

Fra analfabet til avanceret ortografibruger

En oversigt over de vanskeligheder man møder som bruger af dansk ortografi

Af Holger Juul

*a
is the first letter of the alphabet
there are twenty-five more
the chicago public library has all of them
in some very interesting combinations¹*

1. Indledning

Der er noget besnærende simpelt ved det alfabetiske princip. Man skal blot lære en god snes bogstaver, som hver repræsenterer en bestemt lyd. Præmien for denne ulejlighed er ikke af de små. Man bliver i stand til at oversætte et hvilket som helst stykke talesprog til skrift eller omvendt – en færdighed som har vist sig så nyttig, at den næsten er blevet en nødvendighed. Det alfabetiske princip er godt fundet på.

På den anden side kan man ikke sige at en alfabetisk ortografi som den danske er et redskab der er let at gå til. Før de får opøvet en tilfredsstillende færdighed i at bruge dette redskab, har de fleste børn brug for intensiv opmuntring og støtte i en lang periode. Ikke så få oplever vedvarende læse- og stavevanskeligheder. Selv blandt universitetsstuderende kan man træffe usikre stavere. Vejen fra analfabet til avanceret ortografibruger er ofte lang og besværlig.

Hensigten med denne artikel er at give en oversigt over de vanskeligheder man må klare som bruger af dansk ortografi hvis man vil opnå status af »superbruger«. Disse vanskeligheder kan groft inddeles i tre grupper.

1. Forudsætninger

Før man kan blive i stand til at oversætte (eller: »omkode«) fra tale til skrift og fra skrift til tale, må man lære at forbinde talesproglige enheder og skriftsproglige (grafiske) enheder med hinanden. Etableringen af for-

bindelser forudsætter naturligvis at man er i stand til at identificere disse enheder. De vanskeligheder som identifikationen kan indebære, behandles i afsnit 2 (talesproglige enheder) og afsnit 3 (skriftsproglige enheder).

2. Principielle vanskeligheder i omkodningen

Selv om man er i stand til at identificere og forbinde tale- og skriftsproglige enheder med hinanden, så kan omkodningen mellem tale og skrift alligevel støde på vanskeligheder. Det er jo ofte sådan at en enhed har flere forskellige mulige modsvarigheder. Og det betyder at omkodningen ikke altid kan foregå helt mekanisk enhed for enhed. Dette problem, som viser sig i begge omkodningsretninger, både fra tale til skrift og fra skrift til tale, behandles i afsnit 4.

Et beslægtet problem er dét at skriftsproget på mange områder er fattigt på information. Selv om hver enkelt grafisk enhed i en trykt sætning entydigt kan omkodes til talesproglige enheder, så kan sætningen alligevel have adskillige fortolkningsmuligheder. Der er nemlig mange talesproglige distinktioner som slet ikke markeres skriftligt. Den usikkerhed i omkodningen fra skrift til tale som skriftens informationsfattigdom kan medføre, er emnet for afsnit 5.

3. Effektiviseret omkodning

Selv om man principielt er i stand til at omkode mellem skrift og tale, så er det altid muligt at omkodningen kunne forløbe hurtigere eller mere præcist. De mangeartede forhold som kan have betydning for omkodningens effektivitet, er emnet for afsnit 6.

Jeg håber først og fremmest at oversigten kan være en hjælp for lærere der underviser i læsning og stavning, ved at gøre det mere klart hvor elevernes problemer kan ligge, og hvor man måske skal gøre en særlig indsats.

Desuden håber jeg at oversigten kan inspirere til øget forskningsaktivitet. I mange tilfælde er det indtil videre uklart præcis hvor alvorlige de vanskeligheder der nævnes i artiklen, er.

Endelig vil jeg gerne gøre det klart at de vanskeligheder der møder danske læsere og stavere, ikke nødvendigvis er direkte sammenlignelige med de vanskeligheder der møder læsere og stavere i andre lande. Det er et forhold som det er vigtigt at være opmærksom på i disse år hvor der gennemføres stadig flere tværsproglige sammenligninger af læsefærdighed (Elley 1992; Sommer, Lau et al. 1996; Jensen og Holm 2000; Sey-

mour 2000). Specielt finske læsere har klaret sig godt i internationale undersøgelser. Deres resultater har gennemgående været bedre end danske læsere. Derfor er artiklens syvende og sidste afsnit viet en kort sammenligning af dansk og finsk ortografi.²

Som det allerede er fremgået, anlægger jeg et funktionelt syn på den danske ortografi. Den betragtes her som en teknologi der muliggør grafisk repræsentation af talt sprog. Teoretisk set kan man naturligvis også vælge at betragte skriftsproget som et mere eller mindre autonomt semiotisk system (Cummings 1988; Harris 1995). Imidlertid er der mange forhold i praksis som taler mod autonomi i absolut forstand. Nogle få eksempler:

Læsetilgængelsen er langt vanskeligere for døve end for hørende.

Det er meget svært at lære en fremmed ortografi uden at lære det tilhørende talesprog samtidig.

Ordblinde kendetegnes typisk af problemer med at fokusere på sprogets lydige side.

Stavefejl rammer især ord hvor man ikke kan lytte sig frem til den korrekte stavemåde.

Såkaldt semasiografiske skriftsystemer (dvs. skriftsystemer hvis grafiske enheder ikke lader sig relatere til bestemte talesproglige størrelser) har altid et meget begrænset funktionsområde (Sampson 1985; DeFrancis 1989).

Alle disse forhold peger på at ortografibrugere faktisk i vid udstrækning relaterer skriftsproget til talesproget.

2. Identifikation af talesproglige enheder

Fordelen ved minimale enheder

De fleste børn der skal lære at læse, har nok begreb om visse isolerede sproglige enheder. De fleste kan opdele en sætning i ord, og ord i stavelser. Nogle kan måske også opdele stavelser i mindre dele, fx sådan at de kan isolere rimdelen i ordpar som *gris* og *spis*.

Men det er en meget krævende opgave for en begynder at forbinde ord, stavelser og rim direkte med skriftsproglige enheder. De skriftsproglige

enheder som svarer til ord, stavelser og rim, er jo grafisk meget komplekse – de kan bestå af adskillige bogstaver i en helt bestemt rækkefølge.

Hvis man foretager en yderligere opdeling af de talesproglige enheder i minimale lydlig enheder, bliver det derimod muligt at etablere *simple forbindelser* mellem talesprog og skriftsprog: én lyd – ét bogstav. Dermed bliver der også *flere holdepunkter* i omkodningen. Man kan tage én lyd eller ét bogstav ad gangen. Og tilmed er der langt *færre forbindelser* at tilegne sig på dette sproglige niveau. Antallet af lydlig mindsteenheder (fonemer) i dansk er meget mindre end antallet af stavelser.

Nu er formålet med skriftsproget jo imidlertid ikke at repræsentere rent lydlig information, i hvert fald ikke udelukkende. Det overordnede formål må være at repræsentere meningsfulde ytringer. Derfor er det naturligtvis også nødvendigt at forbinde skriftsproglige enheder med betydningsbærende talesproglige enheder. Etableringen af forbindelser mellem lydlig og grafiske enheder er først og fremmest vigtig fordi disse forbindelser kan støtte etableringen af forbindelser mellem betydningsbærende og grafiske enheder.

Når det gælder opdelingen af ytringer i betydningsbærende enheder, forekommer ordet at være den vigtigste enhed. Ordet afgrænses jo grafisk af ordmelletrum og modsvares således af en veldefineret grafisk enhed. Og at tage udgangspunkt i større betydningsbærende enheder (fx syntaktisk sammenhørende ordgrupper) ville være håbløst alene fordi antallet af mulige ordkombinationer er uendeligt. Derimod gælder det igen, at der kan være fordele ved at foretage en yderligere opdeling i *minimale* betydningsbærende enheder (morfemer): simple forbindelser (når man ser bort fra den manglende grafiske afgrænsning), flere holdepunkter i omkodningen, og alt i alt færre forbindelser at tilegne sig.

Konklusionen er at evnen til at opdele det talte sprog i minimale lydlig eller betydningsbærende enheder giver nogle logiske fordele. Hermed er ikke sagt at *al* omkodning mellem skrift og tale i praksis er baseret på minimale enheder, eller at omkodning på dette niveau *altid* er at foretrække (se afsnit 6). Det er blot påpeget at det kan være uhyre væsentligt, både for begyndere og viderekomne, at *have adgang* til forbindelser mellem tale- og skriftsproglige enheder som er så små som muligt.

Uanset på hvilket niveau man slutter forbindelserne mellem tale og skrift, så er der tre generelle problemer som man skal klare. For det første er der selve segmenteringen. Den talesproglige enhed som skal forbindes med en grafisk enhed, må isoleres fra sin sammenhæng.

For det andet må denne enhed skelnes fra andre enheder som kunne forekomme i samme sammenhæng. Dette kalder jeg differentieringsproblemet.

For det tredje kan analysen af sproglige enheder ofte foregå på flere forskellige abstraktionsniveauer. Her kan problemet være at få adgang til netop det abstraktionsniveau som muliggør de simpleste forbindelser til mellem tale- og skriftsproglige enheder.

I det følgende ser jeg på hvert af disse tre problemer i forhold til hhv. lydlig (fonologiske) og betydningsbærende (leksikalske) enheder.

Segmenteringsproblemet

Fonologisk. Bogstavet *g* i *gris* svarer til en *g*-lyd. Det er så enkelt. Men det er det jo kun hvis man kan isolere *g*-lyden fra resten af ordet. Og dét er bestemt ikke lige let for alle. For nogle er det faktisk svært. Problemet kan iagttages hvis man beder børn i førskolealderen om at sige *gris* uden *g*-lyd. Det er næppe alle der uden videre når frem til resultatet *ris*. Og denne segmenteringsfærdighed tjener jo heller ikke noget formål i sig selv. *G*-lyden er jo ganske meningsløs. Og betydningsmæssigt har ordet *gris* hverken noget at gøre med *ris* eller *is*.

Det fonologiske segmenteringsproblems betydning er dokumenteret meget grundigt af læseforskere gennem de sidste par årtier. Førskolebørns evne til at opdele ord i minimale lydlig enheder har gang på gang vist sig at være en stærk prædikator af deres senere læseudvikling (Scarborough 1998). Også voksne der har oplevet læsevanskeligheder kan fortsat have relativt svært ved at segmentere ord fonologisk (Elbro, Nielsen et al. 1994).

Leksikalsk. Leksikalsk segmentering kan dels være det at opdele ytringer i enkelte ord, dels det at opdele ord i morfemer. Generelt er der ikke noget stort problem forbundet med at isolere de enkelte ord i en ytring. Dog kan der ofte ses uenigheder, hvis man beder børn om at tælle ordene i en sætning, især når det gælder grammatiske småord og sammensatte ord (Elbro 1990).

At segmentere ord i morfemer er ikke så stor en udfordring som at segmentere dem fonologisk. Morfemer har en betydningside som man kan hænge dem op på, og deres lydside er oftest lettere at udtale i isolation end et enkelt lydsegment er det. Men blandt andet fænomenet schwa-assimilation kan gøre det svært at vide hvor det ene morfem ender og det næste begynder. I normal udtale assimileres den tryksvage vokal schwa

til nabokonsonanterne, sådan at fx *kvinden* ender på en lang n-lyd og *tamme* på en lang m-lyd. Grænsen mellem stamme og bøjningsendelse bliver flydende.

I modsætning til hvad der er tilfældet med ord og lydsegmenter, er der ikke nogen velafgrænset grafisk enhed som svarer til morfemer. Men de principielle fordele ved at forbinde minimale enheder kan naturligvis også høstes her. Har man fx én gang lært stavemåden *kant*, kan kendskabet til denne bogstav-morfem-forbindelse genudnyttes når morfemet indgår i morfologisk komplekse ord: *kanterne*, *kantede*, *trekantsdrama*, *firkantet*, *udkanten*, *kantspiller* osv.

Specielt identifikationen af bøjnings- og afledningsmorfemer kan være vigtig. Den kan gøre det lettere at identificere de tilbageværende stammer, og den kan lette etableringen af tekstsammenhæng. Den kan også være vigtig i stavning, fordi danske ord som bekendt ofte staves i overensstemmelse med deres morfologiske opbygning. Hvis man eksempelvis skal lære at bruge stavemåden *sødt* frem for den fonologisk mere ligefremme stavemåde **søt* kan det være en hjælp at opdele ordet i stammen *sød* plus bøjningsendelsen *-t*

Også når det gælder morfologisk segmenteringsfærdighed er der iagttaget sammenhænge med læseudvikling (Elbro 1990). De er dog ikke så stærke som sammenhængene med fonologisk segmenteringsfærdighed, og det er ikke ganske klart, hvad der er årsag, og hvad der er virkning.

Differentieringsproblemet

Fonologisk. Det er ikke nok at kunne opdele ord i segmenter. Hvis man skal vide hvilket bogstav hvert segment skal forbindes med, må man naturligvis også kunne skelne de enkelte lydsegmenter fra andre lydsegmenter som kunne forekomme i samme sammenhæng. For ordet *læser* må man kunne skelne l-lyden fra andre mulige initiale konsonanter, fx n-lyd i *næser*; man må kunne skelne den lange æ-lyd fra den lange a-lyd i *låser*, s-lyden fra b-lyden i *løber*, og »nå-lyden« (den tryksvage bagtungevokal) fra schwa (den tryksvage mellemtungevokal) i *læse*. Den lange trykstærke vokal må desuden skelnes fra den tilsvarende korte vokal i *læsser*, idet denne distinktion markeres indirekte (ved fordobling af det følgende konsonantbogstav).

Nødvendigheden af at kunne skelne lydligte forskelle som de nævnte kan måske være så indlysende at det forekommer omsonst at nævne den. Og er det ikke misvisende at tale om et *problem* her? I praksis er der jo sjældent nogen, hverken voksne eller børn, som har besvær med at kende

forskel på minimale ordpar som disse, hvis ellers hørelsen og akustikken er normal. Men faktisk *er* der flere gange fundet en behersket, men signifikant sammenhæng mellem såkaldt »fonemdiskrimination« og læseudvikling (Elbro 1990).

Forklaringen kan være at evnen til i praksis at udnytte en sproglig distinktion ikke med nogen nødvendighed ledsages af en bevidsthed om samme distinktion (Gombert 1992). Det er velkendt at studerende som skal lære at skrive lydskrift næsten altid har svært ved at skelne sikkert mellem stød og fravær af stød, som ortografien jo ikke har vænnet dem til at hæfte sig ved. Det gælder selv om de samme studerende udmærket kan høre forskel på ord som *mor* og *mord*, der lydligt kun adskilles af stødet.

I denne sammenhæng er det også værd at bemærke at markering af vokallængde er en af de klassiske hurdler i stavetilegnelsen. Det gælder i øvrigt ikke kun i dansk ortografi (Nauclér 1980; Jacobsen 1994). Det kan ikke udelukkes at differentieringsproblemet er en del af forklaringen.

Leksikalsk. At kunne skelne forskellige ord og morfemer fra hinanden er naturligvis lige så nødvendigt som at kunne skelne lydlig størrelser – og her er problemet muligvis helt marginalt. I stavning er der dog nogle fejltyper som til dels kan skyldes problemer med leksikalsk differentiering. Når morfemer med en abstrakt betydningsforskel har lydligt sammenfald, er det et velkendt problem at holde dem ude fra hinanden. Det gælder ikke mindst infinitivformer som *køre* over for præsensformer som *kører* (Juul og Elbro under forberedelse).

Også skriftsproglige distinktioner der afspejler forskelle mht. ordklasse og etymologi volder problemer. Tre eksempler: Proprier som *Steen* skal have stort begyndelsesbogstav, i modsætning til substantiver som *sten*. Den danske *ind-* skelnes fra det latinske/engelske *in-*, sådan at man skriver *indføre* med *d*, mens der *ikke* skal være *d* i **indfluere* og **indput*. Ordet *shampoo* er lånt fra engelsk og skrives derfor med *sh*, mens *champagne* er fransk og staves med *ch*.

Abstraktionsproblemet

Fonologisk. En opdeling af et ord i minimale lydlig enheder giver ikke altid mulighed for at etablere simple bogstav-lyd-forbindelser. En almindelig udtale af ordet *larve* kan opdeles i tre fonetiske segmenter. En l-lyd med den simple modsvarighed *l*, en lang åben a-lyd som svarer til *ar* og en kort u- eller o-agtig vokal som svarer til *ve*. Imidlertid kan det lade sig gøre at analysere begge de vokaliske segmenter som manifestationer af

to *abstrakte* lydligge segmenter (eller: underliggende fonemer, se Grønnum 1998). På dette abstraktionsniveau svarer skriftens *r* til et selvstændigt fonem /r/ som i udtalen er smeltet sammen med et foregående /a/ til den lange åbne a lyd. Tilsvarende svarer skriftens *e* til /ə/ (schwa) som er smeltet sammen (ved schwa-assimilation) med et foregående /v/-fonem til den korte u-agtige vokal. Hermed kommer hvert af de fem bogstaver i ordet til at svare til hvert sit abstrakte segment.

Den der kan slutte forbindelser på dette abstrakte niveau, vil formentlig have lettere ved at overskue forbindelserne mellem skrift og lyd, end den som forbinder skriften direkte til det mere konkrete fonetiske niveau. Men det er absolut ikke givet at alle har samme adgang til viden om de til tider ret abstrakte fonologiske strukturer som ortografien kan siges at afspejle (Juul 1999).

Valget mellem alternative abstraktionsniveauer (og der kan være adskillige muligheder) kan således udgøre et problem i sig selv. Dette understreges af at reglerne for hvordan det konkrete og det abstrakte niveau er relaterede, kan være meget udviklede. Det er således ikke bare sådan at en lang åben a-lyd *altid* skal analyseres som underliggende /ar/, og en kort u-lyd altid som /və/.

Leksikalsk. Også ortografiens afspejling af ords *morfologiske* struktur kan forekomme mindre forvirrende hvis man har adgang til en morfologisk analyse på et vist abstraktionsniveau.

For det første er der morfemer hvis betydning simpelt hen bare er abstrakt. Det er fx ikke helt let at udpege det semantiske fællestræk for verber i infinitiv. Og hvis man ikke forstår hvad en infinitiv er, har man ikke så meget ud af at vide at den næsten altid ender på *-e*.

For det andet er der tilfælde hvor samme morfem udtales forskelligt i forskellige sammenhænge. Her bliver problemet at se igennem variationen til den fælles betydning og de tilbageværende fonetiske ligheder. Det er næppe alle fødselsdagsbørn der gør sig klart at en *lagkage* består af *lag* – og at ordets første stavelse derfor ikke staves *lav*. For nogle børn er det en slem overraskelse at få at vide at den første stavelse i *leverpostej* (udtalt »levpostej« eller »lerpostej«) er en reduceret udtale af ordet *lever*.

Endelig kan det også været et problem at en og samme udtale kan repræsentere forskellige morfemer. Der er ikke noget at sige til det, hvis et barn fortolker førsteleddet i *lagkage* som morfemet *lav* – ud fra den betragtning at kagen ikke er særlig høj. Eller man kan tro at sidsteleddet i *leverpostej* må være *steg*.

I et eksperiment jeg gennemførte i 1998 var abstraktionsproblemet ganske tydeligt. Deltagerne skulle danne nye pigenavne i analogi til navnepar som *Anders-Andersine*, *Mikkel-Mikkeline* og *Josef-Josefine*. Stort set alle deltagere var i stand til at segmentere pigenavnene morfologisk og genbruge endelsen i nye pigenavne: *Adolfine*, *Ulrikine* osv. Overraskelsen var at der ikke var enighed om endelsens fonologiske form. Mange dannede navne som *Adolfsine* og *Ulrikfine*. De skrevne former *Ander-sine*, *Mikkeline*, *Josefine* passer jo med en konstant formel: stamme + *-ine*. Men denne analyse viste sig at være en abstraktion som ikke alle nåede frem til (Juli 1999).

Abstraktionsproblemet bliver ikke mindre af at ord, og ikke mindst deres bøjningsendelser, forekommer i et utal af mere eller mindre reducerede udtaleformer. Der er stor afstand fra den distinkte udtale af et udtryk som fx *sådan én* til den reducerede og helt normale udtale der lyder som »sneen«. Det at forbinde bestemte betydninger med bestemte fonologiske strukturer kan således være et problem i sig selv på grund af den betydelige morfo-fonetiske opacitet der kendetegner dansk (Bleses 1998).

3. Identifikation af skriftsproglige enheder

Tredelingen af de principielle vanskeligheder som blev benyttet ovenfor, kan også anvendes når det gælder identifikationen af de skriftsproglige (grafiske) enheder som de talesproglige enheder skal forbindes med.

Segmentering. Alle skriftsystemer gør bestemte sproglige størrelser lettere tilgængelige ved at standardisere de grafiske størrelser der svarer til dem. I alfabetiske skriftsystemer er der to standardiserede grafiske størrelser: bogstavet, der generelt svarer til netop ét lydsegment, og ordet – afgrænset af ordmellelrum – der generelt svarer til netop ét ord. Når det gælder om at dele skriften op i overkommelige bidder, er det derfor let at identificere grafiske enheder der svarer til lydsegmenter og til ord.

Når det gælder identifikationen af grafiske modsvarigheder til lydsegmenter, er der imidlertid nogle velkendte fælder. Nogle lydsegmenter svarer til kombinationer af bogstaver. En velær nasal svarer fx normalt til *ng* som i *sang*. Og nogle bogstaver modsvarer slet ikke noget lydsegment, som fx det stumme *d* i *sand*.

Mere problematisk bliver segmenteringen af skriften hvis enhederne skal modsvare andre sproglige enheder end lydsegmenter og ord. Færdighed i at opdele ord i stavelser eller morfemer kan muligvis være nyttig for ortografibrugere – men ortografiens grafiske udformning tilbyder ikke megen støtte til denne opgave.

En konsekvens af denne manglende afgrænsning er at man i visse tilfælde ikke kan fastslå hvilken morfologisk segmentering der er den rigtige – jf. flertydigheder som *træ-kvinden* vs. *træk-vinden*. Problemet er måske størst hvor flertydigheden involverer bøjningsendelser. I et ordpar som *kanterne* vs. *lanterne* kan man fx ikke umiddelbart se om *-erne* hører til stammen, eller om der er tale om en bøjningsendelse. I ord som *Hans* vs. *Jans* kan man ikke se om *s*'et er en genitiv eller ej. Det er muligt at den almindelige brug af (uautoriserede) stavemåder som *Jan's fodbold støvler* for *Jans fodboldstøvler* ikke udelukkende skal ses som en afsmitning fra engelsk ortografi. Måske afspejler de et behov for at markere morfologiske grænser.

Differentiering. Det indebærer en oplagt fordel hvis man hurtigt og sikkert kan genkende ord som grafiske helheder – altså uden nødvendigvis at bruge tid på identifikationen af hvert enkelt bogstav. Den grafiske udformning af ord i alfabetiske skriftsystemer modarbejder imidlertid denne form for ordgenkendelse. Alfabetiske skriftsystemer har kun en god snes grafiske mindsteenheder (bogstaver) at sætte sammen til ordbilleder. Bogstaverne må endda kun arrangeres i én dimension i forhold til hinanden (den vandrette), og mange kombinationer er udelukkede på forhånd fordi de tilsvarende lydsekvenser er umulige. Det betyder at ordbillederne for en overfladisk betragtning fremstår meget ens – »monotone« med en kinesisk iagttagers ord (Sampson 1985). Mange ord adskiller sig minimalt fra hinanden (fx *brik* vs. *birk*, *smøre* vs. *snøre*). De bliver svære at skelne sikkert uden en detaljeret analyse fordi mængden af redundant information er så begrænset.

På bogstavniveau kan der især være usikkerhed omkring bogstaver der er hinandens spejlbilleder: *b*, *d*, *p* og *q*.

Abstraktion. Også for bogstavernes vedkommende kan der være fordele ved at være opmærksom på mere abstrakte elementer. Der er jo en betydelig variation i bogstavets udseende, især når man ser på håndskrifter. Her kan det være vigtigt at være opmærksom på de grafiske bestanddele der går igen (buer, løkker, overlængder etc.).

Der er nogle få eksempler på at bogstavers grafiske fællestræk kan relateres til fonetiske fællestræk ved de modsvarende lydsegmenter: Bogstaver der modsvarer vokaler har normalt hverken over- eller underlængder (eneste undtagelse er y). Bogstaver med kombinationen løkke plus over- eller underlængde står normalt for lukkelyde (*pbdgq*). Og bogstaver som holder sig inden for x-højden og har bue eller hak, står for stemte kontinuanter (*mnuvw*). Der findes et skriftsystem (det koreanske) hvor man kan komme betydeligt længere ad denne vej (Sampson 1985, DeFrancis 1989). Det er dog uvist hvor stor en rolle denne mulighed spiller i praksis.

4. Flertydige enheder

Hvis man er i stand til at identificere og forbinde talesproglige og grafiske enheder, så har man forudsætningerne i orden. Men den praktiske opgave det er at omkode enheder af den ene slags til enheder af den anden slags, er naturligvis ikke løst dermed. Det kræver øvelse at udnytte sin viden om forbindelserne mellem tale og skrift, sådan at man umiddelbart kan slutte fra en given enhed til dens tale- eller skriftsproglige modsvarighed.

Her spiller det naturligvis en rolle hvor tit man i praksis har brug for at udnytte sin viden om en bestemt forbindelse. Specielt på det leksikalske niveau er der jo mange enheder som man ikke har brug for at omkode særlig tit. Færdighed i at opdele ord i mindre enheder kan derfor være særlig vigtig når det gælder mindre hyppige ord.

Færdighed i at slutte fra en given enhed til dens modsvarighed kan imidlertid også støde på en mere principiel vanskelighed. Det er jo velkendt at forholdet mellem tale- og skriftsproglige enheder ikke altid er entydigt. I mange tilfælde kan en enhed have flere mulige modsvarigheder, og dermed bliver omkodningen behæftet med en vis usikkerhed: Hvilken af de mulige modsvarigheder er den rigtige i sammenhængen?

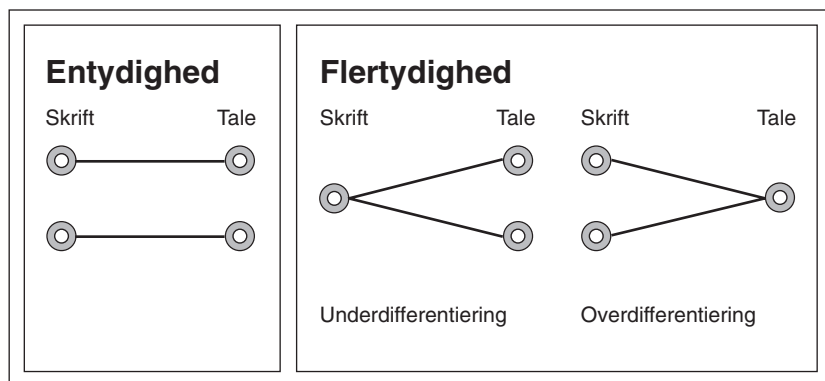
Dette problem forekommer i begge omkodningsretninger og kan ramme både fonologiske og leksikalske enheder.

Flertydighed i retningen fra skrift til tale

Når der til samme skriftsproglige enhed kan svare flere forskellige talesproglige enheder, kan man sige at ortografien er *underdifferentierende*. En forskel mellem talesproglige enheder modsvarer ikke af en forskel

mellem de tilsvarende grafiske enheder (se Figur 1). To eksempler på fonologiske flertydigheder:

Bogstavet *g* udtales som en *g*-lyd i *smagt* og *Holger*, men som en *j*-lyd i *spagt* og *Helge*. Tilsvarende udtales bogstavet *y* som en *y*-lyd i *lytte*, men som en *ø*-lyd i *flytte*.³



Figur 1. Flertydighed giver en principiel usikkerhed i omkodningen mellem skrift- og talesproglige enheder. Det er ikke tilstrækkeligt at identificere en flertydig enhed. Man må også vælge mellem dens mulige modsvarigheder. Når en skriftsproglig enhed kan have flere modsvarigheder, kan man tale om underdifferentiering. Skriften ignorerer en forskel mellem talesproglige enheder. Når en talesproglig enhed kan have flere modsvarigheder, kan man tale om overdifferentiering. Skriften repræsenterer én og samme talesproglige enhed på flere forskellige måder.

Underdifferentiering er især læserens problem. Den gør det sværere at nå frem til ordets præcise lyd, og dermed kan ordet blive sværere at identificere. For staveren er det ikke noget principielt problem at lade være med at markere en distinktion, fx sådan at *ø*-lyde og *y*-lyde noteres med samme bogstav. Men staveren kan alligevel få et problem, af den grund at flertydigheden ikke er *generel*. Således markeres distinktionen mellem *y*-lyd og *ø*-lyd jo faktisk i ord som *bytte* vs. *bøtte*. Resultatet bliver en flertydighed i retningen fra tale til skrift: *ø*-lyd modsvarer nogle gange af *y*, andre gange af *ø*. Flertydighed i denne retning vender jeg tilbage til nedenfor.

Det skal bemærkes at der i eksemplerne her ikke umiddelbart er noget i bogstavernes kontekst der kan afsløre hvilken udtalemulighed der er den rigtige. Heldigvis er det dog ofte sådan at valget mellem alternativene kan afgøres ud fra konteksten. Det måske vigtigste eksempel på dette er distinktionen mellem kort og lang vokal. En længdeforskel markeres

som bekendt ikke på selve vokalbogstavet, men den kan normalt udledes alligevel, nemlig ud fra antallet af følgende konsonanter: *klasse* har en kort a-lyd, mens *klase* har en lang a-lyd.

Det er altså sådan at flertydigheder nogle gange kan overvindes hvis man udvikler mere avancerede omkodningsstrategier – i det aktuelle tilfælde hvis afkodningen af vokaler tager højde for konteksten. Avancerede omkodningsstrategier er et emne jeg vender tilbage til i afsnit 6. Men selv for den der tager højde for konteksten, er der tilfælde hvor det ikke er entydigt om vokalen skal være kort eller lang. Eksempelvis har mange danskere forskellig vokallængde i følgende ordpar, selvom vokalerne ortografiske kontekst er nøjagtig den samme: *Kina* vs. *Nina*, *rype* vs. *type*, *fable* vs. *rable*, *bedt* vs. *fedt*, *otteren* vs. *etteren*.

At flertydighed kan være et reelt problem i afkodningen, fremgår tydeligst når det gælder uregelmæssige ord, dvs. ord der har én eller flere usædvanlige bogstav-lyd-forbindelser. For begyndere er uregelmæssige ord som *de* og *det* oftest sværere end regelmæssige ord som *du* og *den*. Men også hos voksne kan man se en effekt af uregelmæssighed, når det gælder mindre hyppige ord. Afkodningen af uregelmæssige ord er lidt langsommere end afkodningen af regelmæssige ord med samme hyppighed (Parkin 1984).

Et andet indicium på de praktiske konsekvenser af flertydighed kommer fra en interessant sammenligning af tyske og engelske børns afkodning. Her blev jævnaldrende børn bedt om at læse ord som i tysk og engelsk har næsten samme lydlig struktur, næsten samme betydning og næsten samme hyppighed – eksempelvis tysk: *Brot* og engelsk: *bread*. De engelske børn lavede langt flest fejl. En sandsynlig forklaring er at de engelske stavemåder bød på langt flere flertydigheder end de tyske (Landerl, Wimmer et al. 1997).

Ser man på det leksikalske niveau, er der flertydighed i tilfælde af homografi: *tunge* kan være et adjektiv eller et substantiv, *bløde* et adjektiv eller et verbum, *koste* et substantiv eller et verbum osv.

Der kan også være tale om »indlejrede homografer«: *sol* svarer til en selvstændig leksikalsk enhed i *solskin*, men ikke i *soldat*. Andre eksempler er *mos* i *abrikosmos* vs. *mikrokosmos* og endelsen *-erne*, som er en bøjningsendelse i *kanterne* og *vanterne*, men ikke i *lanterne* og *havterne*. Konsekvensen er også her at man i omkodningen fra skrift til tale kan få problemer hvis man går mekanisk frem. Tilstedeværelsen af pseudomorfer som *sol* i *soldat* kan virke forstyrrende (Hansen 1983; Elbro 1990).

Flertydighed i retningen fra tale til skrift

Det modsatte af underdifferentiering er *overdifferentiering*. Ortografien er overdifferentierende når der til samme talesproglige enhed kan svare flere forskellige skriftsproglige enheder (se Figur 1). Ortografien har i dette tilfælde indført en skelnen som ikke modsvarer en forskel mellem talesproglige enheder. Det betyder at repræsentationen af én og samme talesproglige enhed er ikke konstant.

Hvor underdifferentiering er læserens problem, er overdifferentiering staverens problem. Hvis samme talesproglige enhed kan repræsenteres på flere måder, hvilken skal man så vælge at bruge?

Igen nøjes jeg med at give nogle få eksempler, fænomenet er jo velkendt. Det er svært at stave! Til en simpel s-lyd kan der svare mange forskellige bogstaver eller bogstavkombinationer: *sekund, cerut, zebra, scene, xylofon, psykolog* etc. En kort e-lyd kan modsvarer af *e* som i *fedt*, af *i* som i *midt* eller måske af *et* som i *det* og *filet*. Og så videre og så videre.

Også når det gælder den indirekte markering af vokalers længde kan man finde anledning til at blive usikker: *Otte* har lang vokal ligesom *flåte*, men skrives alligevel med dobbelt *t*. *Cykel* har kort vokal ligesom *fakkel* men skrives med enkelt *k*. Her er der dog tale om ret alenestående undtagelser. Når det derimod gælder konsonantfordobling efter (kort) tryksvag vokal kan selv kompetente staver få brug for en ordbog – fx for at fastslå at der er dobbelt *r* og enkelt *s* i *karrusel*, mens der i *kolossal* er enkelt *l* og dobbelt *s*.⁴

Leksikalsk kan man tale om overdifferentiering i to tilfælde. For det første kan et og samme ord have mere end én tilladt stavemåde, fx alternativerne *majonæse* og *mayonnaise*, *ledeligt* og *leddegigt*. Her kan man som staver jo blot vælge den form man bedst kan lide, så umiddelbart er der ikke nogen vanskelighed. I tilfældet *led(d)eligt* kan man så endda vælge den form der passer bedst med ens egen udtale, for her er der udbredt variation mht. længden af vokalen i *led-*. På den anden side har man ikke et tilsvarende valgfrihed mellem fx *gedeost* og **geddeost*. Det kan således være uklart hvornår man må vælge selv, og hvornår man ikke må – et problem der måske især er generende hvis man gerne vil korrigerer andres stavfejl!

For det andet kan samme leksikalske enhed have forskellig stavemåde i forskellige morfologiske sammenhænge:

bus + *-en* = *buss-en*, ikke *bus-en*.

stiv + *-t* = *stift*, ikke *stiv-t*

fire + *-ten* = *fjorten*, ikke *fireten*

Variationerne her afspejler som oftest variationer i den leksikalske enheds fonologiske form i forskellige sammenhænge (allomorfi). Konflikten mellem leksikalsk motiveret konstans og fonologisk motiveret differentiering vender jeg tilbage til efter næste afsnit.

Differentiering og konstans

Ovenfor har jeg forsøgt at vise hvordan fænomenerne underdifferentiering og overdifferentiering kan gøre omkodningen vanskeligere. Men det bør nævnes at disse fænomener også kan have deres fordele:

Det er måske nok ærgerligt at den trykstærke vokals længde ikke kan aflæses i ord som *fable* vs. *rabble* – men til gengæld sparer staveren jo så besværet med at markere længdeforskellen. Og læseren får jo nok alligevel identificeret ordene til sidst. Det er måske nok besværligt at s-lyden i *psykolog* skal have et stumt *p* – men til gengæld bliver det tydeligere at ordet ikke er af dansk (men derimod græsk) oprindelse. Desuden bliver ordet lettere at genkende når det forekommer i fremmede ortografier som jo oftest også har bevaret *p*'et. Og alt i alt bliver ordet jo mere forskelligt fra andre ord med initial s-lyd – hvilket kan gøre ordet lettere at genkende hurtigt.

Der kan altså udmærket gives argumenter for at under- og overdifferentiering tjener fornuftige formål. Men kun til en vis grænse. Under- og overdifferentiering bryder nemlig med skriftsprogets to helt basale principper: differentieringsprincippet og konstansprincippet.

Underdifferentiering er et brud på differentieringsprincippet, som siger at forskellige ytringer skal repræsenteres forskelligt. Hvis en ortografi var maksimalt underdifferentierende, så ville den slet ikke afspejle nogen talesproglige distinktioner overhovedet. Og hvad kan man bruge en ortografisk repræsentation af en ytring til, hvis man den ikke adskiller sig fra repræsentationen af en hvilken som helst anden ytring? Ingenting, naturligvis. Derfor kan man sige at en ortografi bliver stadig mindre egnet til at tjene sit overordnede formål, at repræsentere ytringer grafisk, jo færre forskelle mellem ytringer den afspejler.

Overdifferentiering er et brud på konstansprincippet, som siger at samme talesproglige enhed skal repræsenteres på samme måde hver gang den

forekommer. Hvis en ortografi var maksimalt overdifferentierende, så ville repræsentationen af en bestemt enhed aldrig være den samme i to forskellige sammenhænge. Derfor ville man heller ikke kunne generalisere viden om en bestemt forbindelse mellem tale og skrift fra én sammenhæng til en anden. Ortografien ville igen være blevet ubrugelig.

Spørgsmålet er så bare hvor grænserne går? Hvornår er under- og overdifferentiering faktisk hensigtsmæssigt? Og hvornår bliver konsekvenserne for ortografiens praktiske anvendelighed fatale? Her gælder det for ortografier som for alle andre redskaber at det er svært at udtale sig generelt om udformningens hensigtsmæssighed. Den afhænger af hvad redskabet helt præcist skal bruges til.

Det fonologiske vs. det leksikalske niveau

Et andet spørgsmål der rejser sig her, er hvorvidt brud på principperne om differentiering og konstans er lige alvorlige på det fonologiske og det leksikalske niveau? Dette spørgsmål vil jeg opholde mig lidt ved, fordi disse brud jo ikke altid kan undgås på begge niveauer samtidig.

I tilfælde af *allomorfi*, altså når samme morfem har forskellig fonologisk form i forskellige sammenhænge, så må ortografien »vælge« mellem en differentieret fonologisk repræsentation og en konstant leksikalisk repræsentation.

I tilfælde som *stiv-stift* (ikke **stivt*) og *fire-fjorten* (ikke **fireten*) tilgodeser stavemåden den fonologiske differentiering. Selv om formerne *stiv* og *stif-*, *fire* og *fjor-* i en eller anden abstrakt forstand repræsenterer de samme leksikalske enheder, så staves de alligevel forskelligt i overensstemmelse med udtalen.

Omvendt er der også tilfælde hvor det er den leksikalske konstans som er tilgodeset. Det gælder ordpar som *sød-sødt* (ikke **søt*), *bruge-brugt* (ikke **bråkt*), *seks-seksten* (ikke **sajsten*).

Den omvendte konflikt mellem niveauerne ses i tilfælde af *homofoni*. Her må ortografien »vælge« mellem den fonologiske konstans og den leksikalske differentiering.

Den fonologiske konstans er tilgodeset når fx ordet *tunge* skrives på samme måde, uanset om der er tale om et substantiv eller et adjektiv, eller når passivendelsen, fx i verbet *dannes*, skrives på samme måde som genitivendelsen i *Hannes*.

Derimod er den leksikalske differentiering tilgodeset når enslydende ordpar staves forskelligt, fx *hal* – *halv*, *sagfører* – *savfører*, *tjat* – *chat*, *skålene* – *skålende*.

For brugerne af dansk ortografi er problemet nok ikke så meget selve bruddene på principperne, men snarere dét at det ikke altid er indlysende *hvilket* princip det er der skal ofres. Som vist er det nogle gange det ene princip, andre gange det andet.

Ser man på hvilke brud der giver de største problemer i praksis, så er det afgjort bruddene på det fonologiske niveau. Fonologisk motiverede stavfejl som **sajsten* for *seksten* og **skålene* for *skålende* er ganske almindelige. Derimod er det tvivlsomt om en morfologisk motiveret fejl som **fireten* for *fjorten* overhovedet forekommer. Og afskaffelsen af den indtil 1948 gældende distinktion mellem substantiver som *Tunge* og adjektiver som *tunge*, kan vist kun betegnes som en stor succes.

At omtale leksikalsk differentiering som et egentligt princip kan i øvrigt siges at være misvisende. Vi har jo ikke indført det stumme *v* i *halv* for at gøre ordet forskelligt fra ordet *hal*. Begrundelserne for det stumme *v* er nogle helt andre, nemlig hensynet til traditionen, og det udtalte *v* i det afledte verbum *halvere*.

Når det gælder den leksikalske konstans er det derimod berettiget og almindeligt at tale om et princip (Hansen 1991). Men det skal bemærkes at konsekvenserne af brud på dette princip er begrænsede. To tredjedele af stammen *sød-* er trods alt bevaret i den forkert stavemåde *søt* – så forbindelsen mellem *sød* og *sø-* er ikke helt obskur. Når det derimod gælder brud på det fonologiske konstansprincip, så er der fx ikke noget fælles visuelt træk som kan afsløre at *tj* i *tjat* og *ch* i *chat* repræsenterer samme fonologiske størrelse.

Alt peger således på den konklusion at principperne om differentiering og konstans er vigtigere på det fonologiske niveau end på det leksikalske.

5. Skriften er fattig på information

Man kan nu forestille sig en læser som er i stand til at afkode helt korrekt både på det fonologiske og det leksikalske niveau. Vedkommende bliver præsenteret for denne sætning:

Ræven så pigerne gå hjem fra skoven

Pointen med dette tankeeksperiment er at vise at man ofte står tilbage med adskillige fortolkningsmuligheder, når man skal forbinde en skreven sætning med en helt specifik talesproglig ytring – også selv om man har identificeret hver eneste leksikalsk og fonologisk enhed helt korrekt.

I sætningen ovenfor er der ikke nogen ortografisk angivelse af

- om *Ræven* er et dyr eller et menneske (semantiske træk – fx »menneske« vs. »dyr« – repræsenteres ikke).
- om det er *Ræven* der ser *pigerne* eller omvendt (information om grammatiske roller som subjekt og objekt eller semantiske roller som agens og patiens repræsenteres heller ikke).
- om ordene *gå hjem* udgør et komplekst verbal (med enhedstryk) – eller om *gå* er et selvstændigt verbal (med emphatisk tryk og stød) (det angives ikke hvilke ord der hører sammen i fraser. Tryk og stød repræsenteres ikke systematisk. Emfase *kan* markeres vha. kursivering, men markeringen er ikke obligatorisk).
- om *Ræven* befandt sig i skoven, eller om det var derfra *pigerne* gik hjem (det angives ikke hvilke syntaktiske led der hører tættest sammen. Intonation og pausering repræsenteres heller ikke og kan derfor heller ikke afsløre hvordan sætningen skal fortolkes).

Man kan sige at der her er tale om en helt basal form for underdifferentiering. Forskellene mellem de nævnte fortolkningsmuligheder repræsenteres bare ikke, og sådan er dét – selv om det jo ikke ville være svært at finde på grafiske løsninger på problemerne. Men nu er det jo sådan at man også som afkoder af *talte* ytringer må tage stilling til mange af disse fortolkningspørgsmål. Og man kan mene at det er helt unødvendigt at belemre skriften med distinktioner som heller ikke er til stede i det fonetiske input vi får når vi lytter til talt sprog. Andre vil måske tværtimod spørge hvorfor skriftsproget absolut skal overtage talesprogets mindre heldige egenskaber?

Prosodisk underdifferentiering

Den basale underdifferentiering gælder imidlertid ikke kun distinktioner som er fraværende i det talte sprog. Når dansk ortografi ikke har midler til at repræsentere de prosodiske træk stød og tryk, så fører det til en reel og »ortografispecifik« flertydighed. I udtalen er der fx ingen tvivl om

hvorvidt ordet *mast* er et substantiv (uden stød) eller perfektum af verbet *mase* (med stød). (Første linje af kongesangen bliver sjældent afsunget uden at der er nogen på bageste række der skal more sig med en bevidst misfortolkning.)

Stød og tryk tjener også andre funktioner end adskillelsen af leksikalske størrelser. De fungerer også som vigtige signaler om syntaktisk og morfologisk struktur. Som læser må man analysere skrevne ord og sætninger uden at støtte sig til disse signaler. Og når en tekst skal læses højt, er det et problem for sig at få dem sat ind igen på de rigtige steder. Selv trænede oplæsere kan falde i »tryktabsfælden«:

Han fandt nogle gamle skotøjsæsker frem.

Han fandt nogle gamle skotøjsæsker på loftet.

I den første af disse to sætninger skal man helt hen til ordet *frem* før man kan vide at verbet *fandt* skal udtales *uden* tryk. Tilsvarende er det først når man har læst ordene *på loftet* at man kan vide at verbet i den næste sætning skal udtales *med* tryk.

Endnu et eksempel på basal underdifferentiering er den manglende markering af vokallængdeforskelle i sidste stavelse af et ord. Vokalen er lang i *hus* og *café*, men kort i *bus* og *René*, og dét må man virkelig selv regne ud.

Den manglende ortografiske afspejling af prosodiske distinktioner (stød-, tryk- og længdeforskelle) ligger måske på en måde i kortene. Det geniale ved det alfabetiske princip er jo at man ved at skelne fonologiske enheder på segmentniveau kan opnå en meget tilfredsstillende grad af differentiering også på det leksikalske niveau. Alfabetet er ikke skabt til prosodiske distinktioner.

Leksikalsk og segmental underdifferentiering

Imidlertid kan man jo *heller* ikke være sikker på at skriften afspejler alle relevante leksikalske og segmentale distinktioner. Det blev der givet eksempler på i afsnit 4. Her fokuserede eksemplerne på distinktioner som i princippet afspejles i skriftsproget – men som alligevel ikke bliver det. Men faktisk er der også her distinktioner som ignoreres *helt principielt*.

Det skyldes først og fremmest at ortografien – eller rettere ortografibrugerne – henholder sig til standardiserede ordformer. Denne standardise-

ring gælder til dels også på det leksikalske niveau. Der er eksempelvis (stadig) mange danskere som bruger nægtelsen *inte* og det personlige pronomen *a*. Men der er vist ikke mange der bruger disse ord når de skriver. Her bruger man de ord som hører det standardiserede skriftsprog til: *ikke* og *jeg*.

Også på det segmentalfonologiske niveau er der mange forskelle som ikke afspejles ortografisk. Det gælder udtaleforskelle som har at gøre med talerens alder og regionale/socialle baggrund såvel som forskelle betinget af det udtalemæssige distinkthedsniveau.

Problemet her er at der kan være ganske stor afstand fra den udtale man naturligt bruger til den udtale som skriften afspejler. Man kan ikke bare skrive som man taler: *Ska vik ta på bibloteke?* Omvendt kommer det let til at lyde unaturligt når skriftsproget skal læses højt: *Skal vi ik-ke ta-ge på bib-li-o-te-ke?*

Nogle finder ikke blot et praktisk, men også et ideologisk problem her. Skriften kan være med til at cementere den opfattelse at der findes én og kun én sprogform som er den korrekte (Kristiansen 1999). På den anden side kan man også mene at den ortografiske standard betyder at tekstens budskab står klarere. Standardiseringen tilgodeser den hurtige genkendelse af leksikalske enheder. Og når man skriver, bliver man ikke bedømt på sin sociale baggrund så snart man har taget ordet.

Endnu to former for segmental underdifferentiering skal kort nævnes.

1) Skriften repræsenterer ords segmentalfonologiske struktur på et abstrakt niveau (se afsnit 2). Det vil sige at der er mange konkrete fonetiske forskelle som ikke afspejles ortografisk. Bogstavet *l* kan repræsentere en stemt konsonant som i *lade* eller en ustemt konsonant som i *plade*. Bogstavet *r* kan repræsentere en konsonant som i ordet *ru* eller en halvvokal som i ordet *ur*. Denne form for underdifferentiering kan være hensigtsmæssig, netop fordi disse fonetiske forskelle ikke bidrager direkte til at adskille leksikalske størrelser fra hinanden. Men til gengæld bliver det vigtigt at man som ortografibruger er i stand til at abstrahere fra det konkrete fonetiske niveau.

2) Visse vokalforskelle bidrager faktisk direkte til den leksikalske adskillelse af ord – men af simpel mangel på vokalbogstaver er det ikke alle relevante distinktioner der kan afspejles ortografisk. Vi har tre forskellige rundede fortungevokalkvaliteter: *fyr!* (imp.) vs. *før!* (imp.) vs. *før* (adv.) – men vi har kun to bogstaver: *y* og *ø*. Yngre københavnere har fire distinktiøse forskellige bagtungevokaler i *tud* vs. *fod* vs. *båd* vs. *skod!* –

men vi har kun tre bogstaver: *u*, *o* og *å* (eksempler fra Grønnum 1998). Og de fleste har to forskellige *a*-lyde i *klandre* og *vandre* – men ortografisk har vi kun ét *a* (Brink 2000; Grønnum 2000).

6. Effektiviseret omkodning

Man skal ikke have lært så forfærdelig mange forbindelser mellem tale- og skriftsproglige enheder, før man kan få sine første succesoplevelser som ortografibruger. Når man først er blevet i stand til at etablere forbindelser, er meget af den videre udvikling et spørgsmål om at samle sig erfaringer: Hvilke tale- og skriftsproglige enheder kan forbindes med hinanden? Hvilke enheder er det vigtigt at kunne identificere? Jo flere forbindelser man kender, og jo oftere man har udnyttet sit kendskab til dem, jo hurtigere kan omkodningen mellem tale og skrift forløbe.

Når det gælder flertydigheder, spiller erfaringen også en vigtig rolle. I praksis er flertydighed ofte et ganske ubetydeligt problem fordi erfaringen fortæller os hvilken modsvarighed der er den mest almindelige. Og hvis denne mulighed ikke passer i sammenhængen, kan alternative omkodningsmuligheder som oftest også identificeres ud fra erfaringen.

Men det ville være en forenkling at sige at det at blive »avanceret ortografibruger« udelukkende er et spørgsmål om at øve sig. Man kan godt nævne nogle mere specifikke færdigheder der kan have betydning for hvor hurtigt og præcist omkodningen forløber i praksis. Det er færdigheder af denne slags der er emnet for dette afsnit. Fælles for de fleste af dem er at de har at gøre med en videreudvikling af den sproglige opmærksomhed.

Forbindelser på leksikalsk niveau

Rent kommunikativt kan omkodningen mellem tale og skrift sagtens lykkes selv om den foregår udelukkende på det fonologiske niveau. Men hvis omkodningen skal være hurtig og fejlfri, er det ofte nødvendigt at man kan forbinde skriftsproglige enheder direkte til det leksikalske niveau. Hvis man skal stave helt korrekt, må man vide at lige netop ordet *zebra* staves med *z* og ikke med *s* eller *c*. Det betyder ikke *nødvendigvis* at man må etablere en forbindelse på helordsniveau – blot at fonologiske uforudsigelige stavemåder, som *z*'et i *zebra*, må forbindes med de bestemte ord de forekommer i.

Noget tilsvarende må ske i afkodningsretningen. Det stumme *t* i *det* og

*e'*ets usædvanlige *i*-lyd i *de* kan være forvirrende for begynderen – men de fleste lærer efterhånden at de normale bogstav-lyd-forbindelser ikke gælder i lige netop disse ord. Alternativet kan igen være en forbindelse på helordsniveau, og det er måske det mest oplagte når det gælder korte og højfrequente ord som disse. Men mindre kan gøre det. Når det gælder *d'*erne i *det* og *de* modsvares de helt regelret af *d*-lyde – og der er ikke nogen principiel grund til at man skulle undlade at udnytte denne regelmæssige forbindelse i sin afkodning. Hvis der derimod slet ikke etableres nogen forbindelse til det leksikalske niveau, er der ikke andet at gøre end at gætte sig frem. Derved kan afkodningen blive både langsommere og mindre præcis. For den effektive omkodning er det leksikalske niveau altså ikke til at komme udenom.

Hvis man ser sagen fra den modsatte side, kan man spørge om ikke forbindelserne på fonologisk niveau mister deres betydning, hvis omkodning alligevel må tage udgangspunkt i leksikalske størrelser? Det gør de kun hvis de leksikalske størrelser behandles som helheder – og det er som netop vist ikke nogen nødvendighed. Hvorfor skulle man afskære sig fra at udnytte sin viden om fonologiske forbindelser, hvis denne viden kan gøre omkodningen hurtigere og mere præcis?

Hvis man forestiller sig at man i omkodningen så helt bort fra forbindelser på fonologisk niveau, så ville mange leksikalske variationer blive til de rene mysterier: Hvorfor skal *s'*et i *bus* fordobles i bøjningsformen *bussen*, når *s'*et i *lus* ikke skal fordobles i *lusen*? Hvorfor forandres vokalen i *skyd* i perfektum *skudt*? For øvrigt måtte man sikkert opgive inddelingen af bogstaver i konsonanter og vokaler, eftersom disse kategorier er fonologisk baserede. Det er karakteristisk at det selv for erfarne læsere er næsten umuligt at koble den fonologiske afkodning fra (Van Orden 1987).

Enheders kombinationsmuligheder

Når det gælder om at opdele ord i stavelser eller morfemer kan det være nyttigt at være opmærksom på de mindre enheders kombinationsmuligheder. Når man ser et ord med to stavelser som fx

franskrød

så kan stavelsesgrænsen kun gå mellem *k* og *b*. Det er indlysende at det må være sådan hvis man ved at dansk fonologi ikke tillader konsonantgrupper som *nskb* og *kbr* i samme stavelse.

Udgangspunktet for opdelingen kan naturligvis også være det grafiske

niveau. Her er det fx sådan at *sb* ikke forekommer først i en stavelse, selv om der ikke fonologisk er noget i vejen med konsonantgruppen [sb]. Derfor kan man hurtigt udlede at der må være en stavelsesgrænse mellem *s* og *b* i *busbane*. Tilsvarende kan man straks placere en morfologisk grænse i ord som *skovvej*, *havvand*, *tovværk* når man ved der aldrig forekommer to *v*'er efter hinanden i samme morfem.

Viden om mulige bogstavkombinationer kan også give vigtig feedback i stavning. Med denne viden kan man straks se at der er noget galt med stavemåder som fx **sbise*.

Leksikalsk kan viden om kombinationsmuligheder være nyttige når man skal læse et ord som fx

Hvalparken

Det er meget normalt at substantivet *parken* står som sidsteled i en sammensætning, ikke mindst i stednavne. Det er derfor oplagt at vælge den morfologiske analyse *hval-parken*. Det kan være praktisk at man straks kan fravælge den alternative analyse, *hvalp-arken*, fordi man ved at *arken* er et meget sjældnere sidsteled i sammensætninger.

Morfologisk opmærksomhed i bred forstand

Flertydigheder i retningen fra tale til skrift kan i mange tilfælde overvindes hvis man som staver er opmærksom på bestemte karakteristika ved det ord der skal staves. I mangel af en bedre betegnelse vil jeg her tale om morfologisk opmærksomhed i bred forstand.

Hvis man er i stand til at identificere ords grammatiske funktion, bliver det muligt at undgå en række af de klassiske stavefejl: præsensformer ender på *-r*, infinitiver på *-e*, præsens participier på *-ende*, etc. (se afsnit 2) (Juul og Elbro under forberedelse).

Hvis man er opmærksom på at et morfem kan have varierende udtaleformer (allomorfi), kan man tage dette som en advarsel om at stavemåden kan være bestemt af princippet om morfemkonstans (se afsnit 4). På trods af *w*-lyden er **lavkage* forkert fordi førsteledet er *lag*. Etc.

Hvis man identificerer ordklassen, kan man vide om ordet skal skrives med stort (nemlig hvis det er et proprium), eller måske med stumt *h* (nemlig hvis det er et spørgende pronomen).

Endelig hvis man identificerer et ord som et låneord, kan man tage dette som en advarsel om at der måske skal bruges usædvanlige stavemåder. Måske kan man endda vælge den rigtige stavemåde ud fra sin vi-

den om den långivende ortografi, eller ved at udvikle en intuition om ortografiske subsystemer. Denne viden kan naturligvis også være relevant når det gælder flertydighed i afkodningen. I franske lån udtales *ch* fx på samme måde som *sj*, mens *ch* i engelske lån udtales som *tj*.

Der er imidlertid mange faldgruber når man begynder at inddrage sin morfologiske viden i staveprocessen. Det er let at falde i med fejl som **Dansker* for *dansker*, **døede* for *døde*, **Jeres* for *jeres* og **shampagne* for *champagne*.

Udnyttelse af konteksten

Som nævnt i afsnit 4 kan flertydigheder nogle gange elimineres hvis der tages højde for den sammenhæng en enhed forekommer i. I afsnit 4 gav jeg markeringen af vokallængde som eksempel, men der kan gives mange flere eksempler.

I staveretningen kan man formulere den regel at b-, d- og g-lyd efter en s-lyd næsten altid modsvares af *p*, *t* og *k*: *spade*, *stade*, *skade* – ikke **sbade*, **sdade*, **sgade*. En lidt mere sofistikeret regel siger at sj-lyd efter en lang a-lyd normalt modsvares af *g*: *bandage*, *garage*, *bagage* – ikke **bandasje*, **garasje*, **bagasje*.

I afkodningsretningen er standardeksemplet *c*-reglen: Et *c* udtales som en s-lyd når det eksempelvis følges af et *y* eller et *i* som i *cykel*, *citron*, mens udtalen er en k-lyd før *o* og *a* som i *cola*, *café*. Tilsvarende er udtalen af *u* også forudsigelig i visse sammenhænge: u-lyd før *t* som i *dutte*, *sutte*, men en å-lyd før *m* som i *dumme*, *summe* (Juul og Elbro 2001).

Et sidste eksempel på en kontekstregel viser en interessant mulighed for interaktion mellem det leksikalske og det fonologiske niveau i afkodningen: Vokalen i enstavede ord på *-st* kan ofte udtales på to måder, med eller uden stød. Det gælder ord som *læst*, *vist*, *mast* og *øst* (Becker-Christensen 1988). Hvis man forestiller sig at afkodningen begynder fra højre (Hansen 1983), så vil man i nogle sammenhænge kunne begynde med at identificere det finale *-t* som en verbal bøjningsendelse. Når dette er tilfældet, er der stor sandsynlighed for at vokalen skal udtales *med* stød. Denne fonologiske information kan så bidrage til den endelige identifikation af det pågældende ord som leksikalsk helhed!

Man kan forestille sig at gevinsten ved at inddrage konteksten også kunne opnås på en anden måde, nemlig ved at etablere forbindelser mellem komplekse enheder. Den forøgede præcision ville jo også kunne opnås hvis man eksempelvis forbandt endelserne *-utte* og *-umme* direkte med

de modsvarende fonologiske endelser – frem for at tage udgangspunkt i *u*'et og se på dets kontekst.

For den erfarne læser er det oplagt at der kan være noget vundet ved at basere omkodningen på mere komplekse enheder. Jo flere trin man tager ad gangen, jo hurtigere kommer man ned ad trappen. Til gengæld bliver faren for at snuble større. Der er imidlertid den principielle fordel ved kontekstregler, at de indfanger lige netop den regelmæssighed som kan give den øgede præcision i omkodningen. Når det gælder vokallængde, så følger det fx af en helt generel kontekstregel at udtalen af en vokal skal være kort hvis den følges af en dobbeltskrevet konsonant plus *e* før ordgrænse. Hvis præcis den samme information om vokalens længde skulle gøres tilgængelig ved at man tog udgangspunkt i endelserne som komplekse helheder, så skulle den tilegnes for hver eneste forekommende kombination af vokal plus konsonant:

Endelsen *-utte* har kort vokal, *-umme* har kort vokal, *-asse* har kort vokal, *-amme* har kort vokal osv. osv. Resultatet ville være det samme, men vejen derhen ville være en hel del længere.

7. Sammenligning af dansk og finsk

Det er klart at mange af de ovenfor nævnte vanskeligheder ved den danske ortografi også findes i alle andre alfabetiske ortografier. Men ser man efter i detaljen, kan forskellene mellem to forskellige ortografier være ganske betydelige. Det vil jeg i det følgende forsøge at illustrere med en sammenligning af dansk og finsk.

Identifikation af enheder

Finsk har flere lange ord end dansk – og færre korte. Ifølge Korkeamäki (1997) er der kun ti finske ord der skrives med to bogstaver. Umiddelbart må man derfor formode at finske læsere har et større behov end danske for at opdele ord i mindre, mere overskuelige enheder. På den anden side har finner også et par fordele når det gælder om at segmentere ord:

Hvad angår leksikalske enheder, kan man altid finde et rod morfem længst til venstre i ordet, eftersom præfikser stort set ikke forekommer. Opdelingen af stavelser i lydsegmenter er en mindre udfordring i finsk fordi stavelsesstrukturen er simple – her er langt færre konsonantgrupper, og dermed færre bogstaver og lyde per stavelse. Og når bogstaver og

lyde skal forbindes, har finnerne den fordel at stumme bogstaver stort set ikke forekommer.

Finsk har et mindre inventar af funktionelt forskellige lydsegmenter (fonemer) end dansk. Følgelig er der i praksis også lidt færre bogstaver at holde styr på. Hvor danske arveord skrives med 24 forskellige bogstaver (*c, q, w, x* og *z* forekommer kun i låneord), kan finnerne nøjes med 21. Derimod spiller længdedistinktioner en større rolle end i dansk, eftersom de i finsk også er relevante for tryksvage vokaler og konsonanter.

Finsk ortografi synes også at stille mere beherskede krav til brugernes abstraktionsevne end dansk. Finsk har fx ikke nogen parallel til den danske schwa-assimilation, som er en af de vigtigste årsager til den store afstand mellem ordenes almindelige, reducerede udtaler og de distinkte udtaler som skriften afspejler. Også når det gælder om at identificere leksikalske enheder med fonologiske variationer, er behovet mindre. Finsk afspejler nemlig den fonologiske variation, frem for den leksikalske identitet. Hvis ordene *bage*, *bagt* og *bagværk* skulle staves efter finske principper, ville vi skrive noget i retningen af *baaje*, *bagt* og *bavværg*.

Flertydighed

På det segmentalfonologiske niveau kan hverken over- eller underdifferentiering siges at udgøre nogen alvorlig vanskelighed i finsk. Også når det gælder markeringen af prosodiske distinktioner er man godt hjulpet i finsk. Der er ikke noget der hedder stød. Trykkets placering ligger fast på første stavelse, og markering er derfor overflødig. Og hvad angår længdeforskelle markeres de direkte (lange segmenter skrives med dobbeltbogstav) og *uden* undtagelser.

På det leksikalske niveau forekommer der flertydigheder i finsk fordi ortografien konsekvent følger principperne om fonologisk differentiering og konstans, også i tilfælde af allomorfi og homofoni (se afsnit 4). Men dermed har man sandsynligvis valgt det mindste af to onder. Som tidligere påpeget udgør forekomsten af brud på principperne om fonologisk differentiering og konstans en af de særlige vanskeligheder ved dansk ortografi, og det samme gælder den danske ortografis inkonsekvens i tilfælde af allomorfi og homofoni.

Effektiviseret omkodning

Det leksikalske niveau er ikke nær så uomgængeligt for finske ortografi-brugere som for danske. Ganske vist er der også udtalevariation i finsk, sådan at man ikke altid bare kan stave ordene som man selv siger dem.

Men forudsat at standardudtalen er bekendt, er uforudsigelige bogstavlyd-forbindelser en stor sjældenhed.

Dermed bliver der også mindre brug for at udvikle færdigheder der kan gøre omkodningen mere præcis i tilfælde af flertydighed. Der er ikke som i dansk vundet noget særligt ved at benytte grammatisk baserede staveregler eller ved at udnytte konteksten i omkodningen.

Konklusion

Naturligvis er læsning og stavning på dansk som aktiviteter betragtet nært beslægtede med læsning og stavning på finsk. Men *helt* de samme aktiviteter er det ikke, og de er derfor heller ikke helt sammenlignelige. Spørger man hvilken ortografi der stiller de skrappeste krav til sine brugere, vinder den danske ortografi en klar sejr. Hvad enten det gælder om at identificere og forbinde enheder, om at fortolke flertydigheder, eller om at udvikle færdigheder der sigter på at effektivisere omkodningen mellem tale og skrift, så bliver konklusionen den samme. Som dansker skal man kunne lidt mere før man kan kalde sig avanceret ortografibruger.

Tak til Carsten Elbro og Matti Leiwo for nyttige kommentarer og oplysninger. Tak for økonomisk støtte til Statens Humanistiske Forskningsråd (#9602075) og til Europa-kommissionen (COST STY 98/4014).

Litteratur

- Becker-Christensen, C.: *Bogstav og Lyd. Dansk retskrivning og rigsmålsudtale*. København 1988.
- Bleses, D.: The role of input, productivity and transparency in Danish children's acquisition of past tense morphology, i: *Odense Working Papers in Language and Communication* 17, 1998.
- Brink, L.: Om den sproglige forskel mellem Hansen-a og Larsen-a, i: *Mål & Mæle* 23(1), 2000.
- Carney, E.: *A Survey of English Spelling*. London 1994.
- Cummings, D. W.: *American English Spelling. An Informal Description*. Baltimore 1988.
- DeFrancis, J.: *Visible Speech*. Honolulu 1989.
- Elbro, C.: *Differences in Dyslexia. A Study of Reading Strategies and Deficits in a Linguistic Perspective*. Copenhagen 1990.

- Elbro, C., I. Nielsen og D.K. Petersen: Dyslexia in Adults: Evidence for Deficits in Non-word Reading and in the Phonological Representation of Lexical Items, i: *Annals of Dyslexia* 44, 1994.
- Elley, W. B.: *How in the world do students read?* Hamburg 1992.
- Gombert, J. É.: *Metalinguistic Development*. London 1992.
- Grønnum, N.: *Fonetik og fonologi. Almen og dansk*. København 1998.
- Grønnum, N.: Ét kort a i dansk: en mere abstrakt analyse, i: *Mål & Måle* 23(2), 2000.
- Hansen, E.: *Skrift, stavning og retstavning*. København 1991.
- Hansen, P. M.: An orthography normalizing program for Danish, i: *ARIPUC* 17, 1983.
- Harris, R.: *Signs of Writing*. London 1995.
- Hertel, H.: *Den daglige bog*. København 1983.
- Jacobsen, B.: »The mystery of quantity«. Quantity as a Specific Spelling Problem, and probably a Reading Problem«, i: *Acta Linguistica Hafniensia* 27(1), 1994.
- Jensen, T. P. og A. Holm: *Danskernes læse-regne-færdigheder – i et internationalt lys*. København 2000.
- Juul, H.: Er der r i ordblind?, i: *Mål & Måle* 22(1), 1999.
- Juul, H. og C. Elbro: Når bogstavets lyd afhænger af sammenhængen, i: *Psykologisk Pædagogisk Rådgivning* 38(1), 2001.
- Juul, H. og C. Elbro: Learning to spell in a deep orthography. Under forberedelse.
- Korkeamäki, R.-L.: What can be learned about reading acquisition in the Finnish language, i: *Cross-Language Studies of Learning to Read and Spell*, v. C. K. Leong og R. M. Joshi. Dordrecht 1997.
- Kristiansen, T.: Om farerne ved en ortografi der ikke genspejler talesproglig variation: det danske eksempel, i: *Språkbrukeren – fri til å velge? Artikler om homogen og heterogen språknorm*, v. H. Omdal. Kristiansand 1999.
- Landerl, K., H. Wimmer og U. Frith: The impact of orthographic consistency on dyslexia: A German-English comparison, i: *Cognition* 63, 1997.
- Nauclér, K.: *Perspectives on misspellings. A phonetic, phonological and psycholinguistic study*. Malmö 1980.
- Parkin, A. J.: Redefining the regularity effect, i: *Memory & Cognition* 12(3), 1984.
- Sampson, G.: *Writing Systems*. London 1985.
- Scarborough, H. S.: Early identification of children at risk for reading disabilities, i: *Specific reading disability: A view of the spectrum*, v. B.K. Shapiro, P.J. Accardo og A.J. Capute. New York 1998.
- Seymour, P.H.K.: COST A8 Foundation Literacy Pilot Study. Upubliceret, Dundee 2000.
- Sommer, M., J. Lau og J. Mejding: *Nordlæs – en nordisk undersøgelse af læsefærdigheder i 1.-3. klasse*. København 1996.
- Van Orden, G. C.: A ROWS is a ROSE: spelling, sound, and reading, i: *Memory & Cognition* 15, 1987.

Noter

1. Citeret efter plakat afbildet i Hertel (1983).
2. Det må her bemærkes at Elley (1992) (den meget omtalte undersøgelse, hvor danske elever fra 3. klasse klarede sig forholdsvis ringe) foretog en femdeling af 15 deltagende sprog efter en vurdering af deres »fonetiske regelmæssighed«. Denne skalering viste *ikke* nogen entydig sammenhæng med resultaterne fra disse lande.
3. Se Becker-Christensen (1988) for en udførlig beskrivelse af bogstavers udtalemuligheder i dansk.
4. Når det gælder flertydigheder i omkodningen fra lyde til bogstaver, foreligger der desværre ikke nogen fyldestgørende oversigt over forholdene i dansk. Et forbilledligt engelsk værk er Carney (1994).