Ordrug og udtaleforandringer
Pharao, Nicolai

Published in:
NyS

Publication date:
2010

Document version
Peer-review version

Citation for published version (APA):
Ordbrug og udtaleforandringer

NICOLAI PHARAO


Udbredelsen af lydforandringer blandt sprogbrugere og i deres ordforråd står centralt i analyser af lydforandringsprocesser i både den sprogbrugsbaserede lingvistik og sociolingvistikken, og i begge tilgange er der fremsat hypoteser om faktorer i udbredelsesprocessen og disse implikationer for ordenes repræsentation i det mentale leksikon. Disse hypoteser vil jeg teste her ved at undersøge udviklingen i bortfald af [w] for syllabisk [ð] i københavnisk talesprog.
1 LYDFORANDRING OG ORD I HUKOMMELSEN

1.1 Sprogbrugsbaserede studier af lydforandring


1.2 Ordbrug og mental representation af udtale
bevarer en meget abstrakt repræsentation af ords udtale i det mentale leksikon, så opbevares alle detaljer af hver eneste forekomst af hver ordform direkte i det mentale leksikon i en eksemplar-baseret model. Reproductionen af en ordform konceptualiseres således som en sky af alle de huskede forekomster, de såkaldte eksempler, og denne sky vokser, hver gang ordformen høres eller siges. Reproductionerne kan således opdateres løbende i hver enkelt sprogruers hukommelse, og efterhånden som udtalen af en lydsekvens rutiniseres i højfrequente ord som følge af generelle principper for automatisering af ofte gentagede bevægelser (jf. ovenfor), vil representationen af højfrequente ord ændres, sådan at der bliver flere og flere forekomster i skyen, der ikke længere indeholder den oprindelige lydsekvens. For et ord som ’every’ er der simpelthen (mange) flere eksempler af formen [’evəri] end af formen [’evəri], og derfor bliver formen uden schwa mere almindelig. Lige er det, fordi ordene bruges ofte, at deres representation ændres hurtigere.

For at denne forbindelse mellem sprogrug og representation af udtale i det mentale leksikon skal kunne underbygges, er det imidlertid nødvendigt, at effekten af ordfrekvens, som blev vist på gruppenniveau med eksemplet på schwa-assimilation og -bortfald i amerikansk engelsk ovenfor, også findes hos enkelte talere og ikke bare er en effekt, der kan ses på gruppenniveau. Derfor vil jeg i min analyse af [w]-bortfald undersøge sammenhængen mellem ordfrekvens og tendens til [w]-tab ikke kun for de grupper, der indgår i undersøgelsen, men så vidt muligt for hver enkelt taler. Hvis der hos den enkelte taler kan vises en sammenhæng mellem ordfrekvens og [w]-bortfald, vil dette styrke hypotesen om, at frekvenseffekten afspiller det mentale leksikons struktur, og at dette mentale leksikon netop består af et enormt lager af mange næsten ens udgaver af samme ord, snarere end et lille, invariant lager af abstrakte former med kategoriske manifestationsregler. Individens udtaleadfærd er imidlertid også interessant for en forståelse af, hvordan man kan undersøge igangværende lydforandringer.
1.3 Individer og lydforandringer


Men det er ikke sikker, at sprogbrugeres udtale forbliver stabil gennem hele livet. Flere og flere undersøgelser tyder på, at i hvert fald en del af befolkningen i ethvert sprogsamfund fortsætter med at deltage i igangværende lydforandringer op gennem voksenlivet (se Sankoff (2006) for en oversigt og Gregersen, Maegaard & Pharao (2009) for et dansk eksempel). Så samtidig med at ordformer kan være ustabile, dvs. variable mht. en given lydforandringsproces, så kan den enkelte sprogbruger også være ustabil og fortsætte med at ændre udtale, også efter
han eller hun er blevet voksen. Hvorvidt en persons udtale forbliver stabil eller er i fortsat forandring, kan kun undersøges ved at studere den samme taler flere gange gennem livet, dvs. ved at udføre et såkaldt real time panel-studie. Det er interesserant ikke bare i sig selv, men også fordi det er endnu et område, hvor man kan teste den sprogbaserede fonologis hypotese om, at reduktioner spredes sig fra højfrekvente til lavfrekvente ord. Hvis dette er en konsekvens af inkrementelle ændringer af ordenes repræsentation i det mentale leksikon, så bør forandringer, der fortsætter gennem livet, også følge ordfrekvens. Det vil sige, at samtidig med, at en lydforandring må forventes at tage til og måske blive obligatorisk i de højfrekvente ord, så bør man også forvente, at lydforandringen spredes sig til lavfrekvente ord, hvor den ikke tidligere forekom, i hvert fald for talere, der ændrer udtale livet igennem.

2 BORTFALD AF [w] I KØBENHAVSK

For at teste de hypoteser, der er fremsat i sociolinguistikken og den brugsbaserede fonologi, har jeg udført kvantitative studier af den fonetiske proces [w]-bortfald for syllabisk [ð], dvs. udældelse af det oprindelige [w] i ord som "blevet, skrevet, lavet", der i dag ofte udtaltes "[bleðː] 'sgæðː, læðː]". Denne reduktion af det oprindelige [w] har i tidligere undersøgelser vist sig at være en igangværende lydforandring, og det gør den derfor oplagt til at teste sammenhængen mellem ordfrekvens og spredning af en reduktion fra generation til generation og hos samme taler livet igennem. Undersøgelsen vil derfor forsøge at besvare følgende spørgsmål:

1. Er [w]-bortfald mere udbredt i højfrekvente ord end i lavfrekvente ord?
2. Er [w]-bortfald mere udbredt i relativt lavfrekvente ord hos yngre talere end hos ældre talere?
3. Kan sammenhængen mellem ordfrekvens og [w]-bortfald ses på individniveau?
4. Er der talere, der fortsætter med at deltage i processen, efterhånden som de bliver ældre, og følger de i så fald også ordfrekvensmåler?
Dette vil ikke kun give en opdatering af vores viden om udviklingen af [w]-bortfald, men vil også kunne besvare mere generelle spørgsmål om de processer, der er på spil i igangværende lydforandringer.

2.1 Bortfald af [w] før og nu


2.2 Lydforandring – hos hvem?


2.3 Fremgangsmåde


For at undersøge indflydelsen fra de faktorer, tidligere undersøgelser har fundet relevante, blev hver forekomst også kodet for trykgrad og foranstående vokals kvalitet. Tryk blev kodet som en binær egenskab – enten havde en forekomst hovedtryk, eller også havde den ikke – mens vokalkvaliteten blev kodet med reference til en af 10 mulige vokalkvaliteter. Desuden tilføjedes oplysninger om de enkelte ords leksikalske tryk efter den manuelle kodning var foretaget. Hvor det leksikalske tryk i principippet er fordelt på tre kategorier: hovedtryk, bitryk og nultryk, hvor bitryk defineres som reduceret hovedtryk på
andetled i sammensætninger (efter Grønnum 2005: 250), viste det sig
dog, at ingen af forekomsterne af [w] for [ø] optræder i stavelser, der
er specificeret for leksikalsk nultryk. Hver forekomst i datasettet er
således suppleret med en oplysning om, om det er leksikalsk specificeret
for hoved- eller bitryk, hvilket vil give mulighed for at undersøge,
om der er større eller mindre tendens til [w]-bortfald i andetleddet i
sammensætninger end i usammensatte ord.

For at kunne undersøge sammenhængen mellem tendens til [w]-
bortfald og hvor ofte et ord bruges, har jeg foretaget en frekvensunder-
søgelse af ord i løbende tale. Frekvensberegningerne er foretaget på
baggrund af de sociolingvistiske interview i DGCSS’ korpora. Materialet
indeholder 3,1 millioner løbende ord, og for hvert ord er den relative
hyppighed blevet beregnet ved at dividere antallet af forekomster med
det samlede antal løbende ord. Da den ortografiske udskrivning af
interviewene er blevet suppleret med en fonetisk annotation, der angiver
hvert ord's distinkte udtale, var det ikke nødvendigt at kontrollere for
homografi. Beretningerne afspejler således direkte de distinkt udtalte
ordformers relative hyppighed. Frekvenserne er efterfølgende blevet
konverteret til en logarithmisk skala, og hvor forekomst af (wø) i de
tildyttede interview er siden blevet suppleret med oplysninger om denne
logtransformerede ordfrekvens. I denne undersøgelse blev ordene delt
ind i fire frekvenskategorier, og her angiver jeg ordfrekvensen med
henvisning til, om ordet er højfrekvent, mellemfrekvent, lavfrekvent
eller meget lavfrekvent.

Oplysninger om informanternes kön og fødselsår er blevet trukket
ud af DGCSS’ database, sådan at sammenhængen mellem kön og [w]-
bortfald og ikke mindst alder og [w]-bortfald kunne analyseres.

Alle disse faktorer, både sproginterne, sprogbrugs- og de sociale
faktorer, her kön og alder, da alle informanter var classificeret som
middelklassekøbenhavnerne, kan tænkes at have indflydelse på [w]-
bortfald, og de kan naturligvis have det samtidig. For at undersøge
deres effekt har jeg benyttet en statistisk metode, der tillader, at man
undersøger effekten af hver enkelt faktor, samtidig med at man kon-
trollerer for alle de andre faktorer. Den metode beskrives kort nedenfor.
2.4 At sammenligne fordelinger af varianter


Multipel logistisk regression er fordelagtigt i undersøgelsen af et fænomen som lydforandringer, idet det er muligt at undersøge indflydelsen af flere faktorer samtidig. Når tidligere undersøgelser har vist, at foranstående vokals kvalitet og stavelsens tryk påvirker sandsynligheden for [w]-bortfald, ville man (hvis man ikke brugte multipel logistisk regression) være nødt til at undersøge forskellen i tendens hos de to generationer med separate chi-i-anden-prover for hver af de relevante kontekster⁴. Hvis man ser bort fra forskellen i kontekster i den statistiske analyse, vil man nemlig ikke kunne være sikker på, at en eventuel signifikant forskel mellem generationerne ikke skyldes en forskel i, hvor mange tryksvage forekomster, der var i hver af de to grupper. I multipel logistisk regression vurderes alle faktorer som nævnt samtidig, og derfor kontrolleres der for alle andre faktorer i evalueringen af hver enkelt faktors indflydelse på sandsynligheden for [w]-bortfald. Dvs. at når effekten af talerens alder på sandsynligheden for [w]-bortfald testes i regressionsmodellen, så tages der højde for påvirkningen fra de øvrige faktorer. Metoden sikrer således, at man ikke får udråbt en faktor til at være væsentlig, hvis den tilsyneladende effekt kan tilskrives andre faktorer. Resultatet kan ses med det samme for hver faktor, og virkningen af hver enkelt faktor kan også ses, når man konstruerer sin model for de data, man har. Metoden kan selvfølgelig kun vurdere de faktorer, man kan opstille på baggrund af de oplysninger, man har i sine data. I de interview, jeg har haft adgang til, har det for eksempel (endnu) ikke været muligt at beregne informanternes taletempo, så derfor kan jeg ikke undersøge, om tend-

Korpusundersøgelser af udtalevariation giver ikke kun problemer med at kontrollere for den sproglige kontext. Det er heller ikke muligt at kontrollere for, at de enkelte talere producerer lige mange forekomster af hvert ord, og det kan skabe heterogenitet i de grupper, der skal undersøges. Man kan for eksempel forestille sig, at kun enkelte individer i den yngre generation taler længe og dermed får produceret mange flere forekomster af variablen (wød). Disse individer vil så påvirke gruppens gennemsnit i meget høj grad og vil skabe et skævt indtryk af gruppens generelle adfærd. Det er naturligvis muligt at granske data og på den måde blive opmærksom på, om der er bestemte individer, der skiller sig ud og bør behandles separat. Men det er også muligt at kontrollere for den enkelte talers indflydelse på fordelingen af varianter inden for
gruppen ved at bruge regressionsmodellering med blandede effekter (såkaldte ”mixed models” eller ”mixed-effects models”). Ud over at inddrage faste faktorer som trykgrad, foranstående vocals kvalitet og aldersgruppe kan man inddrage tilfældige faktorer som den enkelte taler og det enkelte ord. Netop ved at inddrage den enkelte taler som en tilfældig faktor, kan man kontrollere for heterogenitet i gruppen af talere. Da effekten af talerkarakteristika som alder og køn påvirkes af, hvilke individer der indgår i gruppen af ældre og yngre talere, og hvilke individer der indgår i grupperne af mænd og kvinner, vil vi meget nodig risikere at stå i den situation, at den effekt af alder, vi kan vise, kun bør tilskrives nogle få medlemmer af den yngre generation. Vi vil m.a.o. gerne have større sikkerhed for, at effekten holder for alle medlemmer af sprogsamfundet, og det kan man få ved at bruge taleren som tilfældig faktor. På den måde tages der højde for hver enkelt talers indflydelse i den statistiske evaluering af effekten af de faste faktorer. Det betyder ikke, at der ikke kan være variation inden for grupperne, blot at den effekt, vi tilskriver gruppen, er relativt robust på trods af den interne variation.

Det at have talerne med som tilfældig effekt er relativt nyt i sociolingvistikken og er ikke helt ukontrroversielt. Sociolingvistikken har længe haft som mere eller mindre eksplicit mål at opstille sprogsystemer, herunder fonologier, for sprogsamfund og netop ikke for enkelte talere. Formålet er at fremstille, hvad der er fælles for de mange enkeltpersoner i et sprogsamfund. Ved at inddrage individuel variation i den statistiske analyse antager man samtidig, at der kan være så stor forskel på individers grammatik, at det ikke er muligt at opstille en grammatik for sprogsamfundet defineret ud fra en objektiv a priori parameter som køn, alder eller socialklasse. Ud over at det kan diskuteres, hvilken relevans sådanne objektive mål har for de mennesker, der kategoriseres ud fra dem, så mener jeg netop, at sprogsmålet om et sprogsamfunds grammatik er et empirisk spørgsmål, hvilket betyder, at man i stedet for at antage tilstrækkelig homogenitet inden for en prædefineret gruppe må undersøge variationen og være parat til at forkaste de prædefinerede gruppers rolle i den sproglige adfærd. Hvis en fast effekt som fx alder forbliver signifikant, selv når der tages højde for variationen inden for generationerne, så er resultatet kun styrket i forhold til en model, der ser
bort fra den mulige fejlkilde. Jeg ser således ikke noget problem i den implicitte konsekvens af at bruge taler som tilfældig effekt, nemlig at hver sprogbrugers mentale leksikon og grammatik må være specifik for den enkelte taler. Mere kontroversielt er det nok at bruge hvert enkelt ord som tilfældig effekt. Ikke desto mindre har jeg i alle mine analyser taget de enkelte ord med som tilfældig effekt, igen fordi jeg mener, det er bedre med den kontrol, det giver i den kvantitative analyse af data udtrukket fra et spontantalekorpus, hvor man som analytiker netop ikke har ret meget kontrol over data.

Valget af multipel logistisk regression til de statistiske analyser bunder således i et ønske om at forstå samspillet mellem alle de faktorer, der undersøges her, og hvordan de påvirker spredningen af [w]-bortfald. Brugen af taler og ord som tilfældige effekter foretages hovedsageligt for at kontrollere for en ønsket variabilitet. Resultaterne af analyserne vil således være mere konservative, men dermed også mere robuste, end hvis jeg havde brugt logistisk regression udelukkende med faste effekter, eller hvis jeg havde valgt at foretage separate chi-i-anden-prover på dele af data.

3 BORTFALD AF [W] I 1987 – APPARENT TIME-
UNDERSØGELSE
Her gennemgås resultaterne af den statistiske analyse af (wa?) i optagelserne fra 1987. Efter en gennemgang af de overordnede resultater går jeg videre til at se på, om processen stadig kan ses som en igangværende lydforandring, ved at sammenligne de to generationer, og dernæst indrager jeg ordfrekvenser for at kunne besvare spørgsmålene om, hvorvidt en igangværende lydforandring spredet sig gradvist men regelmæssigt gennem ordforrådet. Denne analyse foretages både på gruppeniveau og for de enkelte talere.
Helt generelt er der en kraftig tendens til [w]-bortfald allerede i optagelserne fra 1987: I 233 af 303 forekomster, eller 76 % af tilfældene, udelades [w]. Ifølge regressionsanalysen er fødselsårtil en signifikant faktor i modelleringen af [w]-bortfald (p = 0,025), og regressionskoefficienten for fødselsårtil er positiv (β = 0,0698), dvs. at sandsynligheden for [w]-bortfald stiger med fødselsårtil. Inddeler
man gruppen af informanter i to generationer (som beskrevet ovenfor), finder man også en statistisk signifikant forskel mellem de 12 ældre talere og de 10 yngre talere (p = 0,001): For gruppen af yngre sprogbrugere falder [w] bort i 88 % af alle tilfælde, mens det for de ældre sprogbrugere ”kun” sker i 73 % af tilfældene. Lydforandringen er som ventet fortsat i gang i 1987, og den ser endda ud til at være intensiveret. Men man kan også se, at [w]-bortfald endnu ikke er obligatorisk, heller ikke i den yngre generation.

Inden jeg går videre til analysen af sammenhængen med ordfrekvens, vil jeg for hver af de to generationer analysere effekten af den fonetiske kontekst på tendensen til [w]-bortfald. Husk, at faktorerne tryk og foranstående vokals kvalitet blev taget med i regressionsmodellen. For alle 22 informanter som helhed var begge disse faktorer signifikante, men for at undersøge, om de spiller forskellige roller i de to generationer, har jeg splittet datasettet op i to, nemlig forekomsterne produceret af ældre talere og forekomsterne produceret af yngre talere. Herved kan man undersøge, om tryk og foranstående vokals kvalitet stadig spiller den rolle, tidligere undersøgelser har vist, og om de to faktorer har samme betydning i de to generationer.

3.1 Tryk og [w]-bortfald

Først vil jeg se på tryk, både opfattet tryk og leksikalsk tryk. Opfattet tryk blev kodet som en binær egenskab ved stavelser i forbindelse med variantætningen (jf. afsnit 2.3 ovenfor), og forekomster kunne enten være trykstærke eller trykløse. Samtidig er hver enkelt ordform specifikkeret for leksikalsk tryk, altså det tryk, som ordet er ”fødtt med” i den leksikalske repræsentation, og som kan reduceres i sammenhængende tale (jf. fx Grønnum 2005 s. 250 ff.). Det relevante i denne sammenhæng er forskellen på hovedtryk og bitryk, da der i materialet ikke optræder forekomster af [w] i stavelser, der er specifikkeret for leksikalsk nultryk, og modelleringen for begge typer tryk er således som binære faktorer. Det leksikalske tryk er således en angivelse af ordets tryk løsrevet fra en kontekst og udtalt i isolation, mens det opfattede tryk viser, om den pågældende forekomst af et ord faktisk blev udtalt med tryk i den sammenhæng, det oprådde i. Der er naturligvis en vis korrelation mellem opfattet og leksikalsk tryk, idet 95
% af de stavelser i materialet, der har bitryk, realiseres som tryklose, ganske som man ville vente. Men der er også mange af de stavelser, som leksikalk er specificeret for hovedtryk, der realiseres som tryklose. Størstedelen af disse kan tilskrives forekomster af ordet ’blevet’, der oftest realiseres trykloset i sammenhængende tale. Spørgsmålet er så, om det er det tryk, ordet udtales med, altså det opfattede tryk, eller det tryk, ordet er specificeret for leksikalk, der hænger bedst sammen med tendensen til [w]-bortfald. For at undersøge dette blev begge typer tryk repræsenteret som faktorer i regressionsmodellerne. Og for begge generationer viser der sig en signifikant effekt af leksikalk tryk, men opfattet tryk spiller kun en rolle hos den ældre generation. Hos denne gruppe af talere er der større sandsynlighed for [w]-bortfald i tryklose stavelser end i trykstærke (β = 1,29 og p = 0,005). I de tryklose stavelser er dog sådan, at sandsynligheden for [w]-bortfald falder i stavelser med leksikalk bitryk (β = -2,22 og p < 0,0001). Hos den yngre generation er det kun leksikalk tryk, der bidrager signifikant til modellen (β = -2,17 og p = 0,04), og igen er der tale om, at sandsynligheden for [w]-tab falder i stavelser specificeret for bitryk sammenlignet med stavelser specificeret for hovedtryk.

3.2 [w]-bortfald og foranstående vokals kvalitet

Selvom der er stor tendens til [w]-bortfald, så er det stadig sådan for begge generationer at [w] ikke altid falder bort, heller ikke i stavelser uden tryk hos de yngre. Som nævnt viser tidligere undersøgelser, at [w] oftere falder bort efter fortungevokaler end efter bagtungevokaler, medmindre vokalerne er høje, for her var tendensen i hvert fald tidligere modsat. Det var altså mere almindeligt, at [w] faldt bort efter [e, ø, ø] og [æ], dvs. i ord som 'levet, løbet, slæbet, lavet' end det gjorde efter [o, a, ə] og [ɑ] altså i ord som 'toget, tovet, storvægot'. Den forskel går igen hos begge generationer i optagelserne fra 1987: Hvor 83 % af de [w]’er der forekommer efter en fortungevokal, bliver sletthans, så sker det kun for 14 % af dem, der forekommer efter en bagtungevokal. Forskellen er signifikant ifølge regressionsmodellerne, idet faktoren er signifikant både for de ældre talere (p < 0,0001 med β = 3,35) og for de yngre talere (p = 0,007, β = 5,79).

Der er altså en tendens til, at de fonetiske faktorer, der kan blokere [w]-bortfald, stadig er virksomme, omend trykket ser ud til at have mistet denne funktion i de yngres udtale.

3.3 Ordfrekvens

Men hvad nu med ordfrekvens? Det var jo den primære faktor, der skulle testes i undersøgelsen. Regressionsanalysen viste, at ordfrekvens var en signifikant faktor (p = 0,005) og går i den ventede retning, idet β = 3,537 og altså positiv, dvs. sandsynligheden for [w]-bortfald stiger med ordfrekvensen. Den sorte del af søjlerne i figur 1 nedenfor viser andelen af ordforekomster med [w]-bortfald for hver af de repræsenterede frekvenskategorier i materialet.
Det ses tydeligt, at der er en sammenhæng mellem ordfrekvens og tendens til [w]-bortfald, og den går i den forventede retning: [w]-bortfald er hyppigst i højfrekvente ord som 'laver', knap så hyppigt i mellemfrekvente ord, som 'skrevet', og meget sjældent i sjældne ord som 'rever'. Desuden kan man ved at sammenligne fordelingerne for de to generationer se, at de yngre talere har en relativt højere tendens til at udelade [w] i høj- og mellemfrekvente ord end de ældre talere. Det er desværre ikke muligt at se, om denne tendens også er højere i de sjældneste ord med (wød), da ingen af de yngre talere producerer nogle af disse ord i interviewene, hvilket er grunden til, at der kun er tre søjler i diagrammet for denne generation. Husk på, at regressionsanalysen kontrollerer for de andre signifikante faktorer, samtidig med at den modellerer effekten af ordfrekvens. Dvs. at den effekt, vi kan observere her hos de to generationer, ikke kan tilskrives tryk eller vokalkvalitet. Ganske vist er det sådan, at de højfrekvente ord tenderer mod at blive realiseret som tryksvage, men selv når de bliver realiseret med fuldtryk, har de større tendens til at blive ramt af [w]-bortfald end sammenlignelige, men mindre hyppige ord. Vi kan altså godt støde på forekomster af 'blever' udtalt ['bleøðø], dvs. med tryk og [w]-bortfald, mens det er mindre sandsynligt at høre 'levet' udtalt ['leøðø].
På gruppeniveau er der således støtte for den sprogbrugsbaserede fonologisk hypotese om, at reduktioner spreder sig gradvist gennem ordforrådet, fra højfrekvente til lavfrekvente ord, og at transmissionen fra generation til generation også er påvirket af, hvor hyppigt ordene bruges. For at undersøge, om resultaterne også kan støtte den foreslåede model for mental repræsentation af [w]-bortfald, er det imidlertid nødvendigt at undersøge, om frekvenseffekten også kan findes hos de enkelte sprogbrugere.

3.4 Individ og ordfrekvens
Det er ikke muligt at lave en statistisk signifikantestning af ordfrekvensens rolle for den enkelte taler – der produceres ganske enkelt for få forekomster i hvert interview til, at det giver mening. Men det er muligt at kvantificere andelen af forekomster med [w]-bortfald for hver frekvenskategori hos hver taler, og på den måde give en kvalitativ analyse af sammenhængen mellem ordfrekvens og tendens til reduktion hos individer. Nogle talere producerer dog så få forekomster i løbet af interviewet, at en sådan opgørelse igen ikke er meningsfuld, og i det efterfølgende begrænses analysen til at dreje sig om de talere, der har producet mere end 10 forekomster af variablen (wø).

Gruppen af ældre talere kanindeles i tre undergrupper mht. gruppens gennemsnitlige adfærd: de konservative, de typiske, og de atypiske. Alle talernes tendenser er opgjort i tabellen nedenfor.

TABEL 1. [w]-BORTFALED SOM EN FUNKCIION AF ORDFREKVENS HOS DE ENKELTE TALER I DENÆLDERE GENERATION 1987

<table>
<thead>
<tr>
<th>Informant</th>
<th>Antal ord</th>
<th>% bortfald</th>
<th>% bortfald</th>
<th>% bortfald</th>
<th>% bortfald</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Meget lav-frekvente ord</td>
<td>Lavfrekvente ord</td>
<td>Mellemfrekvente ord</td>
<td>Højfrekvente ord</td>
</tr>
<tr>
<td>LAL</td>
<td>11</td>
<td>-</td>
<td>0</td>
<td>0</td>
<td>75</td>
</tr>
<tr>
<td>TNI</td>
<td>33</td>
<td>-</td>
<td>33</td>
<td>50</td>
<td>87</td>
</tr>
<tr>
<td>FTK</td>
<td>38</td>
<td>0</td>
<td>0</td>
<td>37</td>
<td>94</td>
</tr>
<tr>
<td>CNI</td>
<td>20</td>
<td>-</td>
<td>-</td>
<td>20</td>
<td>100</td>
</tr>
<tr>
<td>ASA</td>
<td>11</td>
<td>0</td>
<td>-</td>
<td>40</td>
<td>100</td>
</tr>
<tr>
<td>EAF</td>
<td>11</td>
<td>-</td>
<td>50</td>
<td>-</td>
<td>100</td>
</tr>
<tr>
<td>HTH</td>
<td>20</td>
<td>-</td>
<td>0</td>
<td>50</td>
<td>100</td>
</tr>
<tr>
<td>MJF</td>
<td>15</td>
<td>-</td>
<td>0</td>
<td>50</td>
<td>100</td>
</tr>
<tr>
<td>CEL</td>
<td>17</td>
<td>-</td>
<td>100</td>
<td>100</td>
<td>87</td>
</tr>
<tr>
<td>MFL</td>
<td>22</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>90</td>
</tr>
<tr>
<td>LGR</td>
<td>6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PFR</td>
<td>9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

For de yngre talere er tendensen mere klar – hos alle talere ses der en sammenhæng mellem ordfrekvens og [w]-bortfald efter nøjagtigt det samme mønster, som blev fundet på grupperniveau. Dog kan dette ikke siges med sikkerhed for to af talerne i den yngre generation, men det skyldes simpelthen, at [w]-bortfald er obligatorisk hos dem, de sletter m.a.o. altid [w] før syllabisk [ŋ]. Ganske som hos den ældre gruppe kan den yngre gruppe af talere inddeles i konservative og typiske talere. Tabellen nedenfor angiver proportionerne for hver enkelt.

**TABEL 2. [w]-BORTFALD SOM EN FUNKTION AF ORDFREKVENS HOS DE ENKELTE TALERE I DEN YNGRE GENERATION 1987**

<table>
<thead>
<tr>
<th>Informant</th>
<th>Antal ord</th>
<th>% bortfald</th>
<th>% bortfald højfrekvente ord</th>
<th>% bortfald mellemfrekvente ord</th>
<th>% bortfald lavfrekvente ord</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIP</td>
<td>18</td>
<td>-</td>
<td>50</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>JJE</td>
<td>12</td>
<td>-</td>
<td>67</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>PKJ</td>
<td>20</td>
<td>0</td>
<td>75</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>DBE</td>
<td>17</td>
<td>0</td>
<td>80</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>JOR</td>
<td>14</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>MPT</td>
<td>16</td>
<td>-</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>ASH</td>
<td>9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>KJK</td>
<td>4</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>NOR</td>
<td>6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>SKO</td>
<td>7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

156

NYS 39
Tabellen viser, at alle talere i den yngre generation har obligatorisk [w]-bortfald i højfrekvente ord. De udtales således altid 'blevet' som ['bleːdə] og aldrig som ['blewəd] eller ['blewəd]. Dog er der også i denne gruppe forskel på, hvor fremskreden processen er hos hver enkelt taler. Hos de første fire er [w]-bortfald endnu ikke obligatorisk, mens den hos JOR og MPT ser ud til at være helt gennemført. Desværre er der hele fire af de 10 talere i gruppen, som producerer så få forekomster, at det ikke giver meget mening at se på fordelingen af dem i forhold til ordenes frekvenskategori. Det er dog værd at bemærke, at det er disse fire talere, der står bag alle forekomster af lavfrekvente ord med [w]-bortfald (jf. figur 1 ovenfor), men med så få observationer kan de ikke grupperes sammen med CEL og MFL fra den ældre generation, som var de to talere der udviste en tendens, der gik imod frekvenseffekten.

Ud over at disse fordelinger for individer viser, at frekvenseffekten gælder for langt størstedelen af informanterne, så viser de også, at grupperne er forskelligartede. Hvor de første analyser viste tendenser til [w]-bortfald på gruppeniveau, kan vi her se, at det ikke bør tolkes som udtryk for, at alle talere fra en yngre generation altid vil have større tendens til at slette [w] for syllabisk [ø] end en hvilken som helst taler fra en ældre generation. Faktisk minder mønsteret hos MIP og JJE fra den yngre generation mere om det hos CNI og PTK fra den ældre generation end om mønsteret hos deres jævnaldrende – og CNI og PTK er relativt konservative i deres udtale, i hvert fald i den forstand, at de ikke er blandt de talere, der har den største tendens til at udelade [w] for [ø]. Det er dog kun i den yngre generation, at der overhovedet er talere, som har obligatorisk [w]-bortfald i alle relevante ord, så selvom der er lighedstræk mellem individer fra to forskellige generationer, så er der også typer af talere, som kun forekommer i den ene af de to grupper. Det understreger, at effekten af alder betragtet på gruppeniveau kun bor tolkes sådan, at sandsynligheden for at finde en taler med en meget høj tendens til [w]-bortfald er størst, når man leder i gruppen af yngre talere, end når man leder i gruppen af ældre talere.

Tilsammen viser analyserne af data fra 1987, at [w]-bortfald stadig var en igangværende lydforandring, og at den, som reductionsprocesser i andre sprog, følger ordfreksven både på gruppeniveau og hos den en-
kelte taler, også når man i analysen kontrollerer for sproginterne faktorer. Analysen støttet således hypoteserne fra den sprogbrugsbaserede fonologi, men hvad med antagelsen bag apparent time-hypotesen, altå den, at sprogbrugere forbliver stabile i deres udtale op gennem voksenlivet? For at kunne teste den er det nødvendigt at undersøge variablen (wø) i real time.

4 BORTFALD AF [W] I 2006 HOS DE SAMME TALERE – REAL TIME-UNDERSØGELSE


Forskellen mellem de to optagelser er signifikant for begge generationer (p = 0,02 for den ældre generation, p = 0,005 for den yngre generation), og ændringen er omtrent lige stor i begge generationer: Hvor [w] faldt bort i 73 % af tilfældene hos den ældre generation i 1987, falder det nu bort i 80 % af tilfældene. Hos den yngre generation er der nu [w]-bortfald i 94 % af tilfældene i modsætning til de 88 % af tilfældene i optagelserne fra 1987.

Den øgede tendens til [w]-bortfald i den ældre generation har dog ikke ført til en ændring i betydningen af de fonetiske faktorer. [w]-bortfald er stadig mere sandsynligt i tryklose stavelser end i trykstærke (p = 0,0006), medmindre den tryklose stavelse har bitryk, idet processen her hæmmes (p = 0,006). Og [w]-bortfald er stadig mere almindeligt efter fortungevokaler end efter bagtungevokaler (p < 0,0001) for de ældre talere. Derimod har trykket helt mistet sin betydning hos den yngre generation, idet hverken opfattet tryk eller leksikalsk tryk længere er signifikante faktorer (hvd. p = 0,3 og p = 0,2). Det, der i 1987 lignede et begyndende tab af "reglen" om, at tryk beskytter mod [w]-bortfald, ser således ud til at være blevet gennemført i den mellemliggende
periode. Det kunne være interessant at undersøge, om det samme
gor sig gældende for den nye generation af unge københavnerne
fra middelklassen, og ikke mindst om processen er endnu mere
fremskreden efter bagtungevokaler, som for den yngre generation i
 dette datasets er den eneste kontekst, der hæmmer processen. Det er
desværre ikke muligt at undersøge med de data, der er tilgængelige på
nuværende tidspunkt.

Det er dog muligt at se, om ordfrekvens stadig spiller en rolle, og
altså spørge, om [w]-bortfald blot er intensiveret i de høj- og mellem-
frekvente ord, eller om processen har bredt sig til sjældnere ord.
Faktoren ordfrekvens er signifikant for begge generationer (p < 0,0001
i begge tilfælde), og koefficienten er positiv for begge (β = 1,15 for de
ældre og β = 1,56 for de yngre informanter), så det er stadig sådan, at
[w]-bortfald er hyppigere i højfrekvente end i lavfrekvente ord. For
den ældre generation viser det sig, at processen har bredt sig til de
meget lavfrekvente ord, dvs. ord som 'pivet' og 'kravet', hvor der ikke
tidligere er konstateret [w]-bortfald hos denne generation (jf. figur 1
ovenfor). Hos den yngre generation er det i de mellem- og lavfrekvente
ord, at processen er taget til, men der forekommer nu også ord af
den meget lavfrekvente slags i interviewe med talere fra den yngre
generation. Det er dog ikke muligt at sige, om processen har bredt sig
til denne frekvenskategori, da ingen af de yngre talere producerede
ord af denne type i de første interview. Men det kan konstateres, at en
lydforandring som [w]-bortfald over tid inden for den samme gruppe
af mennesker også følger ordfrekvens.

4.1 Lydforandring gennem livet hos individer
Til sidst vil jeg igen se på, om en egenskab ved grupper, nemlig den
ovenfor beskrevne stigning i [w]-bortfald hos både den ældre og yngre
generation, også kan spores hos enkeltpersoner. Det vil igen give et mere
nuanceret indblik i, hvor stabil individets fonologi kan siges at være, og
i hvor høj grad der er sammenhæng mellem gruppetilhørshorsforhold, her
generations medlemmsskab, og adfærd i løbet af livet.

Kravet om statistisk signifikans kan ikke fraviges, når det skal undersøges om enkeltpersoner forbliver stabile eller følger sprogsmundets
tendens og intensiverer [w]-bortfald i deres sprogbrug. Derfor blev
tendensen til [w]-bortfald signifikantestet for hver taler ved hjælp af chi-i-anden-prøven (eller Fisher's exact-test, som bruges, når de forventede frekvenser er for små til at tillade brug af chi-i-anden-prøven). Kun for 3 af talerne er andelen af forekomster med [w]-bortfald signifikant forskellige i de to optagelser. Data fremgår af tabellen nedenfor.

TABEL 3 – FORDELINGEN AF VARIANTER AF (wø) FOR DE TALERE DER ENDER SIG OVER TID

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DBE</td>
<td>6</td>
<td>12</td>
<td>70 %</td>
<td>3</td>
<td>38</td>
<td>92 %</td>
<td>0,04</td>
<td>Ungre</td>
<td>Kvind</td>
</tr>
<tr>
<td>LAL</td>
<td>8</td>
<td>3</td>
<td>27 %</td>
<td>0</td>
<td>10</td>
<td>100 %</td>
<td>0,001</td>
<td>Aldre</td>
<td>Kvind</td>
</tr>
<tr>
<td>PTK</td>
<td>16</td>
<td>22</td>
<td>58 %</td>
<td>4</td>
<td>29</td>
<td>88 %</td>
<td>0,007</td>
<td>Aldre</td>
<td>Mand</td>
</tr>
</tbody>
</table>

For en af disse talere, LAL, er det umuligt at sige, om forandringen har fulgt ordfrekvensens vej – hun går fra at have relativt få [w]-bortfald (faktisk er hun den mest konservative af alle taler i hele stikprøven fra 1987, jf. tabel 1 ovenfor) til at have obligatorisk [w]-bortfald. Men for de to andre giver det mening at sammenligne fordelingerne i 1987 med fordelingen i 2006. Husk fra tabel 2 ovenfor, at DBE aldrig havde [w]-bortfald i lavfrekvente ord i optagelsen fra 1987. I optagelsen fra 2006 har hun [w]-bortfald i 80 % af de lavfrekvente ord, hun producerer. Dvs. at hvor hun før udtalte ‘rever’ som ['wa[wø]'], udtaler hun det nu ['wa[wø]]. DBE er den tydeligste eksempel på en effekt af ordfrekvens hos en taler, der forandrer udtalevaner i løbet af sit liv ved fortsat at følge med i den igangværende lydforandring. For PTK er det desværre ikke muligt at afgøre, om den øgede tendens til [w]-bortfald, der ses i sammenligning af den anden optagelse med den første, hænger sammen med ordfrekvens. For selvom PTK faktisk siger flere ord med (wø) i det andet interview, så er de allesammen enten høj- eller mellemfrekvente. De tilhører altså de samme frekvenskategorier som de ord, hvor PTK havde [w]-bortfald allerede i optagelsen fra 1987, blot er andelen steget i de mellemfrekvente.
Evidensen fra de talere, der kan vises at have ændret udtalevaner mht. [w]-bortfald i løbet af deres voksenliv, er ganske sparsom, og det er nødvendigt at undersøge en mere hyppigt forekommende variabel end (wød) for at få indblik i sammenhængen mellem ordfrekvens og lydforandringer gennem livet. Men resultaterne går i det mindste ikke imod hypotesen om ordbrug som en drivkraft i spredningen af en lydforandring gennem individets mentale lektion.

5 SAMMENFATNING
De analyser, der er blevet præsenteret her, viser tendenser hos grupper – grupper af mennesker og grupper af ord. Der er ikke tale om obligatoriske regler, men ser man samlet på fordelingen af ord udtalt med og uden det oprindelige [w], er der nogle mønstre, der træder frem:

[w]-bortfald er en meget udbredt proces i løbende, daglig tale, men den er mere udbredt i højfrekvente ord end i lavfrekvente ord.

Yngre talere har større tendens til at udelade [w], end ældre talere har, og de gør det oftere i lavfrekvente ord, hvor ældre talere har tendens til at bevare [w].

Analyserne viser også tendenser hos enkeltpersoner, og det blev blandt andet vist, at effekten af ordfrekvens kan spores hos de fleste talere. Det er altså ikke bare en gennemsnitlig tendens hos grupper af talere, at [w]-bortfald er hyppigere i højfrekvente end i lavfrekvente ord – det er et forhold, der gælder for den enkelte taler. Men selvom den generelle tendens til oftest at udelade [w] i højfrekvente ord går igen hos (næsten) alle informanter, så er der også en variation inden for de to generationer, informanterne var blevet inddelt i. Det er altså ikke sådan, at tendensen er ens hos alle de ældre og alle de yngre talere – tværtimod kan tendensen hos nogle af de ældre talere mindre om tendensen hos de yngre og omvendt. Der er m.a.o. konservative talere blandt de yngre og progressive talere blandt de ældre, og denne tendens kan man kun få øje på ved at se nærmere på de enkeltpersoner, der indgår i de grupper, som traditionelt anvendes i kvantitativ variationssociolinguistik. At sammenhængen mellem [w]-bortfald og ordfrekvens kan ses på individniveau, støtter den anskuelse, at sprogbrugerens implicitte viden
om deres sprogs lydsystem, deres mentale repræsentation af, hvordan fonemer manifesteres, er tæt knyttet til brugen af de ord, som er bygget op af disse fonemer. Det betyder, at modeller for det mentale leksikon, ordenes repræsentation i hukommelsen, må tage højde for sprogbrugernes daglige brug af ordene samt give mulighed for individuelle forskelle i indholdet af det mentale leksikon. Den type modeller, der er bedst egnet til at inkorporere sprogbrugsaspektet, er de såkaldte eksemplar-modeller. Men der må flere undersøgelser til, for at vi bedre kan forstå begrænsningen af frekvenseffekten og dens indflydelse på det talte ord.

Nicolai Pharao
Danmarks Grundforskningsfonds Center
for Sociolingvistiske Sprogsforandringsstudier
Københavns Universitet
nicolaip@hum.ku.dk
NOTER

1 I denne artikel lydskrives med modificeret IPA i henhold til den konvention, der anvendes i Grønnun 2005), jf. oversigten ”Nogle til ydskrift” i dette nummer af NyS.

2 Hvor fonetikeren beskriver en lydforandring ved at sammenligne, hvilken sproglyd der tidligere blev brugt i en given lydlig kontekst, med den, der bruges på beskrivelsestidspunktet, og således må holde sig til en rent fonetisk notation angivet i [ ]-klammer, beskriver fonologen samme proces som en ændring i manifestationen af et bestemt fonem eller som et skift mellem to fonemer, der også indgik i sprogets foneminventar tidligere. Fonemer er abstrakte kategorier for de enkeltyde, der bruges i tale, og de skrives i ///-klammer.

3 Jeg følger her den sociolingvistiske konvention og angiver variablen i runde paranteser ( ) i modsætning til variablens varianter, som angives i lydskrift, og her kan være [wØ] eller [Ø].

4 Chi-i-anden-proven er, forenklet sagt, en statistisk test, der gør det muligt at undersøge, i hvor høj grad forskelle i fordelinger af udfald inden for et begrænset sæt kategorier kan tilskrives tilfældigheden, eller om det er sandsynligt, at udfaldet skyldes den faktor, fordelingerne adskiller sig på (Baayen:2008). For eksempel kan man undersøge, om forskellen i fordelingen af bevarede og bortfaldne [w]er i de to generationer, der er blevet undersøgt her, lige så vel kunne være opstået ved en tilfældighed, eller om forskellen ser ud til ikke at være tilfældig, og på det grundlag kan tilskrives at være en effekt af netop faktoren alder.
LITTERATUR


Gregersen, Frans (2009) ”The data and design of the LANCHART study”. Acta Linguistica Hafniensia 41. 4-37.


Gregersen, Frans, Marie Maegaard & Nicolai Pharao (2009) ”The long and short of (æ)-variation in Danish – a panel study of short (æ)-variants in Danish in real time”. Acta Linguistica Hafniensia 41. 64-82.


Kristensen, Kjeld (1977) ”Variationen i vestjysk stationsbymål”. Dialektstudier 4, 1. halvbinding.


