forskel i beskrivelserne for en fonografemisk og en grafofonemisk relatering når det drejer sig om komplekse enheder. Mens man med nul-relatering i en almen grafofonemisk relatering kan gå fra en sekvens med flere segmenter til færre segmenter, kan det være mindre nærliggende at indsætte grafemer i en tilsvarende fonografemisk fremstilling.

Ved en bidirektional relatering vil valg i enten den fonografemiske eller grafofonemiske beskrivelsesform være bestemmende også for enhederne i den anden retning. Hvis [v] relateres til <hv>, vil denne relation også gælde grafofonemisk med <hv> som enhed relateret til [v].

5.4 Diagrammer og søgepaneler

I relationelle oversigter læses repræsentationsmuligheder ud fra valgte udgangspositioner som fx <h> eller <a> eller fx [v] eller [a]. For relateringer der inkorporerer komplekse enheder, hvad enten bi- eller unidirektionelt orienterede, vil der også være forskellige valgmuligheder i udgangspositionen for en og samme segmentale størrelse, fx om relationer med <h> skal søges med udgangspunkt i <h> eller i <hv>, og om relationer med [a] skal søges med udgangspunkt i [a] eller [aj]. I et relationsdiagram kan det se sådan ud:



Betingelser kan som her for <hv> anføres i oversigten, men de komplicerede forhold mellem lyd og skrift vil nok være prohibitive for sådanne annoteringer i skematiske eller diagrammatiske visninger.

I stedet kan en diagrammatiske fremstilling som ovenstående, men uden angivelse af kontekstbetingelser, aflæses med den subtraktive forholdsregel at komplekse enheder rangerer over monosegmentale; dette vil gælde begge sider, både fx <hv> over <h> (i **hval** og **hat**) og [aj] over [a] (i **s*i*te** og **v*e*j**). Et fuldt netværk fremstillet i et relationsdiagram vil dog nok stadig forekomme ret kompliceret, og oversigten kan da

deles i to (som hos Basbøll 2005: 101 hvor relationer med <hv, hj, ld, nd, rd> og <ng, ar, år, or, er, sj> er grupperet for sig).

Todeling er også en oversigtsskabende mulighed i andre former for fremstilling. I Bogstavlyd er søgning i "Bogstav til lyd" og i "Lyd til bogstav" underopdelt i søgepaneler med monosegmental og polysegmental indgang. Efter valg af enkelt bogstav- eller lydenhed vælger man igen mellem kompleks og simpel enhed, med prioritet for den første. Som det forklares i vejledningen: "Når et bogstav (fx *h*) er valgt, fremkommer en klikbar liste med alle "komplekse enheder" med dette bogstav. Ord som *hvem* og *hjem* indgår i søgeresultatet hvis man søger på de komplekse enheder <hv> og <hj>, men *ikke* hvis man søger på *h* som en selvstændig, simpel bogstavenhed." Og: "Ord som *job* indgår i søgeresultatet hvis man søger på den komplekse enhed [dj], men *ikke* hvis man søger på [d] og [j] som selvstændige, simple lydenheder."

5.5 Primære og sekundære enheder og relationer

Ortografiske beskrivelser må rimeligvis sondre mellem mere generelle skrift-lyd-relationer og relationer der kun gælder få ord. Ud fra frekvensmæssigt skøn over enhedernes dækning grupperer Venezky (1970: 52, 54) sin beskrivelses grafemenheder i 56 "major" og 7 "minor relational units" (de sidste udelukkende komplekse enheder). Carney (1994: 270-272) har en opdeling i "core rules", "minor rules" og "marginal rules" efter en regels procentuelle effektivitet til at forudsige et bestemt fonem, henholdsvis 93, 26 og 106 regler hvor de sidste "contributes less than 2.5 per cent to the successful prediction of the phoneme". Bogstavlyd sætter ikke sådanne skel mellem mere og mindre frekvente enheder eller relationer, men under de enkelte enheder er relationerne ordnet efter frekvens, og den enkelte enheds "forudsigelighedsgrad" anføres, dvs. jo færre forskellige relationer, desto højere grad af forudsigelighed.

I dansk kan de fremmede grafemenheder <q, w, x, z> regnes som "minor relational units" set ud fra deres hyppighed; af disse har <w> i Bogstavlyd den højeste score med 0,1 % af ordformerne (Juul 2010: tabel 1). Men, hvilket Juul fremhæver i sin artikel, også hovedparten af de 201 komplekse grafemenheder i Bogstavlyds registrering er lavfrekvente, både med hensyn til deres forekomst i korpusset og med

hensyn til det antal relationer de indgår i. De 201 enheder står for en forekomst på kun 10 % af det samlede antal løbende grafemenheder i ordene i Bogstavlyd, og ”kun 15 [...] tegner sig for mere end 0,1 % af samtlige bogstavenheder” (ibid.: 23-25). Disse 15 er, med forekomster på fra 2,1 til 0,2 %, <er, nd, ng, or, ar, ld, ig, re, dt> og dobbeltkonsonanterne <tt, kk, ll, pp, mm, ss>. Med i Juuls tabel 1 med de 33 hyppigste komplekse grafemenheder er også <år, ds, rd, hv, ug, hj, lv, sj, ag> hver med med 0,1 %; en tredjedel af de i tabellen anførte enheder er konsonantfordoblinger. Af rene fremmede enheder blandt de 33 er der kun <ch, ck, sh>, hver med 0,1 %. Ellers er der – bortset fra enkelte tilfælde som det at <nd> også dækker <nd> [ŋ] i *gourmand* og <or> omfatter [œ:] i *password* – udelukkende tale om komplekse enheder med hjemlige korrespondenser.

Yderligere så reduceres andelen af komplekse grafemenheder når man ser på deres relationelle hyppighed. I Juuls tabel 2 over de 33 ”hyppigste bogstav-lyd-forbindelser”, hvor relationernes hyppighed i Bogstavlyd går fra 13,2 % til 0,5 %, optræder af komplekse grafemenheder kun <er> [ʌ], <ng> [ŋ] og <nd> [n] (henholdsvis med 2,1, 0,6 og 0,5 %).

Ud fra disse tal vil størstedelen af Bogstavlyds komplekse grafemenheder i en delt ortografisk beskrivelse kunne behandles for sig adskilt fra primære enheder og relationer, hvad enten de beskrives som relationelle enheder eller opløselige sekvenser. En sådan gruppe vil omfatte særlige fremmedordsstavemåder og -relationer som i *lease*, *shoppe* og *burger* og relationer med komplekse fonemenheder som i *juice* og *tape* ligesom også enkeltsegmentale relationer som i *caste* og *brunch* – på samme måde som en fonologisk beskrivelses opdeling som den hos Grønnum (2005, 2007). Andre særlige relationer findes i gruppen af forskellige reducerede former af hyppige ord, fx <il> [e] i *til*, <od> [oʔ] i *god* og *stod* og <ive> [i] i *bliver* og *giver*, og i enkeltord som fx *snedker* og *købmand*.

6 KOMPLEKSE ENHEDERS FUNKTIONALITET

6.1 Differentiering, overdifferentiering og substitutionsmuligheder

Komplekse grafemenheder er nødvendige i en ortografi i tilfælde hvor der ikke findes enkeltgrafemiske repræsentanter for foner med fonemisk status, i det hele taget eller i bestemte omgivelser. I dansk drejer det sig om [ɛ] (*sjæl*), [ŋ] (*lang*), [ɑ:] (*kar:se*), [v:] (*tår:ve*) og [ʌ] (*sommer, pare*) repræsenteret ved bl.a. <sj, ng, ar, år, er, re> – enkeltgrafemisk repræsentation er her begrænset til et mindre antal fremmedord (fx *gene, jaloux, beton, caste*). Hertil kommer også geminering af konsonant som indirekte markering af vokalkvantitet.

I andre tilfælde er komplekse enheder ikke funktionelle i den forstand at de generelt eller i enkeltord vil kunne udskiftes med et enkeltgrafem eller en anden mere gængs kompleks enhed, på samme måde som også enkeltgrafemer kan udskiftes med andre grafemer. Forslag til sådanne ændringer i retskrivningen kan man læse om i bl.a. Aage Hansens *Om moderne dansk retskrivning* (Hansen 1969; jf. også Ågerup Jervelund 2007).

Endocentriske enheder kan reduceres til enkeltgrafemer, fx <hv> (fx *hval*), <ds> efter lang vokal (fx *kræds*) og enheder i fremmedord som fx *hockey, briese, nougat, juice, ghetto, guitar*, heriblandt dobbeltkonsonant foran trykstærk vokal, fx *effekt, ballade, kommode, kassere*.

Exocentriske enheder kan i nogle tilfælde udskiftes med enkeltgrafemer, fx *plaid, bureaukrati, zoome* (**plæd, *byrokrati, *zume*). Og nogle exocentriske enheder er redundante på den måde at de kan udskiftes med andre komplekse enheder, fx <ch> med <sj> (*charme: *sjarme*). Det samme gælder enheder med <d> i fx *fælde, kande, pludselig, plads* og *krudt* som kan udskiftes med geminater (**fælle, *kanne, *plusselig*, med enkeltskrivning i ordudlyd **plas, *krut*). Men i nogle tilfælde fungerer <d> i endo-exocentriske forbindelser som markør også for vokaludtalen, fx *bonde, tonde* (sammenlign *nonne, bonne*) eller som sikker markør for stød, fx *stald* (jf. *bal* og *tal* henholdsvis med og uden stød).

En del især nyere fremmedord i dansk har udtaler som afviger fra den fonologiske struktur i det danskhjemlige ordstof (jf. Grønnum 2005: 209-211), fx afviger en forbindelse som [ɛj] i *container* fra den hjemlige fonotaks. I dette tilfælde kan sekvensen <ai> tænkes udskiftet

med den lydrette men grafotaktisk afvigende forbindelse <æj>: **con-tøjner*. Men der er også tilfælde hvor en fonologisk struktur der er forskellig fra hjemlige ords, ikke lader sig gengive ved lydrette stavemåder. Ord som *server*, *T-shirt*, *password* og *burger* skiller sig fonotaktisk ud fra hjemlige ord ved udtale med [æ:] (eller [æ:ɪ]) (hvor hjemlige ord kun har [æ:] i forbindelserne [æʔv, æ:ʌ] svarende til <ør, øre> (*dør, gøre*) og kun [æ:] i *bøne* og efter [ʁ], fx *røbe*). Udskiftning af <er, ir, or, ur> i disse ord med <ør> eller <ø>: **sørver* eller **søver*, vil involvere nye leksikalsk baserede regler for repræsentationen af <ør> eller <ø>.

I nogle tilfælde hvor indlånte ords lydlig struktur strider mod regler for hjemlige ords lydstruktur, kan en fremmed stavemåde være grafofonemisk funktionel som markør. En enhed som <ck> (i stedet for <kk>) i fx *backe* og *snack* kan således ses at have den funktion at markere udtalen kort <a> [a], hvilken ellers ikke forekommer foran velær konsonant i hjemlige ord (jf. [a] i *hække*). Andre sådanne tilfælde har at gøre med stavelsesstruktur og fravær af stød. De ordudlydende trykstavelser [V:Cσ#] og [VCstemtCɪ#] har stød i hjemlige ord, fx *ti*, *lim*, *dansk* (frasat nogle med /VrCustemɪ/, fx *mark* og *vers*), men fremmedord som *tee*, *spam*, *band*, *mail*, *light* og *layout* har lang vokal eller konsonantkombination uden stød. Stødløshed i disse typer indikeres dog i flere tilfælde grafemisk af komplekse enheder: ved vokalkombination i fx *tee*, *team*, *jeep*, *lied*, *boom*, *velour* og ved diskontinuert *e*-markering <V₁..e> i fx *mayonnaise* [majo'nɛ:s], *flute* [fly:d] og fx *lime* [lajm], *tape* [tejb], *cornflakes* ['kɔɹn,flɛjgs] (hvor <i, a> i de sidste eksempler svarer til diftong). Alternative stavemåder vil i givet fald være fx **tim*, **bum*, **fhyt* og **lajm*, **tæjp*.

Komplekse fonenheder (type 3 og 4, afsnit 4.1) kan i de fleste tilfælde repræsenteres ved særsegmentale relationer, fx [ju] i *tune* ved <ju> (**tjune*) og [nj] i *cognac* ved <nj> (**conjac*). De diftongiske enheder [aj] i *timing*, [ɛj] i *baby* og [ɔw] i *poster* kan repræsenteres ved <aj>, <æj> og eventuelt <ow> (jf. *show*). Men sådanne udskiftninger ligger dog ikke lige for i ord som *lime*, *tape* sb. og *joke* sb. med "silent <e>", jf. ovenfor.

6.2 Afkodning: staver og læser

Ligesom det gælder for enkeltgrafemer, er også komplekse grafemenheders relationelle funktionalitet for de flestes vedkommende forskellig i grafofonemisk og fonografemisk henseende. Generelt er der med den grafemiske overdifferentiering som de komplekse enheder bidrager til, tale om flere valgmuligheder og mindre forudsigelighed i staveretningen end i læseretningen. En relation som [a:] <ar> er med kontekstregler nogenlunde lige entydig i begge retninger (jf. dog bl.a. <ar, arr> i *bare*, *barre*). Men i andre tilfælde er der tale om større flertydighed i staveretningen, fx fonografemisk [v] <hv, v> (*hval*, *valg*) over for grafofonemisk kun <hv> [v]. Alt andet lige (dvs. hvis man ser bort fra forhold som fx viden om frekvens og proveniens) kan enhedernes rolle her antages at være mere problematisk for staveren end for læseren.

På den anden side kan der ved grafofonemisk afkodning af sekvenser være flere udfordringer end det at finde den rette korrespondens til en enhed, nemlig også det at afgrænse de rette grafemenheder (jf. Elbro 2002, 2006, Juul & Sigurdsson 2004). En læser som går analytisk frem, har her problemer der på sin vis modsvarer lingvistens segmenteringsovervejelser. I tilfældet <..hv..> er relateringen entydig (forudsat man kender til den simple regel <hv> [v]). I andre tilfælde som <..nd..> i *pande* og *panda* kræves der kontekstuel information (i begrænset omfang leksikalsk, jf. fx substantivet *bande*) for at afgøre om sekvensen korresponderer med en eller to foner. Andre tilfælde igen som <..ea..> i *areal*, *bearnaise*, *beaujolais*, *beautyboks*, *gear*, *research*, *steak*, *sweater* og <..ui..> i *etui*, *fluidum*, *guitar*, *juice*, *quiz* er mere uoverskuelige med flere mindre forudsigelige afgrænsninger.

I regelsæt som fx i en søgefunktion må man gennemgå forskellige valgmuligheder, fx <nd, n, d> og <gu, ui, u, i>, forud for søgning på de relevante relationer. I et læringsmæssigt materiale kan det være tjenligt at ord med de maksimale sekvenser sammenstilles med deres forskellige aflæsningsmuligheder.

7 NORMERING

7.1 *Traditionsprincippet og sprogbrugsprincippet*

De overordnede retningslinjer for regulering af dansk retskrivning er møntet i de to principper traditionsprincippet og sprogbrugsprincippet (§ 1 i sprognævnsbekendtgørelsen, Bek. nr. 178 af 9.3.2009²). De to principper sikrer en stabil ortografisk norm, men tillader også i vidt omfang fremmede stavemåder i importord idet ”Ord fra andre sprog [dvs. end græsk, latin og fransk], specielt nyere ord, skrives som hovedregel i overensstemmelse med skrivemåden i det pågældende sprog eller med international praksis”. Aktuelt vil dette hovedsagelig sige ord fra engelsk (jf. Galberg Jacobsen 1997: 4, 2010: 239-240, Ågerup Jervelund 2007: 13, 15). I bekendtgørelsen refereres der til 1892-bekendtgørelsen, Bek. nr. 24 af 27.2.1892³, hvis detaljerede anvisninger har været styrende for retskrivningen og fortsat har indvirkning på den. I bestemmelserne i 1892-bekendtgørelsen gives anvisninger på flere forhold som vedrører komplekse enheder.

I de følgende afsnit omtales reguleringer som vedrører komplekse enheder i de seneste fem retskrivningsordbøger fra RO55 til RO12. Oversigter over ændringer i RO55, RO86, RO96 og RO RO01 kan ses i Galberg Jacobsen (2005, 2010) og i de respektive retskrivningsordbøger.

7.2 *Komplekse enheder i fremmedord: regulering og import*

Ifølge 1892-bekendtgørelsen skrives ”Fremmede Ord, der ere fuldt optagne og indgaaede i Sproget og i Udtalen have tabt de fremmede Lyde [...] efter Sprogets almindelige Regler”. At ordene skal have tabt de fremmede lyde, betyder at erstatningsmuligheder skal vurderes ikke blot internt for en enhed men også i forhold til ordets fonologiske struktur som helhed, fx kan *velour* med sin langvokaliske stødløse trykstavelse ikke uden videre skrives **velur* da <ur#> i hjemlige ord har stød. 1892-bekendtgørelsens anvisninger vedrørende fremmedord gælder i sagens natur først og fremmest ord af græsk, latinsk og fransk oprindelse, ud fra et nutidigt synspunkt et hovedsaglig ældre indoptaget ordstof. Om enheder i franske fremmedord hedder det i 1892-bekendtgørelsens § 21 II:

I franske Ord, selv om de ere almindeligt brugte, beholdes den franske Lydbetegnelse, naar i Udtalen særlig franske Lyde (j, g, ch, ll, gn) forekomme; dog bortfalder sædvanligt det stumme e i Enden af Ordene. Ligeledes beholdes Lydbetegnelserne au, eau, ay, oy, oi, ei, ou, e=ø, u, en, em, on, ti, ç, stumt h, stum Endekonsonant [...] Eksempler: jaloux, Jalousi, Chaussé, [...] Derimod friholdes i Almindelighed for de franske Lydbetegnelser ai og eu henholdsvis æ og ø. – f. Eks.: Karantæne, Inspektør [...].

De franske lydbetegnelser <ai> og <eu> har da også siden haft en reduceret status i retskrivningsordbøgernes ordforråd. Tidligere <ai> i bl.a. *distrain* og *Terrain/Terræn* fik med RO55 rene <æ>-former, ligesom RO55 også fastsatte *vigør*, RO86 *causør*, *charmør*, *odør*⁴ og RO96 *miljø* som eneste stavemåder. De franske stavemåder der findes i RO12, hører for de flestes vedkommende til i det mere perifere ordforråd med præg af citatord, fx *sans comparaison*, *vinaigre* og *auteur*, *fleuret*, *meunieresteg* – dog (med fremmed udtale [V:C#]) *cremefraiche*, *mayonnaise* og *voyeur*.

<ou> findes derimod i ord som *boulevard*, *clou*, *desavouere*, *douche*, *fou-ragere*, *gourmand*, *jalousi*, *nougat*, *roulette*, *souschef*, *souvenir*, og (med [V:C#]) *ajour*, *rouge* og *velour*. Enkelte ord som har ”tabt de fremmed Lyde” har fået <u>, i RO55 bl.a. *keutyme*, *rutine* (for tidligere *Coutume* og *R(o)utine*), og *camouflage*, *rembours*, *remoulade*, *ressource* fik i RO86 sideformerne / *kamouflage*, / *remburs*, / *remulade*, / *resurse*.

Andre forenklinger af komplekse enheder i nyere retskrivning har drejet sig om reduktion af de endocentriske vokalenheder <ee, ei, ie> i RO55 *puré* og RO86 *fresia*, *spontaneitet* og *regi* (men stadig <ee> i *jamboree* og <ie> i *sortie*, *hierarki*, fra tysk bl.a. *diesel*, *liebhaber*, fra engelsk fx *briefe*, *cookie*, *pierce*) og af en endocentrisk konsonantenhed <th> i RO55 *tuja* og RO86 *forsytia*, *maratonløb*, *timoté* (men <th> i bl.a. *othellokage*, *thai*). Ændringerne til *puré* og *regi* følger 1892-bekendtgørelsen (”dog bortfalder sædvanligt det stumme e i Enden af Ordene”). Det gør også forenkling af <th>. I bekendtgørelsens § 21 I c siges det at ”I almindelig brugte Ord ombyttes ph med f (Filosofi, Fysik, Sfære, Sfinks), th med t (Teater, Teologi), rh med r (Rytme), ch = χ med k (Katekismus, Kor, Kirurgi, Kemi, Orkester, Skema, Skisma)”; det er bl.a. denne bestemmelse der sætter retskrivningen på kollisionskurs med naturvidenskabelige stavemåder som *lithium*, *phenol* og *chrom*.

Herudover er <qu> blevet til <k> i RO55 *markis* og RO86 *char-tek*, *kommuniké* (men <qu> i bl.a. *bouquet*, *enquête*, *roquefort*). Og i RO01 er *bolsje*, *gesjæft*, *gletsjer*, *rutsje* endeligt fordansket til <s> for tidligere <(s)ch> (trods bemærkningen i 1892-bekendtgørelsen om at ”ch, g, j eller ti [aldrig ombyttes] med sj (undt. Ansjos)”). Ændring til <lj> i RO55 *emalje*, *medalje* og *føljeton* af tidligere <ill> (*Emaille*, *Medaille*/*Medalje* og *Feuilleton*) følger udtalen ligesom også i RO86 *vanille*/*vanilje*, RO01 *vanilje* (hvorimod *paillet* og *bonillon* holder stand). <gn> i *cognac* havde i RO86-96 den alternative stavning <nj> (*konjak*).

Modsat tendensen til udskiftning af fremmede enheder med hjemlige er der i nogle ord sket tilbageføring til fremmede stavemåder. Ord som *Bulevard*, *G(u)itar*, *Li(e)bhaber*, *Vu/Vy*, *Kræm* ændredes i RO55 til *boule-ward*, *guitar*, *liebhaver*, *vue/vy*, *creme* (det sidste ord igen ændret til */krem* i RO86 og tilbage til *creme* i RO01); *margerit* fik sideformen *marguerit* i RO86, og *shampo(o)* og *vermouth/vermut* fik eneformerne *shampoo* og *vermouth* i RO01. I RO12 er *crepe*, *mayonnaise*, *schnitzel*, *Schweiz*, *spaghetti*, *techno*, *youghurt* og *zombie* nu eneste staveformer hvor der forhen var valgfrihed med stavemåderne *krep*, *majonase*, *snitsel*, *Svejs*, *spagetti*, *teknø*, *jogurt* og *zombi*. Dertil har *mikse(r)* fået sideformen *mixer*, og *remiks* er blevet til *remix*.

De reguleringer af fremmedords stavemåder der har været i perioden 1955 og frem, har kun i få tilfælde berørt komplekse enheder i ord fra engelsk og her med tilbageføring til engelske stavemåder. *Plad/Plaid* blev i RO55 til *plaid*; *pie*⁵ havde i kun kortere tid sideformen *paj* i RO86-96, og *roastbeef* havde efter udtalen [ˈʁʌsd-] den danske sideform *rostbøf* i RO86-01. Henvisningerne fra *djungle* og *djunke* til *jungle* og *junke* forsvandt endegyldigt med RO86.

For nyoptagne ord reguleres der med hensyn til konsonantfordobling i forbindelse med danske bøjningsendelser, fx *chatte*, *auditten*. Men ellers sker der i det hele taget kun i ringe grad ortografisk tilpasning af direkte lån (*krasar* for *quasar*, *nørd* for *nerd*); et ord med <ø> for <ea> er *lørner* som er optaget i RO01 med denne stavemåde (Jarvad 1995: 71-72).

Fra og med RO86 er der i retskrivningen kommet flere mere eller mindre centrale ord til med enheder som <cc, ch, ck, ea, ee, ey, ie, gh, gn, gu, oa, oo, ou, qu, sh, th, ui, ur>, fx *cappuccino, brunch, chatte, feedback, deadline, lease, screene, keyboard, brieft, light, tough, lasagne, baguette, snowboard, hooked, blouson, touch, quiche, shoppe, smoothie, jumpsuit* og *burger*, og med <a, i, j, o, u> svarende til [ej, aj, dj, ɔw, ju], fx *tape, live, junk, poster, unit*. I mange af disse er stavemåden funktionel i den forstand at den gengiver ”fremmed Lyde” i ord med en fonologisk struktur der afviger fra den hjemlige, og således ikke vil kunne reguleres efter det der indtil videre er ”Sprogets almindelige Regler”.

7.3 Enkelt- og dobbeltkonsonant efter trykssvag vokal

Et særligt område er konsonantfordobling efter trykssvag vokal, dvs. i ord hvor den forudgående vokallængde ikke er nøgle til konsonant-skrivningen. RO86 normaliserede stavningen i posttonal stavelse foran bøjningsendelse til tre regler, RO01 er gået videre til to regler, fx *motrikker, fernissen*, men *leksikonet* (RO01 § 10(2)). Men ud over spørgsmål som stavning af sidste [n] i [ˈananas] kan særlig konsonantfordobling efter prætonal vokal være et problem som kan kræve ordbogsopslag også for den øvede staver, fx om der er <l> eller <ll> i [baˈlæːðə], [baˈlɛd], [baˈlɒŋ]. Det hedder om denne type i 1892-bekendtgørelsens § 21 III: ”Medlydsfordobling efter ubetonet Selvlyd beholdes i Reglen, naar den findes i det fremmede Sprog (Allé, Ballon, abonnere, Buffet o. s. v.); dog skrives enkelt Medlyd overalt, hvor Brugen har indført det: Bajonet, Kanel, Karbonade, Pomade, Paryk, traditionel m. m.”.

De fleste sådanne ord har enkeltkonsonant (jf. § 9(2b) i RO86 og RO01, Becker-Christensen 1980: 68-69), og dobbeltkonsonant foran trykstærk vokal er da også forenklet i nogle spredte tilfælde i RO86: *pen(n)alhus, diar(r)é, gonor(r)é, hæmor(r)oider, kar(r)osseri* (sprogbrugsbegrundelsen for *penalhus* er omtalt i Hansen 1980: 22). Omvendt er enkeltkonsonant i overensstemmelse med tegnkonstansprincippet ændret til dobbelt foran afledningsendelse i *komplettere, kontrollant, modellere, rekruttere* (RO1955) og *galop(p)ere, glos(s)ar* (RO 1986) som ved trykskifte har dobbeltskrivning i andre former, fx *komplette* og *glosse*, men uden fordobling fx *profitere* og *trafikant* (jf. herom Hansen 1969: 82 og § 9(2c) i RO86 og RO01).

7.4 Hjemlige komplekse enheder: <hj, hv, ds, dt, ld, nd, rd> med flere

Om stumt <h> i forbindelserne <hj, hv> siger 1892-bekendtgørelsens § 16 at ”H foran j og v skrives i samme Udstrækning som hidtil, altsaa f.Eks. hjem[.] hjælpe, Hjeje, hviske, hvisle, hvæsse, hver, Hvilling, Pig-hvarre, hvæse”, og der har ikke i perioden været ændret ved disse forbindelser.

§ 17 i 1892-bekendtgørelsen giver regler for stumt <d>. Bl.a. bibeholdes <d> i <ld, nd>, fx *Guld, Handling, Mund*, og i <ds> i fx *Klods, Plads*, men udelades ”mellem n og s og foran sk, naar Ordets Stamme ligger uden for Sprogets Grænse”, fx *Krans, Pisk, Sveske* (i det sidste ord selv om <d> kunne have den grafofonemiske funktion at markere udtalen <e> [e]). Stumt <d> i <ds> forsvandt i RO55 i *kyras* og *jyske, vaske* og *væske* (trods forbindelsen med *jyde* og *våd*) og i RO01 efter lang vokal i bl.a. *fråse* og *slabberas* (men findes stadig i fx *palads, kryds* og *stads*). Foran afledningsender med <t> skrives ifølge bekendtgørelsen <d> i fx *Fedt* og *Midte* (ligesom også foran bøjningsendelse, fx *mødte*), men <t> i fx *flittig* og *gotte* sig (det sidste ord har dog i analogi med *god* fået <dt> i RO86); *ti(d)t* fik i RO55 eneformen *tit*.

Med RO55 forsvandt skrivemåden [j] <jg> da *dejs, fejs* og *sejs* blev til *dej, fej* og *sej*. Og med RO86 ændredes *viss* til *vis* hvorefter der ikke mere findes dobbeltvokalisk repræsentation af lang vokal i hjemlige ord.

De hjemlige enheder <hj, hv, ds, dt, ld, nd, rd> optræder i et begrænset antal morfemer, i Bogstavlyd fra 15 morfemer for <hj> til ca. 110 for <nd>, og i sammensætninger og afledninger med disse. Deres dækningsområde vokser med sammensætninger som *hvalsafari, tæskehold* og *talskrivinde* som er nogle af de nye ord med disse forbindelser der er registreret i Jarvad (1999), men de forekommer ikke produktivt med nye simpleksord.

Af andre enheder som kan betegnes som hjemlige, er den exocentriske enhed <s> i nyere retskrivning blevet næret af ændringer fra <(s)ch> som i *bolsje* og ved tilgang af ord som *babusjka* og *borsjij*. <ng> er almindelig i indlån, fx *leggings*; samme gælder geminerede konsonanter, fx *mozarella, bipper, chatte*, <er, re>, fx *computer, eskealere*, og <ar, or>, fx *harddisk, support*.

8 AFSLUTTENDE BEMÆRKNINGER

Komplekse relationelle grafemenheder er en gruppe med forskellige afgrænsningsmuligheder – med et større antal enheder i Bogstavlyds restløse segmentering og med et mindre antal i beskrivelser med nul-relationer. Nul-relatering er nødvendig hvis man ønsker at minimere antallet af relationelle grafemenheder i en ortografisk beskrivelse, og hvis man vil fremdrage generelle træk på tværs af den restløse segmenterings komplekse enheder. Man kan fx ikke tale om at der er noget fælles for <hj, hv> i *hjul* og *hval* uden dermed at implicere nul-relationer. Og for en mere generel beskrivelse af vokaludtaler i ord med ”silent <e>” er det i både engelsk og dansk formålstjenligt at segmentere ord som *tape* og *game* med nul-relation for <e> i stedet for som t|a|pe og g|a|me. Nul-relatering giver desuden luft for en mere rimelig beskrivelse i forhold til relationer som <øb> [ø] i *købmand* der forekommer i den gennemført restløse segmenterings yderkroge. Men den dobbelte betragtningsmåde, <hv> som henholdsvis enhed og sekvens med <h> som nul-relateret enhed, er i det store og hele hensigtsmæssig og praktisk i enhver beskrivelse – som det også fremgår af artiklen her.

En restløs segmentering må nødvendigvis resultere i et stort antal komplekse enheder, heriblandt mange med lav frekvens i ordforrådet. For andre former af ortografiske beskrivelser kan det være et problem at bestemme i hvilket omfang det er rimeligt at regne med nul-korrespondenser. I beskrivelser med betingede regler kan der under hensyntagen til grader af regelsimplicitet groft skelnes mellem forholdsvis uproblematisk opløsning af grafemsekvenser med stumt bogstav, dem der her i artiklen er betegnet som endocentriske, fx <hv> [Øv], exocentriske sekvenser som <ng> [ŋØ] der også kan opløses med generelle regler, og exocentriske sekvenser som <au> [o] der ikke kan beskrives ved hjælp af nul-korrespondens uden ved specielle regler for begge segmenter, og som må antages som uomgængelige inventarielle enheder i beskrivelser med nul-relationer. Betingede regler hvori der indgår nul-relatering, er bedst egnede i grafonemiske beskrivelser. I en fonografemisk beskrivelse kan indsættelse af grafem ved betingede regler synes mindre oplagt end at regne med relatering til en kompleks enhed. Dette har så også betydning for den større rolle komplekse enheder kan have i en bidirektionel fremstilling.

Visse komplekse grafemenheder har en nødvendig funktion i retskrivningen, ikke blot hjemlige enheder som <sj, ng> men også enheder i indlån med en fra hjemlige ord afvigende fonologisk struktur, fx <ea> i *team* og <ck> i *hacke* som ikke uden videre kan substituere med hjemligt <i> og <kk>. I de fleste tilfælde betyder komplekse enheder dog en overdifferentiering i skriften, og de vil kunne reduceres til eller udskiftes med enkeltgrafemer eller andre mere gængse komplekse enheder. For den grafonemiske afkoder kan komplekse enheder indebære vanskeligheder med hensyn til afgrænsningen for relateringen, især hvor der er tale om sekvenser som <..ea..> og <..nd..> der ikke er entydigt aflæselige sådan som fx <..hv..>.

De reguleringer der har været af komplekse enheder i retskrivningen fra RO55 til RO12, har hovedsagelig berørt ældre fremmedord af fransk oprindelse. Nye direkte indlån, dvs. især engelske ord, beholder stort set den fremmede stavemåde. En række hjemlige enheder, bl.a. <hj, hv, ld, nd, rd>, er ikke produktive uden for orddannelse ved sammensætning og afledning af eksisterende simpleksord. Til gengæld har der i periodens retskrivningsordbøger været en stigning i antallet af ord med komplekse enheder som fx <ch, ck, ea, ee, gh, oa, oo, ou, sh, ui, ur> som afviger fra hjemlig grafemisk og grafonemisk struktur. Men som det fremgår ud fra det ordstof som Bogstavlyd dækker, er ord med sådanne komplekse enheder stadig svagt repræsenterede i retskrivningen. Og som Juul afsluttende siger i sin artikel: ”Hovedkonklusionen må derfor være at princippet om én lyd til ét bogstav stadig er bærende i dansk ortografi.”

Christian Becker-Christensen

Dansk Sprognævn

chbech@dsn.dk

NOTER

- 1 Termen grafem findes brugt i flere betydninger (se fx Allén 1965: 29-37; 1971: 25). Her i artiklen bruges grafem med betydningen bogstav (som også hos Venezky 1970: 49-50). Men mere præcist bør grafem forstås som en mindste betydningsadskillende enhed som fremkommer ved intern skriftsprogsanalyse (Allén 1971: 25; Sampson 1985: 25). Og i overensstemmelse hermed burde grafem rettelig specificeres som ”bogstavgrafem” når brugt om bogstaver idet bl.a. også taltegn kan rubriceres som grafemer. Alle alfabetets bogstaver kan vises at have status som betydningsadskillende enheder, lidt vanskeligt er det dog med <q>, der som regel optræder i fast kombination med <u>, og hvor et af de få kommutationspar er *qatarerne* vs. *katarerne*. Et af problemerne ved bestemmelsen af grafem såvel som bogstav er i øvrigt om store og små bogstaver hver især er selvstændige grafemer, eller om majuskularitet skal ses som et suprasegmentalt træk (Allén 1971: 31; jf. Diderichsen 1952: 10 som taler om at store bogstaver er ”signaler” for visse indholdsstørrelser (”substantiver” [...]), og at de derfor ikke skal optages i grafeminventaret).

Grafem findes også brugt defineret i forhold til lydsiden og som omfattende flere bogstaver, således forklarer Carney (1994: xxvii) ”grapheme” som ”any minimal letter string used in correspondences. So, <ea> in *head* may be referred to as a single (but complex) grapheme.” Men bedre er det at have en særlig betegnelse for grafemer som relationelle enheder. Allén (1965: 38-40, 1971: 25) har introduceret termene *fonografem* og *grafofonem* for relationelle enheder i henholdsvis skrift- og lydsiden: ”Begreppet fonografem definieras då som klassen av grafem – och i förekommande fall grafemsekvenser – som motsvarar ett fonem. På analogt sätt definieras ett grafofonem som klassen av fonem – och i förekommande fall fonemsekvenser – som motsvarar ett grafem.” Her i artiklen bruges betegnelserne *relationel fon(em)enhed* (fon(em)enhed der svarer til en eller flere grafemenheder) og *relationel grafemenhed* (grafemenhed der svarer til en eller flere fon(em)enheder).

- 2 Paragraffen har samme ordlyd som i den tidligere bekendtgørelse nr. 707 af 4. september 1997 om Dansk Sprognævns virksomhed og sammensætning.
- 3 Flere af bestemmelserne i Bek. nr. 24 af 27.2.1892 går tilbage til Bek. nr. 83 af 7.6.1889, og de her anbefalede ændrede stavemåder er implementeret i Dansk Retskrivningsordbog 1891. 1892-bekendtgørelsen blev sammen med andre tidligere bekendtgørelser ophævet ved Lov om dansk retskrivning 1997 (Lov nr. 332 af 14. maj 1997).

- 4 Ændringen <eur> til <ør> i *nigør* (RO55) og *causør*, *charmør*, *odør* (RO86) forudsætter en ældre udtale [øʔɐ] (jf. Brink & Lund 1975: 592-593; SDU). Med den nu almindelige udtale [æ:ɐ] i disse ord er <ør> for <eur> ikke i trit med 1892-bekendtgørelsen bestemmelse om at regulering af fremmeord kræver at de har mistet den fremmede udtale – eller man må sige at <ør> nu er en accepteret repræsentation for denne udtale. Samme gælder <ør> i *favør*, *interiør* og *parlør* som har <ø> allerede i Dansk Retskrivningsordbog 1891. Med udtalen af <ør> i *chanffør* forholder det sig anderledes.
- 5 I Bogstavlyd relateres <ie> i *pie* til henholdsvis [ɑ:] og [j] (”p:[p] i:[ɑ:] e:[j]”), men alternativt kan <i> beskrives som korresponderende med [ɑ:j] og <e> med nul, jf. afsnit 5.1.

LITTERATUR

- Allén, Sture (1965): *Grafematisk som grundval för textedering*. Acta Universitatis Gothoburgensis. Göteborg: Almqvist & Wiksell.
- Allén, Sture (1971): *Introduktion i grafonomi. Det lingvistiska skriftstudiet*. Stockholm: Almqvist & Wiksell.
- Basboll, Hans (2005): *The Phonology of Danish*. Oxford: Oxford University Press.
- Becker-Christensen, Christian (1980): Om dobbeltskrevet konsonant. *Ord til andet. Iagttagelser og synspunkter 2. Dansk Sprognævns skrifter 10*. København: Gyldendal. 65-74.
- Becker-Christensen, Christian (1988): *Bogstav og lyd. Dansk retskrivning og rigsmålsudtale*. København: Gyldendal.
- Bek. nr. 83 af 7.6.1889 = *Bekjendtgørelse fra Ministeriet for Kirke- og Undervisningsvæsenet om Retskrivningen, Nr. 83, 7de Juni 1889*.
- Bek. nr. 24 af 27.2.1892 = *Bekendtgørelse fra Ministeriet for Kirke- og Undervisningsvæsenet om Retskrivningen, Nr. 24, 27de Februar 1892*. Kan ses på Dansk Sprognævns hjemmeside dsn.dk.
- Bek. nr. 178 af 9.3.2009 = *Bekendtgørelse om Dansk Sprognævns virksomhed og sammensætning, BEK nr. 178 af 09/03/2009*. Kan ses på Dansk Sprognævns hjemmeside dsn.dk.
- Bogstavlyd* = <http://bogstavlyd.ku.dk>. Udviklet af Holger Juul i samarbejde med Philip Diderichsen. Officielt åbnet 6. maj 2011.
- Brink, Lars & Jørn Lund (1975): *Dansk Rigsmål. Lydudviklingen siden 1840 med særligt henblik på sociolekterne i København*. Bd. 1-2. København: Gyldendal.
- Carney, Edward (1994): *A Survey of English Spelling*. London: Routledge.
- Dansk Retskrivningsordbog (1891) = Saaby, Viggo: *Dansk Retskrivningsordbog*. Udarbejdet i Overensstemmelse med de ministerielle retskrivningsregler af 7^{de} Juni 1889. København: C.A. Reitzels Forlag.
- Dansk Retskrivningsordbog (1923) = *Dansk Retskrivningsordbog*. Udgivet af Undervisningsministeriets Retskrivningsudvalg. Ved Jørgen Glahder. København: Gyldendalske Boghandel. 3. udgave, 1929.
- Den Danske Ordbog*. København: Gyldendal, 2003-2005. Også (med IPA) ordnet. dk.ddo.
- Diderichsen, Paul (1952): *Nye bidrag til en analyse af det danske skriftsprogs struktur*. Selskab for Nordisk Filologi. Årsberetning for 1951-52. 6-22.
- Diderichsen, Paul (1957): *Elementær Dansk Grammatik*. 2. udgave. København: Gyldendal.

- Elbro, Carsten (2002): Det er ret svært at lære at læse dansk. *Mål og Måle*, 25. årg., nr. 1. 10-17.
- Elbro, Carsten (2006): Literacy Acquisition in Danish: A Deep Orthography in Crosslinguistic Light. Joshi, R. Malatesha & P.G. Aaron: *Handbook of Orthography and Literacy*. Mahwah, New Jersey: Lawrence Erlbaum Associates. 31-45.
- Galberg Jacobsen, Henrik (1997): Ordrede forhold. Om retskrivningsloven og sprog-nævnsløven. *Nyt fra Sprognævnet* 1997/3. 1-10.
- Galberg Jacobsen, Henrik (2005): *På sproglig grund. Danske Sprognævn fra a til å 1955-2005*. Odense: Syddansk Universitetsforlag.
- Galberg Jacobsen, Henrik (2010): *Ret og Skrift. Officiel dansk retskrivning 1739-2005*. Bind 1-2. Odense: Syddansk Universitetsforlag.
- Grønnum, Nina (2005): *Fonetik og Fonologi – Almen og dansk*, 3. udg. København: Akademisk Forlag.
- Grønnum, Nina (2007): *Rødgrød med Fløde. En lille bog om dansk fonetik*. København: Akademisk Forlag.
- Hansen, Erik (1980): Om den nye Retskrivningsordbog. *Ord til andet. Iagttagelser og synspunkter* 2. Dansk Sprognævns skrifter 10. København: Gyldendal. 15-34.
- Hansen, Erik (1981): *Skrift, stavning og retstavning*. København: Hans Reitzel.
- Hansen, Aage (1969): *Om moderne dansk retskrivning*. København: C.E.C. Gads Forlag.
- Jarvad, Pia (1995): *Nye Ord – hvorfor og hvordan?* København: Gyldendal.
- Jarvad, Pia (1999): *Nye Ord. Ordbog over nye ord i dansk 1955-1998*. København: Gyldendal.
- Juul, Holger (2010): K-a-tt-e-pi-n-er. Om komplekse bogstav-lyd-forbindelser i danske ord. *NyS* 39. 10-32.
- Juul, Holger & Baldur Sigurdsson (2004): Ortografien som handicap? Stavetilegnelse i Rødovre og Reykjavik. *NyS* 32. 67-88.
- Lyons, John (1968): *Introduction to Theoretical Linguistics*. Cambridge: Cambridge University Press.
- Nudansk ordbog = *Politikens Nudanske Ordbog*. København: JP/Politikens Forlagshus, 2010.
- RO55 = *Retskrivningsordbog*. Udg. af Dansk Sprognævn. København: Gyldendal, 1955.
- RO86 = *Dansk Sprognævn: Retskrivningsordbogen*. København: Gyldendal, 1986.
- RO96 = *Dansk Sprognævn: Retskrivningsordbogen*, 2. udg. København: Aschehoug, 1996.
- RO01 = *Dansk Sprognævn: Retskrivningsordbogen*, 3. udg. København: Alinea - Aschehoug Dansk Forlag, 2001.

- RO12 = Dansk Sprognævn: *Retskrivningsordbogen*, 4. udg. København: Alinea, 2012.
- Rollings, Andrew G. (2004): *The Spelling Patterns of English*. München: Lincom Europa.
- Sampson, Geoffrey (1985): *Writing Systems*. Stanford, California: Stanford University Press.
- SDU = *Den Store Danske Udtaleordbog*. København: Munksgård, 1991.
- Venezky, Richard L. (1970): *The Structure of English Orthography*. The Hague: Mouton.
- Venezky, Richard L. (1999): *The American Way of Spelling. The Structure and Origins of American English Orthography*. New York, London: The Guilford Press.
- Ågerup Jervelund, Anita (2007): *Sådan staver vi – om ortografi og stavefejl*. København: Dansk Sprognævn og Dansk lærerforenings Forlag.