



Københavns Universitet



## Ordbrug og udtaleforandringer

Pharao, Nicolai

*Published in:*  
NyS

*Publication date:*  
2010

*Document Version*  
Peer-review version

*Citation for published version (APA):*  
Pharao, N. (2010). Ordbrug og udtaleforandringer. NyS, (39), 137-165.

# Ordbrug og udtaleforandringer

NICOLAI PHARAO

Sprog er i konstant forandring, ord kommer til og glider ud eller får nye betydninger, og sprogets enkeltlyde ændres. Traditionelt beskrives en lydforandring med udgangspunkt i det segment, den enkeltlyd, der undergår en forandring. Det er for eksempel almindeligt at sige, at oprindeligt [æ]<sup>1</sup> er blevet til [ɑ] efter oprindeligt [ɛ] i rigsmålet (Brink & Lund 1975: 128-132), eller med fonologens ord, at systematikken i enkeltlydsinventaret er ændret, sådan at fonemet /ɛ/ er blevet skiftet ud med fonemet /ɑ/ efter /r/ i moderne københavnsk rigsmål<sup>2</sup> (Grønnum 2005: 332). Det er naturligvis implicit i disse beskrivelser af enkeltlydes forandringer, at det er sprogbrugerne i et afgrænset sprogsamfund, der har ændret udtaleadfærd, sådan at begge ovennævnte udsagn svarer til at sige, at ord, der på et tidligere tidspunkt blev udtalt med [ɛæ], som [ɛæð] for 'ret' og [fɛæmʔ] for 'frem', nu udtales med sekvensen [ɑ], dvs. hhv. [ɑð] og [fɛamʔ]. Men en beskrivelse med udgangspunkt i segmenterne slører imidlertid det forhold, at udbredelsen af den nye variant i det sprogsamfund, hvor processen siges at forekomme, ikke nødvendigvis er total. Lydforandringen er m.a.o. ikke nødvendigvis helt gennemført for alle talere, og det er heller ikke sikkert, at lydforandringen rammer alle de ord, der indeholder det segment eller den sekvens af segmenter, der siges at være blevet ændret. Begge disse aspekter er imidlertid vigtige for en forståelse af, hvordan lydforandringsprocesser foregår, og hvordan de påvirker den mentale repræsentation af ordenes form i den enkelte talers hukommelse.

Udbredelsen af lydforandringer blandt sprogbrugere og i deres ordforråd står centralt i analyser af lydforandringsprocesser i både den sprogbaserede lingvistik og sociolingvistikken, og i begge tilgange er der fremsat hypoteser om faktorer i udbredelsesprocessen og disses implikationer for ordenes repræsentation i det mentale leksikon. Disse hypoteser vil jeg teste her ved at undersøge udviklingen i bortfald af [w] før syllabisk [ð] i københavnsk talesprog.

## 1 LYDFORANDRING OG ORD I HUKOMMELSEN

### 1.1 Sprogbrugsbaserede studier af lydforandring

I den sprogbrugsbaserede lingvistik generelt og dermed også i den sprogbrugsbaserede fonologi som anlagt i Bybees *Phonology and Language Use* (2002) betragtes de lydsystemer, der kan opstilles på strukturelt grundlag, ikke som autonome eller på forhånd givne størrelser, men som regelmæssigheder, der opstår i konkret brug af talesprog. Som følge heraf søges ændringer i systemet forklaret ud fra ændringer i sprogbrugen. Når det drejer sig om lydforandringer, fokuseres der på det fænomen, at lydforandringer foregår gradvist gennem ordforrådet. Dette står i kontrast til en strukturelt baseret analyse, hvor lydforandringer beskrives som kategoriske, dvs. de er processer, som appliceres undtagelsesløst på alle ord, der indeholder den lydsekvens, der rammes af den proces, man ønsker at redegøre for (fx Zwicky 1972).

Den sprogbrugsbaserede tilgang henter især støtte i studier af reduktioner, der begynder som forenklende processer i spontan, løbende tale og senere breder sig gennem ordforrådet. Et klassisk eksempel i litteraturen er assimilation af schwa før /r/ i amerikansk engelsk som i ordene 'every, memory, mammary' (da.: 'hver, minde, bryst(kirtel)'), der i distinkt udtale alle ville ende på sekvensen [əɹi]. En analyse af ordenes udtale i spontan tale viser imidlertid (Bybee 2002: 40-42), at det højfrekvente ord 'every' aldrig indeholder et schwa, men faktisk reduceres til et tostavellesord, altså ['evɹi]. Det middelfrekvente ord 'memory' indeholder også kun sjældent schwa, men her bliver [ɹi] dog stavelsesbærende, dvs. den typiske udtale er [mɛməɹi]. Begge disse tendenser modsvares af det meget lavfrekvente ord 'mammary', hvor schwa'et aldrig reduceres, dvs. at det altid udtales [mæməɹi]. Netop det fænomen, at bortfaldet af schwa er obligatorisk i højfrekvente ord, variabelt og sjældent komplet i mellemfrekvente ord og meget sjældent forekommende i lavfrekvente ord, fungerer som modbevis for påstanden (jf. Zwicky 1972) om den kategoriske applikation af fonetiske reduktionsprocesser. Det ses tydeligt, at ikke alle forbindelser af [əɹi] reduceres, men variationen i, hvornår reduktion af schwa'et finder sted, og hvornår processen hæmmes, er ikke helt tilfældig: Den følger med ordfrekvens, sådan at jo højere et ords frekvens, desto højere

er sandsynligheden for, at en reduktionsproces vil blive appliceret på ordet i løbende tale. I sprogbrugsbaseret fonologi forklares denne sammenhæng med, at reduktionsprocesser, som i eksemplet reduktion af schwa, er rutinerede artikulatoriske vaner: Gentagen udtale af sekvensen fører til overlap af de bevægelser af taleorganerne, der skal til for at producere hhv. schwa og [ɹ], og denne øgede tendens til overlap forstærkes, hver gang et ord bruges. Brugen af ord fører således gradvist til en rutinisering af udtalen, og reduktionen er mere udbredt i de højfrekvente ord, *fordi* de bruges oftere. Ligeledes beskyttes sjældne ord mod forandringen, *fordi* de er sjældne. På denne måde former brugen af ord det lydsystem, der kan opstilles af analytikeren. Det indikerer, at sprogbrugernes ”viden om” eller mentale repræsentation af udtalen ikke er uafhængig af de ord, enkeltlydene indgår i. Det har ført til et forslag om en radikalt anderledes mental repræsentation af ords udtale end den, der traditionelt har dannet grundlaget for teorier om leksikalsk repræsentation.

### *1.2 Ordbrug og mental repræsentation af udtale*

Både traditionel, strukturel fonologi og den sprogbrugsbaserede fonologi antager, at ordformer eksisterer som separate enheder i hukommelsen (i det omfang at beskrivelsen overhovedet forpligter sig på psykologisk realitet). Men hvor den traditionelle tilgang (fx Levelt et al. 1999) modellerer repræsentationen af de enkelte ord som bestående af én invariant form, hvis egentlige udtale så dannes gennem applikationen af en eller flere regler, så bevares al variationen direkte i det mentale leksikon i den sprogbrugsbaserede model af repræsentationen af ords udtale (Bybee 2002, Pierrehumbert 2001). Det betyder ikke, at der ikke også samtidig kan finde abstraktion sted i den sprogbrugsbaserede model, og at fx fonemer på den måde kan få psykologisk realitet. Men fokus er på sprogbrugernes viden om den variabilitet, der findes i udtalen af deres sprog, både hos dem selv og hos andre. Dette fanges gennem en såkaldt eksemplar-baseret tilgang til repræsentationen af ord i hukommelsen (jf. Johnson (2007) for en kort indføring i, hvordan eksemplar-teorien for perception af sanseindtryk opstår i psykologien og senere tages i brug i psykolingvistikken). Hvor de traditionelle modeller fjerner al forudsigelig variation og således kun

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. Repræsentationen af en ordform konceptualiseres således som en sky af alle de huskede forekomster, de såkaldte eksemplarer, og denne sky vokser, hver gang ordformen høres eller siges. Repræsentationerne kan således opdateres løbende i hver enkelt sprogbrugers hukommelse, og efterhånden som udtalen af en lydsekvens rutiniseres i højfrekvente ord som følge af generelle principper for automatisering af ofte gentagede bevægelser (jf. ovenfor), vil repræsentationen af højfrekvente 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 eksemplarer af formen ['evɹi] end af formen ['evəɹi], og derfor bliver formen uden schwa mere almindelig. Igen er det, fordi ordene bruges ofte, at deres repræsentation ændres hurtigere.

For at denne forbindelse mellem sprogbrug og repræsentation af udtale i det mentale leksikon skal kunne underbygges, er det imidlertid nødvendigt, at effekten af ordfrekvens, som blev vist på gruppeniveau 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 gruppeniveau. 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 afspejler 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 lydforandring

Som nævnt indledningsvist er der ikke kun forskel på, hvor ofte ord, der indeholder lyde som er under forandring, faktisk påvirkes, dvs. realiseres med den nye variant af en variabel udtale. Der er også forskel på, hvor udbredt den er i forskellige grupper af det sprogsamfund, hvor den er undervejs. Det har en lang række af sociolingvistiske undersøgelser vist (jf. Labov 1994 for en oversigt over mange undersøgelser af især amerikansk engelsk), også sociolingvistiske undersøgelser af dansk (se fx Kristensen 1977, Kristensen & Jørgensen 1994, Gregersen & Pedersen 1991). I størstedelen af disse undersøgelser af igangværende lydforandringer hviler analyserne af forandringerne på den antagelse, der ligger bag den såkaldte *apparent time*-hypotese (Labov 1966(2006): 200), som siger, at en forskel i udtalen mellem to generationer afspejler en igangværende udvikling. Hvis vi fx tager ændringen fra oprindeligt [kæ] til [kɑ] igen, så kan vi sige, at denne proces er en igangværende lydforandring ved fx at sammenligne en gruppe talere på 40 år med en gruppe talere på 20 år. Hvis sandsynligheden for at bruge [kɑ], hvor der tidligere for det meste hørtes [kæ], er højere hos de 20-årige end hos de 40-årige, så kan man tolke mønsteret sådan, at [kæ] er blevet til [kɑ] i løbet af de 20 år, der er gået siden gruppen af 40-årige var 20 år gamle. Men denne tolkning af mønsteret kan *kun* laves, hvis udtalen hos gruppen af 40-årige er forblevet stabil i løbet af perioden. Med denne *apparent time*-tilgang antager man altså, at i hvert fald udtalen stabiliseres på et bestemt tidspunkt i et menneskes liv, i eksemplet at den udtale, de 40-årige har i dag, er den samme som den, de havde for 20 år siden. På den måde bliver sammenligningen af de to generationer til et billede af en forandring og ikke bare til et synkront øjebliksbillede af en udtaleforskel mellem unge og midaldrende talere.

Men det er ikke sikkert, at sprogbrugerens 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 interessant ikke bare i sig selv, men også fordi det er endnu et område, hvor man kan teste den sprogbrugsbaserede fonologis hypotese om, at reduktioner spredt 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 spredt 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 sociolingvistikken og den brugsbaserede fonologi, har jeg udført kvantitative studier af den fonetiske proces [w]-bortfald før syllabisk [ð], dvs. udeladelse af det oprindelige [w] i ord som 'blevet, skrevet, lavet', der i dag ofte udtales ['ble:ðð 'sg̃ræ:ðð 'læ:ðð]. 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 ordfrekvensmønstret?

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

Som nævnt er dette ikke den første undersøgelse af [w]-bortfald i københavnsk talesprog, og der er en række faktorer, som bør tages i betragtning i undersøgelsen af sammenhængen med ordfrekvens. Tidligere undersøgelser har nemlig vist en række sproginterne forhold, der påvirker tendensen til [w]-bortfald. Grønnum (2005) tolker processen som et resultat af lytterens reanalyse, hvor man tilskriver det velære i approksimanten [w] til det efterfølgende [ø], der i københavnsk oftest er velariseret, og denne velarisering drukner så rundingen i [w], der derfor udelades i den lyttendes efterfølgende brug af ord med oprindeligt [wøð] (Grønnum 2005: 334). Processen beskrives også i Brink & Lund (1975: 346-352), hvor den behandles sammen med udeladelse af [j] i samme lydlike kontekst som en del af en monoftongeringsproces, der begynder allerede hos informanter født i 1840'erne. De første [w]-bortfald konstateres hos informanter født i 1880'erne, og hvor bortfald af [j] konstateres som gennemført hos talere født i 1930'erne eller senere, rapporterer Brink & Lund (1975), at udeladelse af [w] endnu ikke er obligatorisk før [ø], heller ikke hos deres alleryngste informanter.

Brink & Lund (1975) bemærker i øvrigt, at den foranstående vokals kvalitet og stavelsens tryk påvirker tendensen til [w]-bortfald. De konstaterer aldrig [w]-bortfald efter [ɑ ʌ] eller [ɑ:] og kun sjældent efter [ɑ:] og [i] eller [i:]. Bortfald efter [i] finder de kun hos yngre generationer, oftest i svagtryksstilling. I beskrivelsen hos Grønnum (2005) gives dog flere eksempler på [w]-bortfald efter [i] eller [i:] i trykstærk stilling. Dvs. at hvor det tidligere som regel var sådan, at der var [w] i 'livet', men ikke i 'kirkelivet' eller 'privatlivet', er det tilsyneladende nu mere udbredt at udelade [w] i alle tre ord. Disse faktorer må der naturligvis tages højde for i nærværende undersøgelse, så både trykgrad og foranstående vokals kvalitet er blevet inddraget i analyserne, jf. nedenfor.



## 2.2 Lydforandring – hos hvem?

I 1980'erne blev talesproget i Nyboder undersøgt i BySoc-undersøgelsen (Gregersen & Pedersen 1991). Denne (og andre) sociolingvistiske talesprogsundersøgelser er ved at blive gentaget, hvilket gør det muligt ikke blot at undersøge, om den tidligere observerede tendens til [w]-bortfald spredtes fra generation til generation, men også, om den spredtes i den enkelte talers sprog gennem livet. Disse optagelser bruges i apparent time- og real time-undersøgelserne af [w]-bortfald. Jeg fokuserer på talerne fra de to aldersgrupper i middelklassen, fordi jeg gerne vil undersøge, om [w]-bortfald er ved at være helt gennemført i det talesprog, som af mange opfattes som standarden for i hvert fald københavnsk udtale, eller det, man kunne kalde det "bedste" sprog (jf. Kristiansen 2009). I det omfang den københavnske middelklassens udtale udgør en uofficiel norm for dansk udtale, er det interessant at høre, om en gruppe af mennesker, der er udvalgt til at være repræsentative for netop denne samfundsgruppe, udviser de tendenser til forandring i udtalen, som er blevet observeret tidligere. Hvis denne lydforandring også er almindelig i middelklassens talesprog, kan det tages som et tegn på, at [w]-bortfald er ved at blive normen snarere end undtagelsen i almindelig daglig tale. Det må dog understreges, at en kvantitativ undersøgelse af middelklassens sprogbrug ikke i sig selv kan sige noget om, hvordan denne varierende udtale opfattes – det kræver naturligvis perceptions- og holdningsundersøgelser. Blot er det sandsynligt, at [w]-bortfald vil være på vej til at blive den accepterede norm, hvis den dominerer middelklassens sprog.

I den oprindelige BySoc-undersøgelse deltog 22 informanter fra middelklassen. 12 af dem var født mellem 1946 og 1962, og 10 af dem var født mellem 1967 og 1973, og de to grupper udgør her hhv. den ældre og den yngre generation. Talerne i hver generation var ligeligt fordelt på køn. De blev alle optaget på bånd i deres hjem i 1987 i samtale med en interviewer fra Københavns Universitet, og det er disse optagelser, der udgør grundlaget for min apparent time-undersøgelse. Optagelserne er foretaget mere end 10 år efter de sidste optagelser, der rapporteres hos Brink & Lund (1975), og giver således mulighed for at undersøge, om processen skulle være blevet endeligt gennemført i den mellemliggende periode.

Fordi BySoc-undersøgelsen er blevet gentaget, er det som nævnt også muligt at undersøge, om tendensen til [w]-bortfald styrkes eller svækkes, når sprogbrugerne bliver ældre. Det er muligt, fordi gentagelsen er udført med nøjagtigt de samme 22 informanter, der deltog i den oprindelige undersøgelse, sådan at man kan sammenligne den enkelte sprogbrugers udtale i 1987 med hans eller hendes udtale i 2006. Disse genoptagelser af informanterne danner grundlaget for mit real time-panelstudie.

### 2.3 Fremgangsmåde

De rå data består af auditivt klassificerede forekomster af det, der herefter omtales som variabelen (wøð)<sup>3</sup>. Variablen omfatter alle forekomster af ord, der i distinkt udtale indeholder fonsekvensen [wøð], fx 'levet, lavet, revet, livet'. Når [ø] assimileres til [ð], der derved bliver stavelsesbærende, [ð], opstår muligheden for bortfald af den distinkte udtales [w] (jf. redegørelsen refereret fra Grønnum 2005 ovenfor). Bemærk, at oprindeligt [w] i moderne udtale falder væk efter de høje bagtungevokaler [u o], uanset om det efterfølges af [ð] eller ej. Sådanne forekomster er derfor ikke medtaget i undersøgelsen, hvilket i praksis betyder, at alle forekomster af ordet 'overhovedet' er udeladt fra datagrundlaget. Alle andre forekomster af (wøð) blev kodet for bortfald. Processen betragtes her som binær, dvs. enten blev [w] klassificeret som bortfaldet eller som bevaret. Aflytningen blev foretaget af mig selv og en anden trænet aflytter med en baggrund i lingvistik og fonetik, således at data ikke kun hviler på observationer fra én lytter. Kun forekomster, hvor vi var enige, er taget med i de statistiske analyser.

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 princippet 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] før [ð] optræder i stavelser, der er specificeret for leksikalsk nultryk. Hver forekomst i datasættet 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 frekvensundersøgelse af ord i løbende tale. Frekvensberegningerne er foretaget på baggrund af de sociolingvistiske interview i DGCCS' 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 ords distinkte udtale, var det ikke nødvendigt at kontrollere for homografi. Beregningerne afspejler således direkte de distinkte udtalte ordformers relative hyppighed. Frekvenserne er efterfølgende blevet konverteret til en logaritmisk skala, og hver forekomst af (wøð) i de aflyttede 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 køn og fødselsår er blevet trukket ud af DGCCS' database, sådan at sammenhængen mellem køn og [w]-bortfald og ikke mindst alder og [w]-bortfald kunne analyseres.

Alle disse faktorer, både sproginterne, sprogbrugs- og de sociale faktorer, her køn og alder, da alle informanter var klassificeret som middelklassekøbenhavnere, 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 kontrollerer for alle de andre faktorer. Den metode beskrives kort nedenfor.

#### 2.4 *At sammenligne fordelinger af varianter*

I den statistiske analyse af de klassificerede forekomster af variabelen (wøð) har jeg brugt multipel logistisk regression med blandede effekter. Det er en statistisk metode, der tillader, at man undersøger flere faktoreres indflydelse samtidig, og derudover giver mulighed for, at man kan tage højde for de enkelte taleres forskellige adfærd. Metoden er ikke ret udbredt i dansk (socio)lingvistik, og derfor giver jeg her en meget kort introduktion til metoden, og hvordan analyseresultaterne vil blive præsenteret i artiklen. Den interesserede læser kan sætte sig ind i metoden, og hvordan den bruges, ved at læse fx Johnson (2008) eller Baayen (2008).

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-prøver for hver af de relevante kontekster<sup>4</sup>. 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 tale tempo, så derfor kan jeg ikke undersøge, om tend-

ensen til [w]-bortfald stiger med taletempoet (hvad man ellers vil forvente). Men jeg kan tage højde for sproginterne faktorer i mine analyser af sprogbrugsfaktorer, og jeg kan også undersøge effekten af disse sproginterne faktorer. For hver faktor får man nemlig ikke bare en oplysning om, hvorvidt den bidrager signifikant til at forudsige [w]-bortfald, hvilket angives af faktorens p-værdi. Man får også at vide, om faktoren øger eller sænker sandsynligheden for [w]-bortfald. Denne effekt udtrykkes med den såkaldte  $\beta$ -koefficient, der er et statistisk mål for, hvordan en given faktor påvirker sandsynligheden for udfaldet af den proces, man har målt. Når  $\beta$ -koefficienten er positiv, indikerer det, at faktoren øger sandsynligheden for, at den givne proces finder sted, mens en negativ  $\beta$ -koefficient angiver, at faktoren sænker sandsynligheden. Altså, hvis man vil vide om talerens alder har en statistisk signifikant indflydelse på om oprindeligt [w] udelades i udtalen, skal man først kende p-værdien for faktoren "alder" i den statistiske model for [w]-bortfald. Når p-værdien er under 0,05 kan faktoren "alder" siges at være tilstrækkeligt solid. Hvis man vil vide, hvordan faktoren "alder" påvirker sandsynligheden for [w]-bortfald, skal man kende faktorens  $\beta$ -koefficient – en positiv værdi vil betyde, at sandsynligheden for [w]-bortfald stiger, jo yngre taleren er, og en negativ  $\beta$ -koefficient vil betyde, at sandsynligheden falder, jo yngre taleren er. I præsentationen af de kvantitative analyser vil jeg derfor angive både  $\beta$ -koefficient og p-værdi for de faktorer, der indgår i undersøgelsen.

Korpusundersøgelser af udtalevariation giver ikke kun problemer med at kontrollere for den sproglige kontekst. 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 variabelen (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 vokals 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 talerkaraktistika 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 kvinder, vil vi meget nødig 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 sociolinguistikken og er ikke helt ukontroversielt. Sociolinguistikken 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 spørgsmålet 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 implicite 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 uønsket 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-prøver på dele af data.

### **3 BORTFALD AF [W] I 1987 – APPARENT TIME-UNDERSØGELSEN**

Her gennemgås resultaterne af den statistiske analyse af (wød) 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 inddrager jeg ordfrekvenser for at kunne besvare spørgsmålene om, hvorvidt en igangværende lydforandring spreder 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 fødselsår en signifikant faktor i modelleringen af [w]-bortfald ( $p = 0,025$ ), og regressionskoefficienten for fødselsår er positiv ( $\beta = 0,0698$ ), dvs. at sandsynligheden for [w]-bortfald stiger med fødselsår. 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 datasættet 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 variantsætningen (jf. afsnit 2.3 ovenfor), og forekomster kunne enten være trykstærke eller trykløse. Samtidig er hver enkelt ordform specificeret for leksikalsk tryk, altså det tryk, som ordet er ”født 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 specificeret 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 optrådte i. Der er naturligvis en vis korrelation mellem opfattet og leksikalsk tryk, idet 95



% af de stavelser i materialet, der har bitryk, realiseres som trykløse, ganske som man ville vente. Men der er også mange af de stavelser, som leksikalsk er specificeret for hovedtryk, der realiseres som trykløse. Størstedelen af disse kan tilskrives forekomster af ordet 'blevet', der oftest realiseres trykløst 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 leksikalsk, 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 leksikalsk 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 trykløse stavelser end i trykstærke ( $\beta = 1,29$  og  $p = 0,005$ ). I de trykløse stavelser er dog sådan, at sandsynligheden for [w]-bortfald falder i stavelser med leksikalsk bitryk ( $\beta = -2,22$  og  $p < 0,0001$ ). Hos den yngre generation er det kun leksikalsk tryk, der bidrager signifikant til modellen ( $\beta = -2,17$  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.

Det betyder, at stavelser, der er blevet opfattet som og altså realiseres som trykløse, generelt har højere tendens til [w]-bortfald, dvs. [w]-bortfald er mere sandsynligt i 'blevet' end i 'levet', medmindre den trykløse stavelse leksikalsk er specificeret for bitryk. I stavelser med bitryk er tendensen til [w]-bortfald lavere, også selvom stavelsen realiseres som trykløs. Sandsynligheden for [w]-bortfald er altså størst i et ord som 'blevet', mindre i et ord som 'oplevet', hvor stavelsen har bitryk, og mindst i et ord som 'levet', der har fuldtryk. Selvom trykket langt fra beskytter mod bortfald af [w], så er der alligevel en tendens til at [w] oftere bevares, når stavelsen har tryk. Denne tendens er forsvundet hos den yngre generation. For dem er der næsten lige stor sandsynlighed for bortfald af [w], uanset om stavelsen har tryk eller ej. Det tyder på, at selvom forandringen endnu ikke er gennemført fuldstændig hos den yngre generation, så har en af de faktorer, der før begrænsede tendensen til [w]-bortfald, mistet sin betydning, og at det nu kun er det leksikalsk specificerede tryk, der spiller en rolle.

### 3.2 [w]-bortfald og foranstående vokals kvalitet

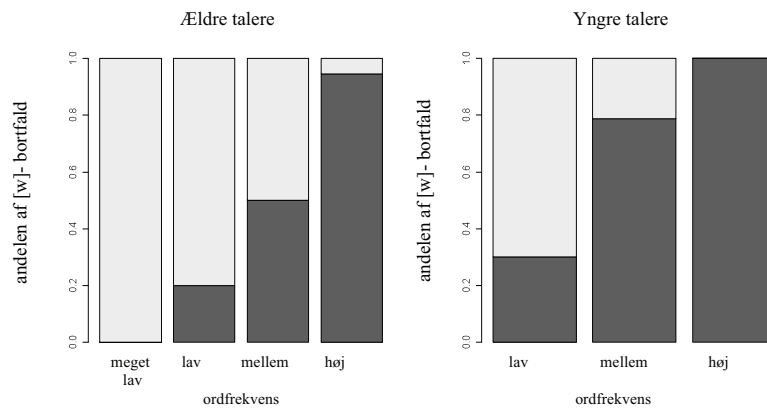
Selvom der er stor tendens til [w]-bortfald, så er det stadig sådan for begge generationer at [w] ikke altid falder bort, heller ikke i stavelser uden tryk hos de yngre. Som nævnt viser tidligere undersøgelser, at [w] oftere falder bort efter fortungevokaler end efter bagtungevokaler, medmindre vokalerne er høje, for her var tendensen i hvert fald tidligere modsat. Det var altså mere almindeligt, at [w] faldt bort efter [e, ø, ε] og [æ], dvs. i ord som 'levet, løbet, slæbet, lavet' end det gjorde efter [ɔ, ʌ, ɒ] og [ɑ] altså i ord som 'toget, tovet, torvet, skrabet'. Den forskel går igen hos begge generationer i optagelserne fra 1987: Hvor 83 % af de [w]'er der forekommer efter en fortungevokal, bliver slettet, 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  $\beta = 3,35$ ) og for de yngre talere ( $p = 0,007$ ,  $\beta = 5,79$ ).

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  $\beta = 3,537$  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.

FIGUR 1 – [w]-BORTFALD SOM EN FUNKTION AF ORDFREKVENS I 1987



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 højfrequente ord som 'lavet', knap så hyppigt i mellemfrequente ord, som 'skrevet', og meget sjældent i sjældne ord som 'revet'. 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 mellemfrequente 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 højfrequente 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 'blevet' udtalt ['ble:ðð], dvs. med tryk og [w]-bortfald, mens det er mindre sandsynligt at høre 'levet' udtalt ['le:ðð].

På gruppeniveau er der således støtte for den sprogbaserede fonologiske hypotese om, at reduktioner spredt 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 Individuer og ordfrekvens

Det er ikke muligt at lave en statistisk signifikanstestning 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 produceret mere end 10 forekomster af variabelen (wød).

Gruppen af ældre talere kan indeles 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]-BORTFALD SOM EN FUNKTION AF ORDFREKVENNS HOS DE ENKELTE TALERE I DEN ÆLDRE GENERATION 1987

Informant	Antal ord	% bortfald Meget lavfrekvente ord	% bortfald Lavfrekvente ord	% bortfald Mellemligfrekvente ord	% bortfald Højfrekvente ord
LAL	11	-	0	0	75
TNI	33	-	33	50	87
PTK	38	0	0	37	94
CNI	20	-	-	20	100
ASA	11	0	-	40	100
EAF	11	-	50	-	100
HTH	20	-	0	50	100
MJE	15	-	0	50	100
CEL	17	-	100	100	87
MFL	22	0	0	100	90
LKR	6	-	-	-	-
PFR	9	-	-	-	-

Informanterne er ordnet efter deres tendens til [w]-bortfald i højfrekvente ord, dvs. jo højere oppe i tabellen, desto lavere tendens til at udelade [w] før [ø] i udtalen af højfrekvente ord. Den første informant, LAL, kan betegnes som en konservativ taler, idet hun har meget lav tendens til [w]-bortfald. Derimod har de næste syv informanter en relativt ensartet og høj tendens til [w]-bortfald. Processen er obligatorisk i højfrekvente ord for næsten alle disse talere, og tendensen falder jævnt med ordfrekvens. For alle disse otte talere, altså såvel den konservative som de gennemsnitlige talere, kan frekvenseffekten vises at fungere på individniveau. Men for de to næste talere er mønsteret et andet: CEL og MFL har begge næsten obligatorisk [w]-bortfald bortset fra i enkelte højfrekvente ord. De går altså imod frekvenseffekten og viser, at den automatiske rutinisering, der er resultatet af gentagen brug af bestemte artikulatoriske gesti, kan overskygges af andre faktorer. Den enkelte taler er m.a.o. ikke uden fri vilje, selvom frekvenseffekten er stærk.

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 gruppeniveau. 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 [ø]. 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

Informant	Antal ord	% bortfald		
		Lavfrekvente ord	Mellemfrekvente ord	Højfrekvente ord
MIP	18	-	50	100
JJE	12	-	67	100
PKJ	20	0	75	100
DBE	17	0	80	100
JOR	14	-	100	100
MPT	16	-	100	100
ASH	9	-	-	-
KKJ	4	-	-	-
NOR	6	-	-	-
SKO	7	-	-	-

Tabellen viser, at alle talere i den yngre generation har obligatorisk [w]-bortfald i højfrekvente ord. De udtaler således altid 'blevet' som ['ble:ðð] og aldrig som ['ble:wð] eller ['ble:wəð]. 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 frekvens-effekten.

Ud over at disse fordelinger for individer viser, at frekvens-effekten 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] før syllabisk [ð] 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] før [ð]. 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 bør 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 reduktionsprocesser i andre sprog, følger ordfrekvens både på gruppeniveau og hos den en-

kelte taler, også når man i analysen kontrollerer for sproginterne faktorer. Analysen støtter således hypoteserne fra den sprogbrugsbaserede fonologi, men hvad med antagelsen bag apparent time-hypotesen, altså den, at sprogbrugere forbliver stabile i deres udtale op gennem voksenlivet? For at kunne teste den er det nødvendigt at undersøge variabelen (wøð) i real time.

#### **4 BORTFALD AF [w] I 2006 HOS DE SAMME TALERE – REAL TIME-UNDERSØGELSEN**

Undersøgelsen af tendenserne i 1987 viser altså, at bortfald af [w] er meget almindeligt, selvom det endnu ikke er obligatorisk, i afslappet løbende tale. Spørgsmålet er nu, hvordan det er gået med [w] siden da hos de samme talere: Har de fået større tendens til at slette deres [w] før [ø]? For gruppen som helhed er svaret ja, og ændringen gælder for begge generationer. Dvs. der er en generel tendens til, at [w]-bortfald intensiveres hos de samme talere fra midten af 1980'erne til starten af 2000'erne.

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 trykløse stavelser end i trykstærke ( $p = 0,0006$ ), medmindre den trykløse 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 (hhv.  $p = 0,3$  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 gør sig gældende for den nye generation af unge københavnere fra middelklassen, og ikke mindst om processen er endnu mere fremskreden efter bagtungevokaler, som for den yngre generation i dette datasæt 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 ( $\beta = 1,15$  for de ældre og  $\beta = 1,56$  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 'pivot' 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 interviewene 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 gruppetilhørsforhold, her generationsmedlemskab, og adfærd i løbet af livet.

Kravet om statistisk signifikans kan ikke fraviges, når det skal undersøges om enkeltpersoner forbliver stabile eller følger sprogsamfundets tendens og intensiverer [w]-bortfald i deres sprogbrug. Derfor blev



tendensen til [w]-bortfald signifikantstestet 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 (wøð) FOR DE TALERE DER ÆNDRER SIG OVER TID

Infor- mant	Antal med [w] 1987	Antal uden [w] 1987	% uden [w] 1987	Antal med [w] 2006	Antal uden [w] 2006	% uden [w] 2006	p	Gene- ration	Køn
DBE	5	12	70 %	3	38	92 %	0,04	Yngre	Kvinde
LAL	8	3	27 %	0	10	100 %	0,001	Ældre	Kvinde
PTK	16	22	58 %	4	29	88 %	0,007	Ældre	Mand

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 talere 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 'revet' som [ˈvæ:wøð], udtaler hun det nu [ˈvæ:ðøð]. DBE er det 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 (wøð) 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 (wøð) 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 leksion.

## 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 minde 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 variationssociolingvistik. At sammenhængen mellem [w]-bortfald og ordfrekvens kan ses på individniveau, støtter den anskuelse, at sprogbrugernes implicite 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 Sprogforandringsstudier  
Københavns Universitet  
nicolajp@hum.ku.dk

## NOTER

- 1 I denne artikel lydskrives med modificeret IPA i henhold til den konvention, der anvendes i Grønnum (2005), jf. oversigten ”Nøgle til ydskrift” i dette nummer af NyS.
- 2 Hvor fonetikerer beskriver en lydforandring ved at sammenligne, hvilken sproglyd der tidligere blev brugt i en given lydig 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 enkeltlyde, der bruges i tale, og de skrives i // -klammer.
- 3 Jeg følger her den sociolingvistiske konvention og angiver variabelen i runde parenteser ( ) i modsætning til variabelens varianter, som angives i lydskrift, og her kan være [w̥] eller [̥].
- 4 Chi-i-anden-prøven 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 tilfældighed, 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

- Baayen, R. Harald (2008) *Analyzing Linguistic Data*. Cambridge: Cambridge University Press.
- Bybee, Joan (2002) *Phonology and Language Use*. Cambridge: Cambridge University Press.
- Brink, Lars & Jørn Lund (1975) *Dansk Rigmål 1- Lydudviklingen siden 1840 med særligt henblik på sociolekttern i København*. København: Gyldendal.
- Gregersen, Frans (2009) "The data and design of the LANCHART study". *Acta Linguistica Hafniensia* 41. 4-37.
- Gregersen, Frans & Inge Lise Pedersen (red.) (1991) *The Copenhagen Study in Urban Sociolinguistics 1*. København: C.A. Reitzel.
- 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.
- Grønnum, Nina (2005) *Fonetik og Fonologi – Almen og Dansk*, 3. udg. København: Akademisk Forlag.
- Johnson, Keith (2007) "Decisions and Mechanisms in Exemplar-based Phonology". M.-J. Solé, P. Speeter Beddor & M. Ohala (red.) *Experimental Approaches to Phonology*. Oxford: Oxford University Press. 25-40.
- Johnson, Keith (2008) *Quantitative Methods in Linguistics*. Blackwell Publishing, Malden, MA, USA.
- Kristensen, Kjeld (1977) "Variationen i vestjysk stationsbymål". *Dialektstudier* 4, 1. halvbind.
- Kristensen, Kjeld & J. Normann Jørgensen (1994) *Moderne sjællandsk*. København: C.A. Reitzel.
- Krisitansen, Tore (2009) "The macro-level social meanings of late-modern Danish accents". *Acta Linguistica Hafniensia* 41. 167-192.
- Labov, William (1994) *Principles of Linguistic Change – Internal Factors*. Oxford: Blackwell.
- Labov, William (1966(2006)) *The Social Stratification of English in New York City*, 2. udg. Cambridge: Cambridge University Press.
- Levelt, Willem J.M., Ardi Roelofs & Antje S. Meyer (1999) "A theory of lexical access in speech production". *Behavioral and Brain Sciences* 22. 1-75.

- Pierrehumbert, Janet (2002) "Word-specific phonetics". C. Gussenhoven & N. Warner (red.) *Papers in Laboratory Phonology 7*. Berlin: Mouton de Gruyter. 101-140.
- Zwicky, Arnold (1972) "Note on a phonological hierarchy in English". R. P. Stockwell & R.S. Macauley (red.) *Linguistic change and generative theory*. Bloomington, USA: Indiana University Press. 275-301.