

En høj, en lav? En stigning, et fald?

En undersøgelse af ændringer i det næstvedske trykgruppemønster

CHRISTOPHE ZERAKITSKY VIES OG NICOLAI PHARAO

ABSTRACT

Det primære træk, der auditivt adskiller danske regiolekter fra hinanden, er intonation, nærmere bestemt trykgruppemønsteret. Næstvedsk intonation er særlig interessant, da trykgruppemønsteret akustisk minder meget om københavnsk, idet begge starter med et kort fald efterfulgt af en stigning, og begge slutter på et fald. Dog lyder de to ganske forskellige, sandsynligvis fordi det initiale fald er dybere i københavnsk, mens det finale fald sætter ind tidligere og er stejlere i næstvedsk. I nyere optagelser fra Næstved lyder nogle af talerne umiddelbart københavnske. En sådan uoverensstemmelse kunne tyde på en igangværende sprogforandring. Med henblik på at undersøge dette udførte vi akustiske analyser af optagelser fra Næstved i 1980'erne og i 00'erne. Formålet var at undersøge, om man kan observere en sprogforandring i Næstved i løbet af de tre generationer, der er repræsenteret i korpusset. Vi sammenlignede tre generationer i de nyeste optagelser, og derudover lavede vi en real time-undersøgelse af de ældste talere, hvor vi sammenlignede deres gamle optagelse med den nye. Vores resultater indikerer, at trykgruppemønsteret har bevæget sig mere i retning af det københavnske mønster, men vi kan ikke vise, at de er sammenfaldende. Der er en tydelig tendens til forandringer på tværs af de tre generationer, men der er også stor interindividuel variation, så særligt real time-forandringerne må siges at være træk ved individer og ikke ved de forskellige grupper. Dette kan forklare, hvorfor nogle talere fra Næstved lyder københavnske, mens andre ikke gør.

EMNEORD: intonation; sprogforandring; dialekter; sociofonetik, Generalised Additive Mixed Models

Hvor stabilt er vores sprog? Vi ved, at sprog forandrer sig, og i hverdagen hører vi forskelle mellem yngre og ældre menneskers sprog, som gør det klart, at ikke alle dele af det kan være lige stabile. Særligt ordvalg og udtale springer os i ørerne og minder os om sprogets foranderlighed. En

meget stor mængde faglitteratur viser, at sprogforandring foregår mellem generationer og kan observeres nu og her i tillæg til sproghistoriske undersøgelser, der dokumenterer forandringer over århundreder. En mindre men voksende mængde faglitteratur har dokumenteret, at også den enkeltes sprog kan forandre sig gennem livet. Det er altså ikke bare dansk udtale, der ændrer sig, men også Annes og Peters udtale, der kan ændre sig gennem livet.

I denne undersøgelse fokuserer vi netop på udtalen, helt specifikt tonegangen i forbindelse med trykstærke stavelser. Vi undersøger den hos tre generationer af talere, der alle er født og opvokset i Næstved, og vi undersøger den både på tværs af generationer og hos enkeltpersoner på to forskellige tidspunkter i deres liv. Tonegangen er særligt interessant at undersøge i dansk, fordi vi ved, at den varierer systematisk på tværs af regiolekter og dermed kunne forventes at være mere stabil end f.eks. enkeltlyde. Samtidig findes der kun få undersøgelser af forandringer i tonegangen, og ingen kvantitative undersøgelser er tidligere foretaget på grundlag af akustiske målinger af grundtonen.

1 TID OG TONEGANG

Undersøgelsen af igangværende sprogforandringer er central i variationistisk sociolingvistik. Langt størstedelen af disse undersøgelser, både af dansk og andre sprog, er foretaget ved sammenligninger af sproglig variation på tværs af generationer, såkaldte *apparent time*-undersøgelser (Sankoff 2006). Logikken i denne tilgang er, at en forskel i 60-årige og 20-åriges sprogbrug, f.eks. deres udtale af den trykstærke vokal i ord som *lande*, *kasse*, *fatte*, eller deres brug af *du* som generisk pronomen, kan tolkes som en sprogforandring, der er foregået i løbet af årtierne mellem de to generationer. *Apparent time*-tilgangen hviler på den antagelse, at folk fortsætter med at tale, som de gjorde, da de var unge, gennem hele livet, dvs. at hvorimod sproget som et fælles system menes at være foranderligt, så er den enkelte sprogbrugers sprog stabilt. En voksende mængde undersøgelser viser dog, at ikke alle menneskers sprog er så stabilt gennem hele livet. Dette har vist sig gennem en række *real time*-undersøgelser, hvor den samme talers sprogbrug på forskellige tidspunkter bliver sammenlignet (Sankoff 2006). Sådanne undersøgelser kræver adgang til samlinger af optagelser foretaget med de samme

personer med flere års mellemrum, således at alt undtagen den observerede sprogbrug holdes konstant.

En sådan samling optagelser er tilgængelig gennem LANCHART-korpusset (Gregersen m.fl. 2014), som omfatter sociolingvistiske interview fra flere forskellige steder i Danmark med sprogbrugere fra tre forskellige generationer. Optagelserne kan klassificeres som spontan tale, da talerne ikke blev instrueret i, hvad de skulle sige, og samtalerne foregik, hvis man ser bort fra, at de blev optaget, som frie samtaler, omend interviewerens spurgte ind til bestemte emner (se yderligere Gregersen 2009). De to ældste generationer, i korpusset benævnt generation 1 og generation 2, består af talere, der deltog i variationistiske undersøgelser i 1980'erne. De samme personer blev optaget igen i 00'erne, hvor man samtidig indsamlede sociolingvistiske interview med unge født i 90'erne. Herved giver LANCHART-korpusset mulighed for at foretage både *apparent time*-undersøgelser på tværs af de tre generationer ved at sammenligne deres sprogbrug i 00'erne, og desuden også *real time*-undersøgelser ved at sammenligne sprogbrugen hos talerne i generation 1 og 2 i 80'erne med de samme taleres egen sprogbrug i 00'erne, sådan som det f.eks. er gjort i Gregersen m.fl. (2014), Gregersen & Pharao (2016) og Maegaard m.fl. (2013). Der er meget at tage højde for i tilrettelæggelsen af begge typer undersøgelser, og interesserede læsere kan dykke dybere ned i disse overvejelser i de omtalte artikler. Det væsentlige for nærværende undersøgelse er, at vi har haft adgang til dette materiale og derfor kan undersøge, om der er sket en ændring i trykgruppemønstret blandt sprogbrugere fra Næstved.

Vi er interesserede i at undersøge forandring i tonegangen i næstvedsk, helt specifikt grundtoneforløbet i trykgruppen, fordi vi fra tidligere undersøgelser ved, at dette forløb varierer systematisk på tværs af regiolektter i Danmark (Grønnum 2022). Det er væsentligt for genkendelse af, hvor talere kommer fra (Kristiansen m.fl. 2011, Tøndering & Pharao 2020). Det er således et væsentligt sociolingvistisk træk i et sprogsamfund præget af udbredt dialektnivellering, og derfor bliver det særligt interessant at undersøge, om forløbet er så stabilt, som man hidtil har antaget. Trykgruppemønstret er en komponent i den model for dansk intonation, som Nina Grønnum (Thorsen) har udviklet på baggrund af omfattende studier af oplæst tale fra 80'erne (Grønnum 1992, 2022), og som siden er blevet undersøgt videre i spontan tale af John Tøndering

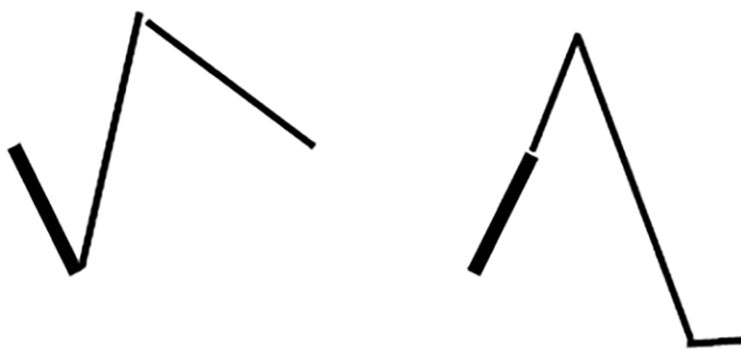
(2010a). Trykgruppemønsteret er det grundtoneforløb, der initieres af en trykstærk vokal og knytter sig til alle efterfølgende tryksvage stavelser i en ytring frem til den næste trykstærke vokal (som initierer en ny trykgruppe). Dette trykgruppemønster er kvalitativt invariant i oplæst tale, om end omfanget af stigninger og fald i grundtonen inden for trykgruppen kan variere som følge af bl.a. placering i ytringen, og om de ord, der omfattes af trykgruppen, er mere eller mindre prominente. En række undersøgelser af Grønnum viser som nævnt, at trykgruppemønsteret varierer fra regiolekt til regiolekt i Danmark. I københavnsk-baseret dansk begynder trykgruppen med et fald ind i den trykstærke vokal og efterfølges af en stigning, der når sin top i den første posttoniske stavelse. I århusiansk-baseret dansk bliver den trykstærke stavelse udtalt med en stigende tone efterfulgt af et fald til første posttoniske stavelse. Dyhr (1993) og Dyhr (1995) undersøgte trykgruppemønstre i spontantale i københavnsk-baseret rigsmål og fandt også, at trykgruppemønsteret er kvalitativt invariant i spontantale. Det samme gør Scheuer (2019), som har undersøgt mønstret i optagelser af fynsk.

Nærværende undersøgelse består af akustiske analyser af trykgruppemønsteret hos sprogbrugere fra Næstved. Det er baseret på spontan tale, idet vi bruger de sociolingvistiske interview fra LANCHART-korpusset, og undersøgelsen vil således bidrage med yderligere viden om trykgruppemønsteret i hverdagssproget. Vi har valgt at undersøge potentielle forandringer netop i Næstved, fordi det i Tøndering & Pharao (2020) bemærkes, at forfatterne under udarbejdelsen af stimuli til deres lytteforsøg blev opmærksomme på, at det var svært at skelne udtalen hos mandlige talere fra middelklassen i Næstved fra københavnske talere, hvorimod mandlige talere fra arbejderklassen i Næstved fortsat havde et tydeligt regiolektalt præg i deres udtale (Tøndering & Pharao 2020 undersøgte ikke kvindestemmer). Denne forskel indikerer, at der potentielt var en sprogforandring på vej i Næstved i 00'erne, og det er det, vi her har undersøgt systematisk. Desuden er der detaljer i analysen af trykgruppemønsteret i Næstved i den undersøgelse, Nina Grønnum (Thorsen) beskriver i Thorsen (1989), som antyder, at en sprogforandring kan have været i gang allerede på det tidspunkt. Vi begynder med at se nærmere på den undersøgelse, inden vi beskriver vores eget undersøgelsesdesign og vores datagrundlag.

2 TIDLIGERE BESKRIVELSER AF TRYKGRUPPEN I SJÆLLANDSK REGIOLEKT

Grønnum (2022) præsenterer en oversigt over trykgruppemønstre i en række danske regionalsprog, herunder trykgruppemønstret i København og i Næstved. Bag denne oversigt ligger en række artikler med detaljerede analyser og beskrivelser af trykgruppemønstret i enkelte regiolekt. Ifølge Grønnum (Thorsen 1989: 75–76) kan trykgruppemønsteret i København og Næstved stiliseres som vist i figur 1 (fra Grønnum 2005: 341). Stiliseringen af det næstvedske trykgruppemønster i figuren er baseret på akustiske målinger samt Grønnums auditive indtryk af grundtonekonturen (Thorsen 1989: 75–76). Mønstret i de gennemsnitlige kurver beskrives som ”rising-falling” og perceptuelt ”high-low”, hvilket dog i samme artikel ændres til ”low-high-low” ((Grønnum) Thorsen 1988: 190). Til sammenligning karakteriserer Grønnum københavnsk som ”low-high-declining” (Thorsen 1989: 76) eller bare ”low-high” (Thorsen 1988: 190).

FIGUR 1. STILISEREDE TRYKGRUPPEMØNSTRE I KØBENHAVNSK (TIL VENSTRE) OG NÆSTVEDSK (TIL HØJRE)



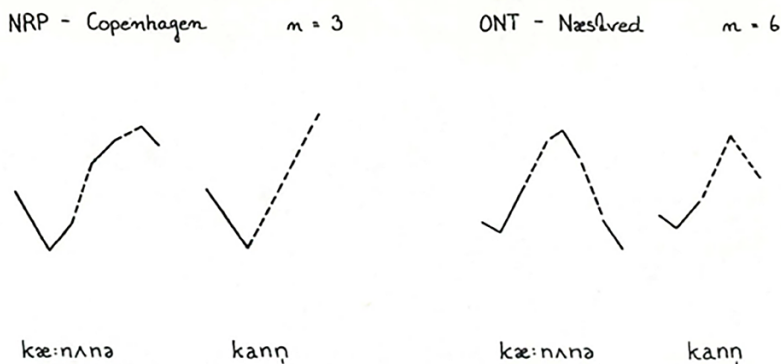
Den fede linje angiver den betonedede stavelse, og resten af linjen angiver følgende ubetonede stavelser. Figuren er lånt fra Grønnum (2005: figur 18.4, s. 341).

Ser man imidlertid på figur 2 nedenfor, der er en reproduktion fra den oprindelige undersøgelse af intonationen i Næstved (Thorsen 1989), ses en sammenligning af den gennemsnitlige F0-kontur i to ord taget fra en taler fra København og en taler fra Næstved. Her får vi en idé om, hvorfor stiliseringen i figur 1 ikke umiddelbart passer helt til beskrivel-

sen ”low-high-low”, men snarere passer med de akustiske observationer af trykgruppemønsteret i Næstved. I sammenligning med københavnsk har taleren fra Næstved et meget lille initialt fald i den betonedede stavelse, og selv i *kanden*, der har kort vokal, påbegyndes stigningen i F0 allerede i den betonedede vokal, hvorimod stigningen starter senere i københavnsk. Det er sandsynligt, at det initiale fald ind i den betonedede stavelse, der ses for taleren fra Næstved i figur 2, skyldes en påvirkning fra den foranstående ustemte klusil (jf. Kirby & Ladd 2016). Ser man nærmere på de gennemsnitskurver, der præsenteres for talere fra Næstved i (Grønnum) Thorsen (1988: 157–160), genfindes dette initiale fald oftest i betonedede stavelser, der begynder med en aspireret (eller affrikeret) lukkelyd (i navnene ’Kamma’ og ’Torben’), men også i trykgrupper der begynder med ustemt frikativ (’søster’ og ’Fakse’), ustemt lukkelyd (’stammer’), sonorant (’Næstved’) og endda med vokal (’Anders’, om end dette kun observeres ytringsinitialt og formodentlig kan tilskrives sprængsans). Dette fald er dog langt fra obligatorisk (jf. (Grønnum) Thorsen 1988: 183–187). Vi mener, at dette underbygger den perceptuelle fortolkning af trykgruppemønsteret i næstvedsk som ”high-low”, idet det initiale fald kan tænkes at være blevet tilskrevet klusilens glot-tiskarakteristik.

I opsummeringen af sin akustiske undersøgelse af næstvedsk skriver Grønnum videre, at ”I have termed the Næstved stress group pattern one of the perceptually high-low type. I would like to revise that, and make a distinction between low-high patterns (Bornholm and Copenhagen), low-high-low (Næstved) ...” (Thorsen 1988: 190). Vi har altså at gøre med en situation, hvor akustiske målinger af grundtonen fortolkes anderledes i perceptionen af trykgruppemønsteret, da det lille fald ind i den betonedede i stavelse i Næstved ikke af Grønnum opfattes på samme måde. Det er som antydnet formentlig, fordi dette fald, som kan observeres i de akustiske data, i langt de fleste tilfælde udløses af ustemte (og ofte aspirerede) obstruenter før den betonedede vokal i trykgruppen. Spørgsmålet for os bliver nu, om der her kan ligge kimen til en lydforandring, m.a.o. en perceptuel omfortolkning af den intrinsiske egenskab til et primært cue.

FIGUR 2. SAMMENLIGNING AF TRYKGRUPPEMØNSTRET I KØBENHAVNSK OG NÆSTVEDSK



Sammenligning af den gennemsnitlige F0-kontur i ordene *kanerne* og *kanden* for en informant fra København (til venstre) og en informant fra Næstved (til højre). Fuldt optrukne linjer angiver vokaler, og stiplede linjer angiver sonorante konsonanter. NRP og ONT er forkortelser, der identificerer talerne i Grønnums (Thorsen) oprindelige undersøgelse”. Figuren er lånt fra Thorsen (1988: figur 30, s. 189).

Det er væsentligt at være opmærksom på, at der med karakteristikkens 'low' hverken hos os og slet ikke hos Grønnum (Thorsen) menes et tonalt *target*, der karakteriserer stavelsen og specificeres som en del af en fonologisk repræsentation, som man kender det fra den såkaldte autosegmentale-metriske model (Ladd 2008, Arvaniti 2022). I denne model behandles intonation som en kombination af høje (H) og lave (L) toner, der associeres med bestemte tonebærende enheder (f.eks. stavelser). Da alle konturer underliggende antages at udgøres af høje og lave primitiver, beskrives stigninger og fald nødvendigvis som kombinationer af disse. Det er ikke denne artikels ærinde at give en dybdegående indføring i autosegmental-metrisk fonologi endsige udvikle en model baseret herpå for dansk. Det er ikke helt klart, hvordan forskellen på det københavnske mønster og det næstvedske mønster ville skulle beskrives i denne formalisme, da grundtonen i den betonedede stavelse for begges vedkommende ligger lavere end i den posttoniske. Hvorom alting er, så står det klart ud fra Grønnums undersøgelser, at der er en forskel i, hvordan tonerne justeres i tidsforløbet i relation til segmenterne: I københavnsk associeres den betonedede stavelse med en lav tone, og den posttoniske associeres med en høj tone. I næstvedsk associeres den betonedede stavelse med en stigning, hvis toppunkt ikke nødvendig-

vis ligger i den betonede stavelse, men stigningen fortsætter ind i den posttoniske stavelse.

Uoverensstemmelsen mellem det auditive indtryk og de akustiske målinger har fået Grønnum til at foreslå en beskrivelse af næstvedsk, der adskiller mønstret fra københavnsk (da de to mønstre lyder forskellige), men som samtidig fanger ligheden mellem de to (fordi de begge akustisk starter med et dyk i F0 efterfulgt af en stigning og til sidst et fald igen). Dette kan ses som et tegn på, at der på undersøgelsens tidspunkt i 1986 var en lydforandring i gang: Der er et kort målbart initialt fald i F0, som ikke tillægges samme vægt auditivt i den senere stilisering af mønstret. Med andre ord kan det tænkes, at trykgruppemønstret i slutningen af 80'erne var ved at introducere det initiale fald ind i den betonede stavelse, som allerede på det tidspunkt var kendetegnende for københavnsk. Grønnums (Thorsen 1988) informanter var fire teenagere (og størstedelen af undersøgelsen er kun baseret på de to kvinder i gruppen) og en 28-årig mand. Spørgsmålet er nu, om F0-konturen for ældre talere i 80'erne passer bedre med det opfattede trykgruppemønster. Altså, ville man hos ældre talere i 80'erne måle en udelukkende stigende F0 i den betonede stavelse? Et andet interessant spørgsmål, der melder sig, er, om det initiale fald, som også Grønnum måler, men tilskriver intrinsiske forhold, er blevet tydeligere med tiden? Hvis der er tale om en sprogforandring i retning af det københavnske, kunne denne lille ændring konventionaliseres af talerne og dermed fremhæves. Det ville betyde, at yngre talere på et senere tidspunkt skulle have et endnu tydeligere initialt fald i den betonede stavelse. Derudover kan man have en hypotese om, at det finale fald er mindre stejlt blandt yngre talere sammenlignet med ældre talere, hvis deres trykgruppemønster bevæger sig i retning af det københavnske.

I vores undersøgelse vil vi prøve at finde ud af, om der er sket en lydforandring i Næstved fra slutningen af 80'erne til 00'erne, mere specifikt om trykgruppemønstret har ændret sig. Det undersøger vi ved hjælp af LANCHART-korpussets optagelser fra Næstved. I de følgende afsnit vil vi beskrive de data, vi har brugt. Dernæst beskriver vi proceduren for analysen af data efterfulgt af en gennemgang af resultaterne. Til sidst runder vi af med en diskussion af vores fund og deres implikationer for teorier om sprogforandring.

3 DATAGRUNDLAG: INFORMANTER OG MATERIALE

For at undersøge om sprogbrugere fra Næstved er begyndt at bruge det københavnske trykgruppemønster, har vi benyttet LANCHART-korpussets (Gregersen m.fl. 2014) optagelser, som vi beskrev i afsnit 1. I det følgende forklarer vi, hvordan vi udvalgte informanterne fra korpusset, og vi giver en oversigt over deres demografiske karakteristika.

3.1 Udvalgelse af informanter

Enkelte informanter blev fravalgt, f.eks. hvis de blev vurderet til ikke at være opvokset i Næstved på trods af deres klassificering i korpusset. Et eksempel på dette er en mand fra generation 3, som havde boet i Odense, indtil han var 12, men som derefter havde vekslet mellem at være hos sin mor og far i Odense og Næstved. Derudover blev enkelte talere fravalgt, fordi de kun optrådte i optagelsen fra 80'erne eller i den fra 00'erne, dvs. fordi der var for lidt materiale med dem.

Det forholdt sig sådan, at talerne fra generation 2 i Næstved ikke var kodet for socialklasse, så den har vi selv foretaget. Vi brugte kriterierne i Monka (2013: 75), navnlig om informanten havde over eller under tre et halvt års uddannelse, og om de havde ledelsesansvar. Vi lyttede starten af hvert interview igennem og forsøgte at inddele informanterne på baggrund af de informationer, de gav i interviewet. Typisk kunne de kategoriseres forholdsvis entydigt på baggrund af deres uddannelse.

I tabel 1–3 ses oversigter over, hvordan de udvalgte talere fordeler sig efter køn, socialklasse (arbejderklasse, AK; middelklasse, MK) og optageårsti. Derudover har vi noteret talernes fødeår og generationens aldersbredde. De to kolonner yderst til højre opsummerer antallet af talere og antallet af udvalgte trykgrupper for hver kombination af demografiske variable. Antallet af trykgrupper er det totale antal observationer, vi endte med, efter at have fraserteret fejlmålinger og ekstreme værdier (± 2 standardafvigelse fra den gennemsnitlige F0 målt i Hz for hver informant). Antallet af observationer før fraserteringen er angivet i parentes. I alt blev 224 trykgrupper fraserteret, så det totale antal endte på 2046 observationer.

TABEL 1. OVERSIGT OVER TALERE FRA GENERATION 1

Køn	Socialklasse	Optagelse	Fødeår	Alder	Talere	Trykgrupper
Kvinder	AK	80'erne	1943–1955	34–46	4	51 (53)
Kvinder	AK	00'erne	1943–1960	46–63	5	120 (128)
Kvinder	MK	80'erne	1945–1959	30–44	5	89 (94)
Kvinder	MK	00'erne	1945–1959	46–61	7	102 (114)
Mænd	AK	80'erne	1948–1962	27–41	3	49 (52)
Mænd	AK	00'erne	1946–1962	43–60	5	128 (137)
Mænd	MK	80'erne	1942–1959	30–47	3	79 (86)
Mænd	MK	00'erne	1942–1959	46–63	6	237 (276)

TABEL 2. OVERSIGT OVER TALERE FRA GENERATION 2

Køn	Socialklasse	Optagelse	Fødeår	Alder	Talere	Trykgrupper
Kvinder	AK	80'erne	1968–1969	17–18	2	15 (16)
Kvinder	AK	00'erne	1968–1969	37–39	2	76 (80)
Kvinder	MK	80'erne	1964–1969	17–22	6	55 (62)
Kvinder	MK	00'erne	1964–1969	37–43	6	171 (190)
Mænd	AK	80'erne	1965–1968	18–21	2	37 (41)
Mænd	AK	00'erne	1965–1968	39–42	3	127 (154)
Mænd	MK	80'erne	1969	17–17	4	55 (58)
Mænd	MK	00'erne	1969	37–38	4	244 (273)

TABEL 3. OVERSIGT OVER TALERE FRA GENERATION 3

Køn	Socialklasse	Optagelse	Fødeår	Alder	Talere	Trykgrupper
Kvinder	AK	00'erne	1990–1991	16–17	5	51 (54)
Kvinder	MK	00'erne	1991	16–16	11	124 (141)
Mænd	AK	00'erne	1989–1991	16–18	6	109 (123)
Mænd	MK	00'erne	1990–1991	16–17	7	128 (138)

3.2 Udvalgelse af lydklip og annoteringer i Praat

Før vi beskriver udvælgelsen af lydklip, vil vi kort redegøre for annoteringerne af lydfilerne i LANCHART-korpusset. For at forklare disse må vi først skrive lidt om programmet Praat (Boersma & Weenink 2023), der er blevet brugt til at annotere lydfilerne.

Praat er et signalbehandlingsprogram, der er lavet specifikt til akustisk analyse af tale. Derfor kan man i dette program indsætte tekstinformationer ved bestemte tidskoder, som svarer til akustiske begivenhe-

der i lydfilen. Et *tier* i Praat er en række med tekstinformation. Disse kan 'stables' oven på hinanden, så man kan have flere niveauer af annotering, som helt eller delvist overlapper i tid (se Tøndering 2010b for en introduktion til Praat). Man kan f.eks. have et tier til sætninger over et tier, der inddeler lydfilen i mindre bidder som ord, stavelser, el.lign. (se figur 3 for et eksempel).

LANCHART-korpusset er annoteret for mange forskellige sociale, lingvistiske og interaktionelle faktorer (Gregersen m.fl. 2014, Gregersen 2009). Alle disse informationer er kodet i separate tiers i Praat. Foruden en ortografisk transskription af informantens og interviewerens tale er alle interviews kodet for det, der er blevet kaldt interaktionsstruktur. Eftersom vi ønskede at udvælge lydclip, hvor informanten talte mest muligt, prøvede vi i første omgang kun at ekstrahere alle lydclip, der er kodet som monologer ("I8" i annoteringen) i interaktionsstruktur. Monologer defineres i LANCHARTs kodningsmanual bl.a. ved, at "taleren selv bestemmer, hvad der tales om gennem en længere passage" (LANCHART 2010: 32). Det er ønskeligt for vores undersøgelse, da det kan forventes at være mere naturlig tale end i situationer, hvor informanten f.eks. besvarer spørgsmål af biografisk karakter. I disse situationer kan man håbe på, at taleren er mindre opmærksom på måden, de taler på, da interviewsituationen kunne tænkes at få dem til følge andre normer for udtale, end de ellers ville have gjort (Labov 1972). Denne taktik viste sig dog ikke at fungere lige godt i alle tilfælde. Nogle talere producerede meget få monologer (eller også er monologerne bare ikke blevet kodet systematisk i hele korpusset), mens andre interviews slet ikke er kodet for interaktionsstruktur, og de indeholder dermed ingen opmærkede monologer. I disse tilfælde brugte vi det materiale, der var til rådighed.

3.3 Bearbejdning af udvalgte lydclip

De udvalgte lydclip blev automatisk fonetisk annoteret ved hjælp af *forced aligner*'en Autophon (Young & Anikwe 2024). Dette værktøj benytter neurale netværk til at koble en lydfil til en ortografisk transskription. Samtidig inddeler den lydfilen i passende grænser og indsætter en fonetisk transskription baseret på en 'oversættelse' fra ortografi til lydskrift. Disse annotationer blev dels brugt til at få et hurtigt overblik over seg-

menterne i ordene, dels for at muliggøre fremtidige segmentalfonetiske undersøgelser. Denne segmentering er ikke i fuldstændig overensstemmelse med manuel segmentering, men overensstemmelsen er inden for 20 millisekunders nøjagtighed for 72 procent af alle segmentgrænser sammenlignet med en manuel segmentering, hvilket er samme standard, som kendes fra automatisk segmentering af f.eks. engelsk (Young 2024). De forekomster, der udgør vores datamateriale, har været igennem en manuel korrigerende af segmentgrænserne (se mere nedenfor).

Alle optagelser, nye såvel som gamle, er optaget i to spor med to mikrofoner – én til informanten og én til interviewer. Det udstyr, man brugte i slutningen af 80'erne til de gamle optagelser, er dog af en betydeligt ringere kvalitet end dem, der blev brugt til de nyeste optagelser i 2005–2007. Det har en række konsekvenser for vores undersøgelse: For det første har aligeringen meget svært ved at annotere de gamle lydoptagelser. Det skyldes mængden af baggrundsstøj, og at mikrofonerne nærmest ikke kan adskille de to talere: Det er næsten ligegyldigt, om man lytter til det ene eller det andet spor, for begge talere (og alt, der sker i baggrunden) kan høres på begge spor.

Det andet, og væsentlig større, problem med de gamle optagelser er, at Praats algoritme har meget svært ved at udregne F0, når der er for meget støj, fordi den i så fald ikke opfanger periodiciteten. I mange tilfælde kunne vi derfor ikke bruge ellers gode trykgrupper, fordi Praat ikke kunne finde grundtonefrekvensen. Det resulterer i, at vi har væsentlig færre forekomster fra de gamle optagelser sammenlignet med de nye (jf. tabel 1–3). Desuden er der større risiko for, at Praats algoritme har regnet forkert i disse optagelser. Det er dog det bedste, vi har at arbejde med, så vi må acceptere, at disse optagelser er relativt støjfulde.

4 UDVÆLGELSE AF TRYKGRUPPER

4.1 *Hvor starter og slutter en trykgruppe?*

En trykgruppes grundtoneforløb begynder ved den betonede vokal og fortsætter frem til næste betonede vokal (Grønnum 2005: 339). Med andre ord hører eventuelle konsonanter i ansats, uanset stemthed, strengt taget til foregående trykgruppe, f.eks. /st/ i /'stæ:ʝəlsə/ *stavelse*. Hvis vi tager en frase som *den tunge stavelse*, så inddeles den i trykgrupper på

følgende vis: *den t{funge st}{avelse}*. Her indrammes trykgrupper vha. tuborgklammer. *Den* og *t* (dvs. *t*'et i *tunge*) hører formelt set ikke til nogen trykgruppe, da de står initialt i intonationsfrasen uden at starte med en betonet stavelse. De to trykgrupper, der følger, starter med den betonedede vokal, så de første to segmenter i *stavelse* hører til samme trykgruppe som /'ɔŋə/ i *tunge*.

Vi valgte dog af flere årsager at definere trykgruppens afslutning som den sidste stavelses *offset*, dvs. uden at medregne eventuelle ansatskonsonanter fra den følgende betonedede stavelse. For det første er der en uoverensstemmelse mellem det *prosodiske ord* i Basbølls fonologiske model (2005: 539) og Grønnums intonationsmodel ift. ordinitiale ansatskonsonanter. Vies (2021) viser, at prosodisk ordinitialt [s] akustisk opfører sig anderledes, når det står før en trykgruppeinitial vokal (dvs. når det står initialt i den betonedede stavelse) sammenlignet med andre positioner. Selvom F0-konturen begynder ved den betonedede vokal, er det altså vanskeligt at afgøre, om prosodisk ordinitiale ansatskonsonanter hører med i samme trykgruppe som efterfølgende vokal eller den foregående trykgruppe: Det kommer an på, om man vil tilskrive andre faktorer end F0 en rolle i trykgruppestruktur. Pga. denne forskel, som afhænger af niveauet i den prosodiske struktur, udelod vi disse konsonanter. Vi er primært interesserede i starten af trykgruppen, for det er her, tidligere undersøgelser finder de største forskelle mellem næstvedsk og københavnsk, så om vi medregner efterfølgende ansatskonsonanter til grundtoneforløbet for foregående stavelser eller ej, har det begrænset påvirkning på vores resultater. Desuden bidrager ustemte ansatskonsonanter ikke til F0-konturen. Som nævnt blev alle ansatskonsonanter udeladt af analysen, og de har således ikke indflydelse på vores målinger.

4.2 Mikroprosodi

Et af de vanskelige aspekter ved at arbejde med grundtoneforløb er, at det påvirkes af rigtig mange faktorer, som kan være svære at adskille fra hinanden. F.eks. kan talerens humør have en indvirkning på størrelsen på udsvingene i F0 (Pichora-Fuller m.fl. 2016). En anden og betydelig mindre påvirkning er mikroprosodi. F.eks. vil F0 være højere efter ustemte konsonanter sammenlignet med stemte konsonanter, og F0 sænkes særligt før stemte konsonanter. Derudover har vokalers højde en

påvirkning på F0: Høje vokaler fører til højere F0 sammenlignet med lave vokaler ((Grønnum) Thorsen 1979, Gussenhoven 2004: 7–9). Af denne grund er trykgruppemønsteret i *isfugl* /'isfu:ʔl/ med to høje vokaler strengt taget ikke direkte sammenligneligt med trykgruppemønsteret i *rovfugl* /'røʁfu:ʔl/, for sidstnævntes første stavelse vil alt andet lige starte med en lavere F0-værdi end førstnævntes på grund af forskellene i vokalhøjde. Hvis vi ønsker at undersøge, om Taler nr. 1 fra Næstved taler mere som en københavnervner end den yngre Taler nr. 2, nytter det derfor ikke noget at sammenligne den ene talers udtale af *isfugl* med den anden talers udtale af *rovfugl* – for F0-stigningen fra *rov-* til *fugl* vil være større end fra *is-* til *fugl*. Derfor kunne man fejlagtigt komme til at konkludere, at Taler 2 taler mere københavnsk end Taler 1, alene på grund af forskellen i F0 fra første til anden stavelse, selvom den i virkeligheden skyldes de iboende F0-forskelle i vokalhøjde i de ord, vi sammenligner.

Derfor ledte vi specifikt efter trykgrupper, hvori alle vokaler havde nogenlunde samme højde gennem hele trykgruppen. Fordi både vokalhøjde og konsonanters stemthed og lukkelydes aspiration påvirker F0, kodede vi hver trykgruppe med den information. F.eks. blev *abe* kodet som 'æ ustemt ə'. Stavelsesbærende sonoranter blev kodet som sådanne, dvs. med angivelse af sonoranten, f.eks. *aben*: 'æ ustemt m'. Senere i analysen blev stavelseskerne inddelt efter højde. Vi valgte i første omgang ikke at kode sonorante konsonanter for højde. Med andre ord kodede vi kun vokalhøjde, og sonorante konsonanter fungerede derfor som en slags jokere. Vores tilgang er opsummeret i tabel 4.

TABEL 4. KODNING AF HØJDE I STAVELSESKERNER

Højde	Segment
Høj	{i, ɪ, y, e, u, ʊ}
Mellem	{ɛ, ø, œ, ə, o}
Lav	{æ, ɐ, a, ɐ, ʌ, ɔ, ɒ}
Sonorant konsonant	{m, n, ŋ, l, ʃ}

1 Man kunne argumentere for, at /ð/ skulle være klassificeret som en vokal. F.eks. transskriberer Schachtenhaufen (2022) lyden som /ɜ/. Vi har valgt at følge Grønnums (2005) konvention og klassificere /ð/ som en sonorant, fordi det gør det lettere at sammenligne med de laboratoriestudier, der er lavet om intonation. Desuden antyder igangværende undersøgelser af /ð/, at lyden nok ligner en vokal akustisk, men ikke i artikulationen (Rasmus Puggaard-Rode, personlig kommunikation).

4.3 Anden relevant information om trykgrupperne

Fordi intonation spiller en væsentlig rolle i markeringen af illokution (Grønnum 2005: 346–348), f.eks. om en sætning fungerer som deklarativ eller interrogativ, valgte vi ikke at inkludere trykgrupper, der indgik i interrogative ytringer, og vi brugte altså kun deklarativer.

Derudover noterede vi trykgruppens placering i intonationsfrasen og antallet af stavelser, der indgik i trykgruppen, og vi tilføjede kommentarer om auditivt opfattet emfase (dvs. om trykgruppen var særligt fremhævet ved brug af fonetisk prominens). Disse faktorer forklarer vi nærmere nedenfor.

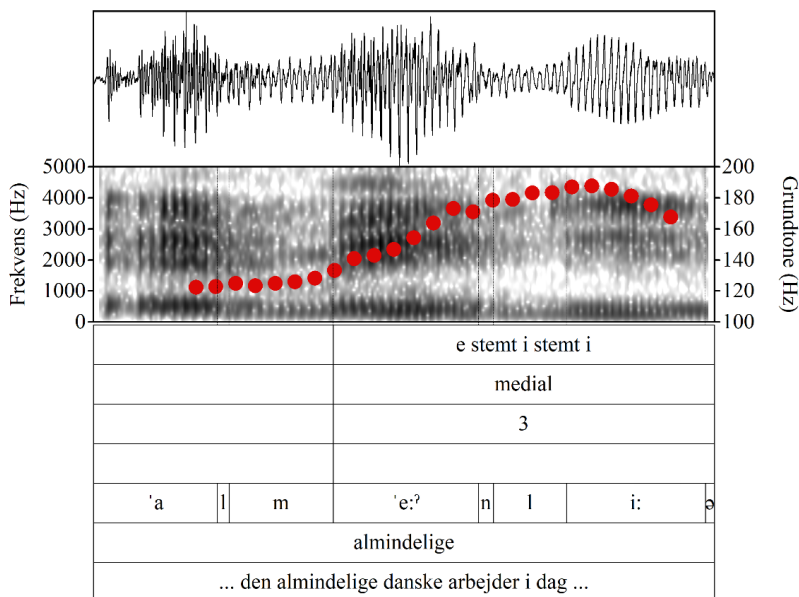
4.3.1 Position og antal stavelser

Endnu et aspekt, der vanskeliggør sammenligninger mellem trykgrupper, er deres position og antal stavelser. Grønnum (2005: 343–348) beskriver, at den globale intonationskontur generelt er faldende i deklarativer, hvilket betyder, at trykgrupper først i intonationsfrasen generelt vil have højere F0 end trykgrupper sidst i intonationsfrasen. Det gælder også for spontan tale: Der er en mindre tendens til, at længere intonationsfraser starter højere, og der er en stærk tendens til, at længere intonationsfraser ender lavere (Tøndering 2008: 108). Eftersom grundtoneintervallet, dvs. faldet i F0 fra en betonet stavelse til den næste, er nogenlunde konstant, vil kortere trykgrupper have et relativt stejlere fald end længere trykgrupper i hvert fald i oplæst tale (Grønnum 2005: 343–344). Det er, så vidt vi ved, ikke undersøgt, om det forholder sig sådan i spontan tale, men Tøndering (2008: 108) viser, at antallet af trykgrupper i en ytring ikke påvirker hældningen, så det kan tænkes, at en lignende (mangel på) sammenhæng findes for antallet af stavelser i en trykgruppe i spontan tale. ((Grønnum) Thorsen 1984 har vist, at forløbet gennem de posttoniske stavelser kun i beskedent omfang påvirkes i trykgrupper på seks stavelser eller mere i oplæst tale, og tendensen ser ud til at være talerspecifik.) I et forsøg på at tage højde for disse potentielt forstyrrende faktorer noterede vi derfor antallet af stavelser i hver trykgruppe, og hvor i intonationsfrasen de befandt sig.

Antal stavelser er ikke altid så ligetil at bestemme i løbende tale, idet ord kan reduceres fonetisk, så nogle stavelser falder bort eller bliver meget korte med centraliserede vokaler. I denne undersøgelse har vi dog

talt ud fra den fonetiske realisering, så hvis *stavelser* blev realiseret som [ˈstæʊsə] i stedet for [ˈstæ:ʊlsə], har vi talt ordet som tostavet i stedet for trestavet. Afgrænsningen af intonationsfraser blev vurderet auditivt og ved visuel inspektion af oscillogrammet og spektrogrammet. Det blev gjort ud fra flere kriterier: Mellem to intonationsfraser finder man nogle gange et såkaldt *reset*, dvs. et spring op i F0, så den første trykgruppe i intonationsfrasen realiseres med højere F0 end foregående trykgruppe. I disse tilfælde blev første trykgruppe efter et reset kodet som 'initial', og den sidste trykgruppe før et reset blev kodet som 'final'. Alt mellem disse blev kodet som 'medial'. I andre tilfælde kunne intonationsfraser afgrænses ved pauser af stilhed mellem dem. Indimellem virkede det også til, at grænser blev markeret med knirkestemme. Ingen af disse kriterier bruges konsekvent af sprogbrugerne (men pauser er ret frekvente), og vi anerkender, at den mediale gruppe godt kunne være mere finkornet. Men da der ikke findes særlig meget litteratur på området, besluttede vi, at det var bedst at holde os til disse rimelig håndgribelige kriterier. Desuden ville det gøre den statistiske analyse mere kompliceret, hvis vi opererede med mange flere niveauer for ytrings- og/eller intonationsfrasepositioner. I figur 3 ses et eksempel på en udvalgt trykgruppe med tilhørende tekstinformation. Klippet er taget fra den gamle optagelse (1986) med en mandlig taler (PRB) fra generation 2.

FIGUR 3. EKSEMPEL PÅ EN TRYKGRUPPE FRA DEN MANDLIGE TALER PRB (GENERATION 2, GAMMEL OPTAGELSE)



Øverst ses et oscillogram. I midten ses spektrogrammet med tilhørende frekvensakse i venstre side. Oveni spektrogrammet er grundtoneforløbet (F0) markeret med røde prikker. Frekvensaksen i højre side hører til F0-målet. Nederst er textgriddet indtegnet. I det første tier er vokalkvalitet og konsonanternes stemthed noteret. Dernæst kommer information om trykgruppens position i intonationsfrasen og antallet af stavelser. Det her tomme felt bruges til kommentarer. Næstnederst er der et tier med lydskrift efterfulgt af et tier med ortografisk transskription på ordniveau, og nederste tier viser ortografi på ytringsniveau (her er det beskåret).

4.3.2 *Emfase*

Grønnum (1980) viser, at trykgrupper med emphatisk tryk følger det samme mønster som ikke-emfatiske trykgrupper, men at F0-udsvinget er større for emfatiske trykgrupper. Det er især iøjnefaldende, at den posttoniske tone er markant højere ved emfase, og faldet gennem de efterfølgende ubetonede stavelser er dybere. Eftersom emfase påvirker trykgruppen, noterede vi alle trykgrupper, som vi vurderede som havende ekstra stærkt tryk, så vores statistiske model kunne tage højde for den variation. Emfase blev noteret som kommentar i textgriddet, dvs. tier nummer fire i figur 3.

5 VIDERE DATABASEHANDLING OG VALG AF STATISTISKE VÆRKTØJER

Ved hjælp af Praat-scripts (Pharao 2024a, Pharao 2024b, Diderichsen 2023) ekstraherede vi alle de intervaller, der indeholdt udvalgte trykgrupper, sammen med de manuelt kodede informationer (vokalkvalitet og konsonanters stemthed, antal stavelser, position i intonationsfrasen, emfase og eventuelle andre kommentarer). Lydsiden af disse klip blev automatisk inddelt i 30 ækvidistante punkter, og i hvert punkt blev F0 målt. Med andre ord målte vi F0 ved 1/30, 2/30, 3/30 ... frem til 30/30 af trykgruppens totale varighed. Hvis en trykgruppe havde en varighed på 300 ms, blev F0 altså målt ved $300/30 * 1$ (dvs. 10 ms), $300/30 * 2$ (dvs. 20 ms), osv., så vi havde 30 målinger udført med samme tidsinterval imellem hver. Dette blev gjort, så vi ikke behøvede at tage højde for forskelle i trykgruppernes varighed. Således blev tid normaliseret, da den absolutte afstand fra punkt 1 til punkt 2 ikke nødvendigvis er den samme for to trykgrupper, men den relative afstand i forhold til de givne trykgruppers overordnede varighed er den samme. Vi bad Praat om kun at lede efter F0 i intervallet 50–300 Hz for mænd og 100–500 Hz for kvinder, fordi Praats manual anbefaler, at man afgrænser det interval, hvor algoritmen skal forvente at finde F0, da man ellers i højere grad risikerer fejlmålinger.

Alle disse data blev derpå importeret i R (R Core Team 2022), som er et statistisk databehandlingsprogram. Her tilføjede vi demografiske oplysninger om talerne, herunder deres socialklasse, generation, alder og køn. Derudover fjernede vi for hvert af de 30 målepunkter alle datapunkter på mere end to standardafvigelser fra gennemsnittet.

Til de statistiske analyser benyttede vi en forholdsvist ny type statistisk model kaldet *generalized additive mixed models* (GAMMs). Da disse, så vidt vi ved, endnu ikke er blevet beskrevet i sammenhæng med dansk lingvistik, redegør vi i næste afsnit kort for, hvordan disse modeller fungerer, og hvorfor vi har valgt at bruge dem frem for mere hyppigt anvendte og bedre beskrevne metoder som f.eks. lineær regression eller variansanalyse.

5.1 GAMMs – *Generalized additive mixed models*

Mange træk ved GAMMs er velkendte fra statistisk analyse af sproglig variation og forandring, da GAMMs er en form for multivariat regres-

sionsanalyse med blandede effekter (Wood 2017). Det er en analysetype, der gør det muligt at undersøge forklaringskraften i et sæt af forhold, der kendes for data, de såkaldte faktorer eller prædiktorer, som antages at have faste effekter, og samtidig tage højde for, at der inden for de faktorer, der testes i modellen, er grupperinger af data, som struktureres af andre forhold, nemlig de tilfældige effekter. Kombinationen af faste og tilfældige effekter er det, der gør denne type til modeller med blandede effekter. Med data, hvor man ønsker at undersøge (eller kontrollere for) faktorer som køn, alder og socialklasse, og hvor man ved, at flere datapunkter kommer fra den samme taler, har man brug for at kunne bruge modeller med blandede effekter. Faktorerne er de faste effekter, hvis indflydelse på mønstrene i data, man ønsker at få estimeret og vurderet mht. signifikans. Man kan f.eks. være interesseret i, om personer født i 60'erne har en statistisk signifikant anderledes udtale af /a/ sammenlignet med personer født i 80'erne. Hvis man alene undersøger fødselsåret som en fast effekt og altså tager alle sine datapunkter fra dem, der er født i 60'erne, og samler dem i samme kategori og sammenligner dem med alle datapunkter fra dem, der er født i 80'erne, uden at tage højde for, at der indgår flere forskellige talere i hver gruppe, og at disse talere har produceret mere end ét datapunkt i hver bunke, så kan man komme til at tilskrive fødselsåret en effekt, som alene bæres af enkelte personer i hver gruppe. En model med blandede effekter tager højde for variationen i tilfældige effekter, dvs. at talerne fra f.eks. 60'erne behandles som en gruppe, hvis punkter ikke er uafhængige af hinanden. Modellen tager altså hensyn til, at nogle talere i gruppen evt. opfører sig forskelligt fra gruppens gennemsnit. Datasættets sammenhæng gør således, at det er fordelagtigt og, ville vi mene, nødvendigt at bruge modeller med blandede effekter, fordi det herved bliver muligt at tage højde for, at der er flere forhold i data, der bidrager til datasættets struktur. Nogle af forholdene er vi interesserede i at undersøge nærmere, og derfor inkluderes de som faste effekter. Andre forhold inddeler på tilfældig vis data i mindre undergrupperinger, og i undersøgelser af sproglig variation er det typisk de undersøgte talere og ord, som vi ikke nødvendigvis ønsker at undersøge, men vi vil gerne kunne tage højde for denne komponent i strukturen.

De data, vi skal behandle, består af tidsserier, idet de er målinger af grundtonen i løbet af ytringer. Vi skal derfor kunne estimere detaljerede

og forventeligt komplekse forløb over tid. Netop fordi vi på baggrund af eksisterende undersøgelser af trykgruppemønstret forventer, at forløbene vil være komplekse, har vi brug for et værktøj som GAMMs, da det gør det muligt at estimere ikke-lineære forløb. Hvor andre regressionsmodeller, der anvendes på kontinuerte målinger, antager et lineært forhold mellem prædiktorer og det, man har målt på, giver GAMMs mulighed for at undersøge, om der er mere komplekse relationer mellem en prædikator og ens målinger. Det er muligt, fordi metoden inddrager såkaldte *smooth terms* i modellen, dvs. ud over at kunne estimere skæringspunkt på y-aksen og hældning på regressionslinjen, kan GAMMs også estimere, hvor godt målingerne passer med en række såkaldte *basisfunktioner*, som er forskellige kombinationer af polynomier, der indsættes med samme interval langs x-aksen. Det er også muligt at estimere ikke-lineære sammenhænge med andre regressionsmodeller, men de kræver, at man specificerer graden af kompleksitet på forhånd. I sådanne tilfælde er det ikke nødvendigt med GAMMs, fordi den fornødne kompleksitet beregnes i forbindelse med estimeringen og således er baseret på data. Desuden begrænses kompleksiteten, dvs. graden af nonlinearitet, vha. en *smoothing-parameter*, som modellen automatisk udleder af data. På denne måde mindskes risikoen for *overfitting*. Metoden er således ideel til undersøgelser af grundtoneforløb, fordi vi har meget god grund til at regne med, at de består af stigninger og fald, og fordi vi gerne vil være sikre på, at estimeringerne ikke skævvrides af enkelte talere, altså af samme grund som regression med blandede effekter i det hele taget er udbredt i undersøgelser af sproglig variation og forandring. GAMMs gør det således muligt ikke bare at tage højde for de enkelte taleres bidrag til de parametriske dele af en model, dvs. de prædiktorer, hvis faste effekter skal undersøges, men også at tillade tilfældige smooths, dvs. at modellere konturerne for hver enkelt taler og tage højde for dem, når konturen for en bestemt gruppe skal estimeres. Med andre ord tager modellen højde for, at basisfunktionerne vægtes forskelligt for forskellige talere. Denne sidste mulighed gør modellerne meget kraftfulde, idet der tages højde for tilfældig variation. Samtidig er det væsentligt at holde sig for øje, at det ikke er helt ligetil at afgøre, hvordan disse tilfældige smooth-termer skal inkorporeres i modellen.

I konstruktionen af en GAMM er det nødvendigt både at medtage de parametriske termer, man kender fra lineære regressionsmodeller, og

de omtalte smooths. Disse smooths er baseret på de tidsserier, der er i data, i vores tilfælde målingerne af F_0 i de 30 ækvidistante punkter i hver trykgruppe, som funktion af en prædiktør, man gerne vil undersøge, f.eks. optagetidspunktet, eller gerne vil kontrollere for i modellen, f.eks. antallet af stavelser i trykgruppen. Hertil kommer som sagt inklusion af faktorer med tilfældig variation, her de enkelte talere, da de bidrager med flere forekomster hver især.

Hvor hypotesedreven brug af lineære og logistiske regressionsmodeller giver mulighed for at vurdere en prædiktors væsentlighed alene ved at se på dens egen p -værdi, når den tilføjes en model, der indeholder alle kendte kontrolfaktorer, så er denne praksis ikke tilstrækkelig, når man skal arbejde med smooths. Det skyldes, at de p -værdier, man kan aflæse for smooths, alene fortæller, om effekten af den givne faktor signifikant forskellig fra en rent lineær modellering af faktoren. Særligt med tidsseriedata med mange målepunkter for hver enkelt forekomst som i vores tilfælde her, er muligheden for at finde et ikke-lineært forhold ganske høj. Derfor er det nødvendigt at vurdere tilføjjelsen af smooths gennem modelsammenligning. Én af måderne, hvorpå man kan afgøre, om en given faktor har en effekt på den variabel, man er interesseret i at modellere, er ved at sammenligne en simpel model med en mere kompleks model. I disse sammenligninger undersøger man bl.a., om den øgede kompleksitet i den ene model kan retfærdiggøres, ved at dens forudsigelser passer bedre til de observerede data. Her benytter man som regel Akaikes informationskriterium (AIC; Cavanaugh & Neath 2019). Dette bruges til at finde et kompromis, hvor man bedst muligt undgår at lave en model, der ikke kan generaliseres til nye data (*overfitting*), eller en model, der er for simpel til at kunne forklare mønstre i data (*underfitting*). Sammenligningerne laves i R med funktionen *compareML()*, der er en form for variansanalyse, fra pakken *itsadug* (van Rij m.fl. 2022).

Vi vil i næste afsnit forsøge at give læseren en konceptuel forståelse af, hvordan disse modeller fungerer, uden at gå i dybden med den bagvedliggende matematik. Interesserede læsere henvises til Wood (2017) for en dybdegående indføring i GAMMs. Ovenstående korte beskrivelse af, hvordan man arbejder med at konstruere GAMMs, bygger på Sóskuthy (2017, 2021), Winter & Wieling (2016) og Wieling (2018), der alle er solide introduktioner til anvendelsen af GAMMs brugt på lingvistiske data og velegnede for læsere, der ønsker at gå videre med metoden selv.

5.2 Fra basisfunktioner til sprællende estimater

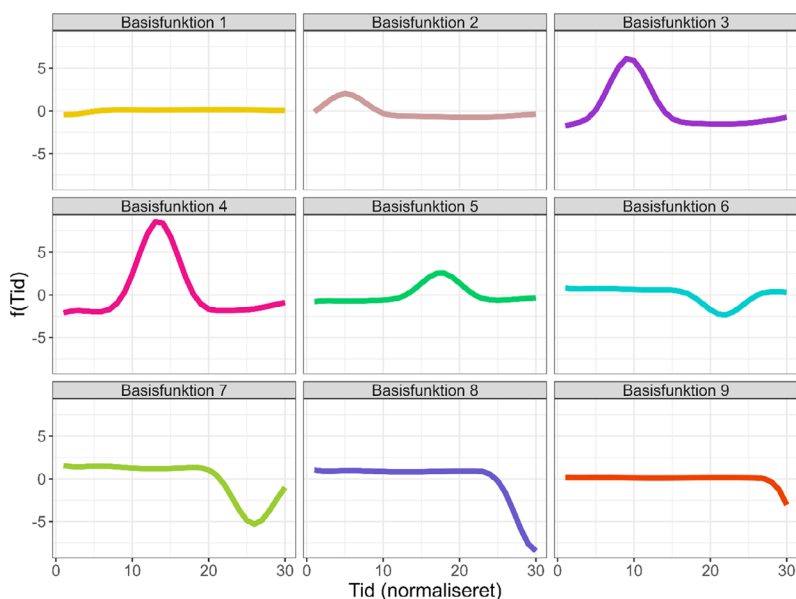
For at kunne forklare, hvordan GAM(M)s fungerer, vil vi beskrive modelleringen af en enkelt talers estimerede grundtoneforløb i trykgrupperne i ét interview. Vi benytter altså en simpel GAM (med ét M, dvs. uden variable med tilfældig variation), da vi modellerer F0-forløbet for én taler. Her har vi arbitrært valgt at kigge på taleren MBH fra generation 2 i de nyeste optagelser fra 2007. I denne omgang inkluderer vi ingen af de faktorer, vi forventer påvirker grundtoneforløbet i modellen. Det undlader vi dels for at holde eksemplet så simpelt som muligt, dels fordi vi ikke ønsker at teste vores hypoteser på en enkelt taler. Den simple model kan beskrives ganske enkelt: Hvilken kombination af ikke-lineære funktioner beskriver bedst ændringen i F0 over tid for taleren MBH? Her skal tid forstås som 'i løbet af trykgruppen', dvs. i de 30 målepunkter. I statistikprogrammet R formaliseres det ved hjælp af *mgcv*-pakken (Wood 2011, 2017) som $F0 \sim s(\text{Tid})$.

Graden af udjævning (*smoothing*) af funktionens linje afgøres af såkaldte knudepunkter eller konvergenspunkter, som bestemmer, hvor mange basisfunktioner, der maksimalt kan indgå i modellen. Basisfunktioner kan bl.a. udgøres af prædefinerede kombinationer af polynomier, der indsættes med lige store intervaller langs x-aksen. Det kunne f.eks. være 3.-gradspolynomier (dvs. $ax^3 + bx^2 + cx + d$). Disse basisfunktioner, som altså er matematiske støttefunktioner og ikke egenskaber ved grundtoneforløbet, ganges med en koefficient, dvs. den størrelse, som de bidrager med i modelleringen af grundtoneforløbet, og dernæst summeres de for samlet set at give et estimat af den overordnede udvikling over tid (målt i de 30 punkter). Hvis der f.eks. er ni basisfunktioner i modellen, vil hver af disse ganges med en vægtning, så den endelige funktion er den vægtede sum af alle basisfunktionerne, dvs. $\text{vægt}_a \times \text{funktion}_1 + \text{vægt}_b \times \text{funktion}_2 \dots + \text{vægt}_i \times \text{funktion}_9$. Det størst mulige antal basisfunktioner i en model er antallet af knudepunkter minus en. Den resterende basisfunktion modellerer skæringspunktet, dvs. den gennemsnitlige værdi over tid, og denne værdi er konstant, så den bidrager ikke til udjævningen af kurven. Der er altså en direkte relation mellem antallet af knudepunkter og antallet af basisfunktioner. Ved at øge antallet af knudepunkter tillader man modellen at benytte sig af flere basisfunktioner, hvilket i sidste ende gør det muligt at modellere flere nonlineære tendenser. Mo-

dellen benytter sig dog ikke nødvendigvis af flere basisfunktioner, fordi man øger antallet af knudepunkter. I vores simple model har vi brugt 10 knudepunkter, hvilket betyder, at der maksimalt kan indgå ni basisfunktioner til estimeringen af MBH's grundtoneforløb. Vær opmærksom på, at antallet af knudepunkter ikke direkte har noget at gøre med antallet af målepunkter (som altid var 30 for hver forekomst). Antallet af knudepunkter bestemmer, hvor detaljeret et non-lineært forløb der kan estimeres af modellen.

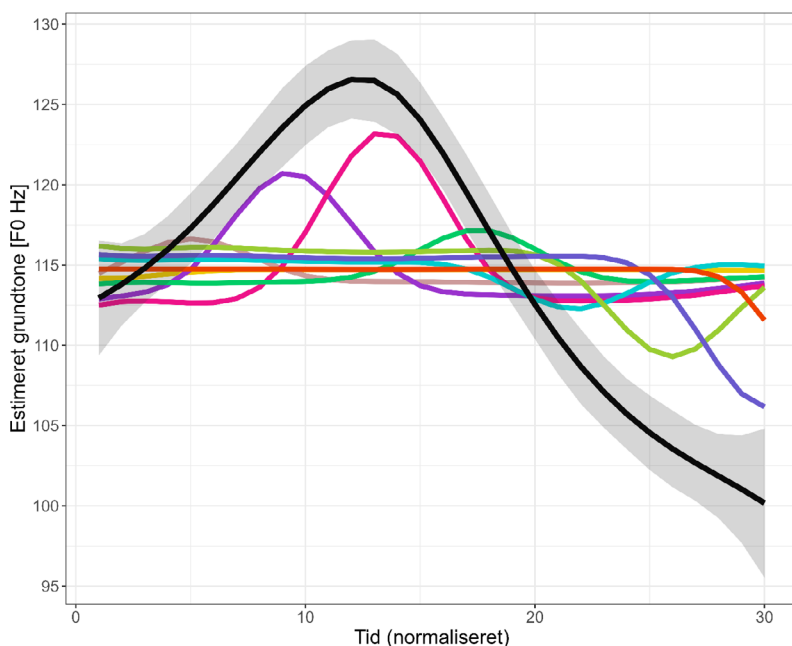
Ved hjælp af R-pakken *Gratia* (Simpson 2024) kan man ekstrahere de vægtede basisfunktioner fra modellens koefficienter. Dem har vi plottet i figur 4. Hver graf viser en af de ni basisfunktioner. X-aksen viser vores normalisering af tid, dvs. de 30 målepunkter, og y-aksen viser værdien $f(\text{Tid})$ for hver basisfunktion. Aksen er normaliseret, så modellens skæringspunkt er trukket fra, og værdierne er centreret omkring nul. Grafen viser således, hvordan basisfunktionerne bidrager til det samlede estimat af, hvordan forløbet afviger fra skæringspunktet som funktion af tid.

FIGUR 4. UNDERLIGGENDE BASISFUNKTIONER FOR ESTIMERINGEN AF MBH'S GRUNDTONEFORLØB



Tager man summen af disse basisfunktioner, får man som sagt det samlede estimat af grundtoneforløbets udvikling over tid. Det viser vi i figur 5, hvor vi har lagt basisfunktionerne sammen med skæringspunktet (den sidste basisfunktion, som ikke bidrager til udjævningen af grafen), så værdierne centrerer omkring grundtoneforløbets gennemsnitsværdi snarere end omkring nul. Den sorte linje viser estimatet, altså de summerede basisfunktioner (inklusive skæringspunktet) for hvert punkt i tidsserien, og det grå bånd viser konfidensintervallerne. Her er altså tale om et estimat af grundtoneforløbet baseret på mange, lidt forskellige, konkrete grundtoneforløb.

FIGUR 5. MBH'S ESTIMEREDE GRUNDTONEFORLØB OG UNDERLIGGENDE BASISFUNKTIONER



6 ZOOM IND PÅ DE ENKELTE TALERE

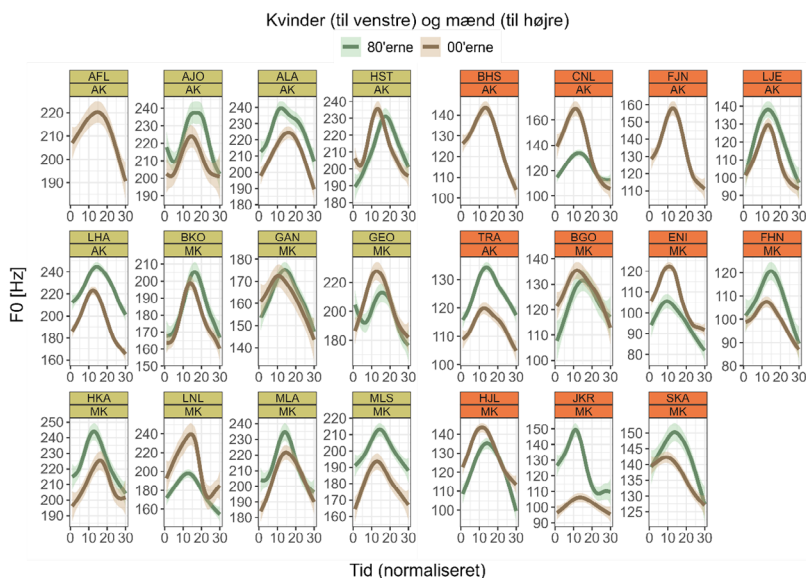
Før vi går videre med modelleringen af det fulde datasæt, vil vi først danne os et overblik over de enkelte talere. Det gør vi for at sikre os, at vores data ser ud til at opføre sig normalt, altså om de ligner forventelige trykgruppemønstre. Da vi for nogle grupper af talere har relativt få

observationer per taler, risikerer vi, at der ikke er nok data til at kunne sige noget meningsfuldt, hvilket kan mudre det overordnede billede. Derfor benytter vi visuel inspektion til at vurdere, om vi for nogle talere skal foretage yderligere ekskluderinger af fejlmålinger eller frasortere dem helt.

Til modelleringen af de enkelte taleres grundtoneforløb benyttede vi os af den samme simple form for GAM (med ét M, da den ikke er med blandede, *mixed*, effekter), som vi gennemgik i forrige afsnit. Dvs. at graferne i dette afsnit bygger på simple modeller af talernes individuelle grundtone som funktion af tid (vi viser ikke basisfunktionerne længere – de indgik blot i figur 5 for at forklare, hvordan modellerne fungerer). Altså tages der på nuværende tidspunkt ikke højde for andre faktorer, der kan tænkes at påvirke grundtonen. Da talerne ikke sammenlignes direkte her, viser vi grundtonen i Hz.

I figur 6 ses grundtoneforløbet for de enkelte talere i generation 1. De kvindelige talere er i venstre side, de mandlige talere er i højre side, og begge er arrangeret efter socialklasse. De grønne konturer viser talernes gennemsnitlige grundtoneforløb i 80'erne, hvor talerne var mellem 27 og 47 år, og den brune kontur viser grundtoneforløbet i 00'erne, hvor talerne var mellem 43 og 63 år. Nogle talere har kun én kontur, hvis de ikke blev optaget i begge årtier. Bemærk, at der her regnes på et gennemsnit over forekomster på både to, tre og fire stavelser. Det betyder, at timingen af målingerne relativt til vokalerne varierer i et vist omfang, og at der dermed introduceres variabilitet i målingerne. Til gengæld gør tidsnormaliseringen det, at vi kan se det generelle mønster uanset trykgruppens længde. Omfanget af variabiliteten kan ses ud fra konfidensintervallerne omkring konturerne, og i appendiks til artiklen kan man se, hvor mange målinger der ligger til grund for kurverne for hver deltager.

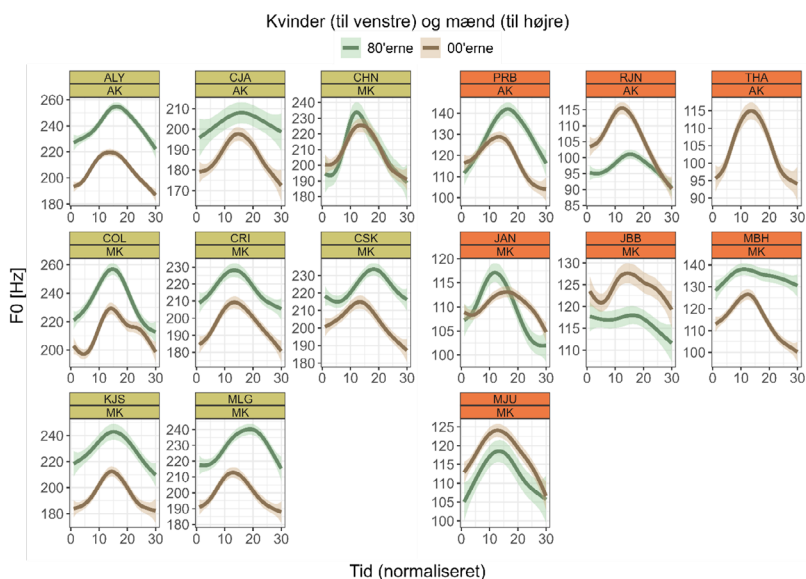
FIGUR 6. GRUNDTONEFORLØBET FOR HVER TALER I GENERATION 1



Over hver graf er talerens anonymiserede kode angivet. Under koden står AK for arbejderklasse og MK for middelklasse. Kvinderne er placeret i venstre side med kakifarvede kasser øverst, og mændene er placeret til højre med siennafarvede kasser øverst. Grundtonemønstre fra 80'erne er grønne, og mønstre fra 00'erne er brune.

Det er ret iøjnefaldende, at to kvinder (AJO og GEO) ser mere københavnske ud i 80'erne sammenlignet med 00'erne, idet de har et relativt markant fald ind i den betonedede stavelse. En tredje kvinde, HST, har til gengæld fået et initialt fald i 00'erne. Mændene i generation 1 ser derimod ud til at følge det næstvedske mønster konsekvent. Der er en vis tendens til, at kvindernes grundtone sænkes med alderen, ligesom flere mænd får en højere grundtone i de nye optagelser, men det er forventeligt (Torre & Barlow 2009) og ikke relevant for vores undersøgelse. Derudover er der ikke umiddelbart noget særligt bemærkelsesværdigt, så vi ser ingen grund til at bearbejde disse data yderligere. Nedenfor viser vi grundtoneforløbene for talerne fra generation 2 fra 80'erne, hvor de var 17–22 år, og fra 00'erne, hvor de var 37–43 år. Også her er der regnet på et gennemsnit over trykgrupper af forskellig længde, se ovenfor om generation 1.

FIGUR 7. GRUNDTONEFORLØBET FOR HVER INFORMANT I GENERATION 2

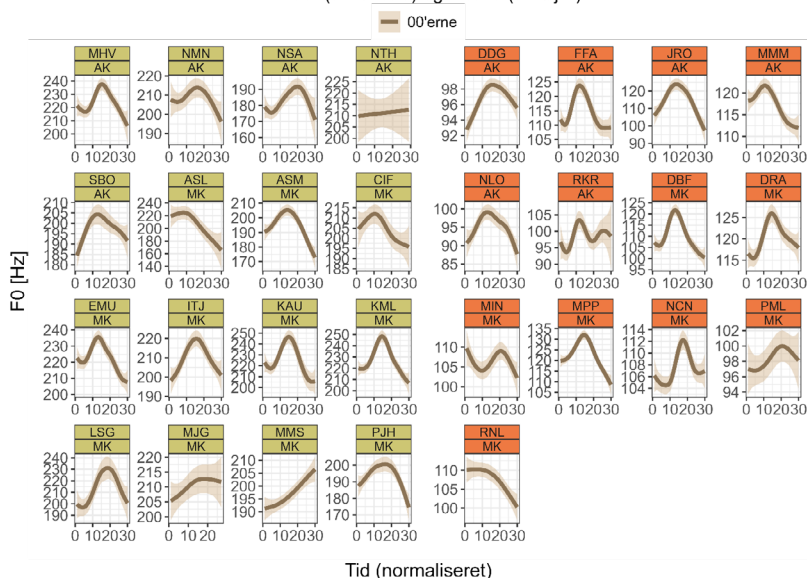


Over hver graf er talerens anonymiserede kode angivet. Under koden står AK for arbejderklasse og MK for middeklasse. Kvinderne er placeret i venstre side med kakifarvede kasser øverst, og mændene er placeret til højre med siennafarvede kasser øverst. Grundtonemønstre fra 80'erne er grønne, og mønstre fra 00'erne er brune.

Kvinderne ser generelt ud til at følge det næstvedske mønster, men CJA's og COL's mønstre bliver måske lidt mere københavnske i 00'erne. Som for generation 1 har vi et lidt underligt tilfælde med CSK og MLG, der ser mere københavnske ud i 80'erne end i 00'erne. Det samme kan man blandt mændene sige om RJN's mønster, mens JAN's og JBB's mønstre virker mere københavnske i 00'erne sammenlignet med 80'erne. Forskellene i real time er således for de enkelte talere ganske forskellige: Mens nogle ser ud til at have et fald ind i den trykstærke stavelse allerede i 1980'erne, som for enkelte aflæses af en ren stigning i 00'erne, er der andre som først viser en tendens til et fald ind i den betonedes stavelse i 00'erne. Ellers er der ikke noget usædvanligt, så disse data beholdes også, som de er. I figur 8 viser vi til sidst grundtoneforløbet for generation 3 under samme betingelser som for de to andre generationer, men her er der naturligvis kun en kurve pr. taler.

FIGUR 8. GRUNDTONEFORLØBET FOR HVER INFORMANT I GENERATION 3

Kvinder (til venstre) og mænd (til højre)



Over hver graf er talerens anonymiserede kode angivet. Under koden står AK for arbejderklasse og MK for middelklasse. Kvinderne er placeret i venstre side med kakifarvede kasser øverst, og mændene er placeret til højre med siennafarvede kasser øverst.

Generation 3, dvs. talere mellem 16 og 18 år, ser som de tidligere generationer lidt blandet ud. Blandt kvinderne ser EMU, KAU, LSG, MHV og NSA ud til at følge et københavnskliggende mønster, mens mønstret for resten ser næstvedsk ud. Blandt mændene følger DBF, DRA, FFA, MIN, NCN og måske RKR et mønster, der ligner det københavnske mest, mens mønstret for resten ser overvejende næstvedsk ud. Vi har dog to talere, NTH og MJG, hvis grundtonemønster næsten ser ud til at være lineære. Kombineret med de store konfidensintervaller og få forekomster (hhv. 11 og 7) tyder disse mønstre på fejlmålinger og/eller mangel på statistisk power. Vi valgte derfor at ekskludere disse to talere fra videre analyse. Taleren MMS har også et tilnærmelsesvist lineært forløb, men denne taler beholdt vi i analysen, da konfidensintervallerne er små.

7 MODELLERINGEN AF DET KOMPLETTE DATASÆT

Som allerede nævnt var eksemplet, som vi gennemgik i forrige afsnit, simplificeret en del i sammenligning med den egentlige analyse. I dette

afsnit forklarer vi fremgangsmåden for vores brug af GAMMs til at modellere hele datasættet. Vi tilføjede løbende nye faktorer og evaluerede deres egnethed trinvist ved brug af modelsammenligning med funktionen *compareML()*, som beskrevet i afsnit 5.1.

Vi startede med at lave en model, der kun inkluderede vores kontrolvariable og tilfældige variable. Dvs. vi inkluderede variablene *auditivt vurderet emfase* (emfase | ikke emfase), *intrinsisk F0* (ens vokalhøjde i alle trykgruppens stavelser | forskellig vokalhøjde i en eller flere stavelser) og *antal stavelser* (2–6) som kontrolvariable, og vi inkluderede taler som en tilfældig smooth-term. Inklusionen af den tilfældige variabel er det, der gør modellen til en GAMM (dvs. en GAM med blandede, *mixed*, effekter). Rå F0-målinger kan ikke sammenlignes på tværs af talere, da forskellige mennesker jo har forskelligt udformede struber. Der findes en del måder at omregne F0 til en skala, der er bedre egnet til sammenligning på tværs af talere. F.eks. omregner Tøndering (2008: 44) F0 i Hz til halvtoner med talerens 10 %-percentil som base. Vi valgte at benytte en oktav-median-skala (De Looze & Hirst 2014). Hirst & De Looze (2021) gennemgår neurofysiologisk evidens for at regne oktaver for den naturlige enhed i perceptionen af toner. Med udgangspunkt i dette benyttede vi formelen til at omregne F0 til en normaliseret oktav-median-skala for hver taler i hver optagelse. Normalisering med medianen sikrer imod uønsket indflydelse fra fejlagtige ekstreme værdier i grundtonemålingerne (Hirst & De Looze 2021: 353). Da talernes stemmer kunne tænkes at ændre sig over tid, udførte vi omregningerne separat for 80'erne og 00'erne for generation 1 og 2.

Næste trin var så at tilføje de faktorer, vi egentlig er interesserede i. Hver gang, vi tilføjede en ny faktor, sammenlignede vi den nye, mere komplekse model med dens forgænger. Metoden her er, lidt simplificeret, at inkludere tilføjede faktorer, hvis de forbedrer modellen, og omvendt ikke at inkludere dem, hvis det observerede fald i AIC ikke er statistisk signifikant (se afsnit 5.1), altså hvis det indikerer, at tilføjjelsen ikke forbedrer modellen. På denne måde fandt vi frem til, at socialklasse ikke er en brugbar prædikator i modelleringen af grundtoneforløbet i korpusset: Forskellen i AIC på de to modeller var 0,23 ($p = 0,48$). Det betyder med andre ord, at vi ikke finder statistisk signifikante forskelle i trykgruppemønstret mellem talere fra arbejderklassen og middelklassen.

For at finde frem til eventuelle sammenhænge mellem de variable, vi er interesserede i, undersøgte vi interaktionseffekter mellem køn, generation og optageår. F.eks. kan man forestille sig, at køns påvirkning på grundtoneforløbet vil være forskelligt i de tre generationer, og for hver generation kan det have en forskellig effekt i de to optageår. I GAMMs findes der ikke en metode til at undersøge interaktioner, der svarer til metoden i lineær regression. Måden, man gør det på, er i stedet at behandle de forskellige variable som niveauer i en overordnet variabel, dvs. én variabel, hvis niveauer koder alle tre faktorer. Derfor kan det ikke lade sig gøre i en GAMM at isolere hovedeffekter i interaktionsvariable; de kan kun tolkes i kombination. Vi har konstrueret en faktor, der kombinerer generation-optageår-køn, og viser i resultaterne nedenfor, hvordan den påvirker modelleringen af det fulde datasæt. Denne treleddede interaktion viste sig nemlig at være den bedste måde at beskrive data på. Den model, vi endte med at bruge, kan altså opsummeres således: F0 (oktav-median-skala) modelleres som funktion af målingstidspunkt, taler som tilfældig smooth-faktor, emfase, intrinsisk F0 og antal stavelser som kontrolfaktorer (både som parametriske faktorer og som smooth-faktorer), og til sidst den treleddede interaktion mellem køn, generation og optagelsesår (både parametriske faktorer og som smooth-faktorer). Vi forsøgte med andre ord at beskrive de lineære og nonlineære effekter af de forskellige faktorer, der i kombination bedst tilnærmer sig det generelle kurvede forløb. Forskelligt fra det, vi viste i figur 5, tager denne mere komplekse model højde for, at hver enkelte talers kurver kan have forskellige vægtninger af basisfunktionerne. Vi testede mængden af autokorrelation i modellen (da man kan forvente at punkterne i en tidsserie korrelerer med punkterne før), og der viste sig ikke at være så høj en grad af autokorrelation, at det kunne blive problematisk, da residualerne per punkt i tidsserien var under 0,1 (og værdier under 0,2 burde ikke være problematiske; se Sóskuthy 2017 for en fuldstændig redegørelse for autokorrelationens rolle i vurdering af tidsserier).

I tabel 5 viser vi modelopsummeringen af modellens parametriske termer, dvs. de globale forskelle mellem faktorerne, uden at der tages højde for forskelle gennem forløbet (som er det, smooth-termerne siger noget om). Opsummeringen her fortæller os, om en faktor bidrager til den generelle grundtonefrekvens, dvs. om den bliver højere eller lavere,

men ikke om den påvirker konturens forløb. F.eks. kan vi se, at forskelle i intrinsisk F0 ikke har statistisk signifikant indflydelse på grundtonefrekvensen, mens emfase har.

TABEL 5. MODELLENS ESTIMATER FOR DE PARAMETRISKE KOEFFICIENTER

Term	Estimat	Standardfejl	<i>t</i> -værdi	<i>p</i> -værdi
Skæringspunkt	1,47	0,11	13,8	< 0,001
Emfase	-1,00	0,05	-21,7	< 0,001
Intrinsisk F0	-0,01	0,02	-0,7	0,47
3 stavelser	-0,44	0,02	-18,7	< 0,001
4 stavelser	-0,87	0,03	-27,5	< 0,001
5 stavelser	-1,13	0,06	-20,1	< 0,001
6 stavelser	-1,32	0,10	-13,1	< 0,001
Generation 2, 80'erne (kvinder)	-0,08	0,15	-0,5	0,58
Generation 1, 00'erne (kvinder)	-0,03	0,05	-0,6	0,55
Generation 2, 00'erne (kvinder)	0,13	0,15	-0,9	0,37
Generation 3, 00'erne (kvinder)	-0,28	0,13	-2,1	0,04
Generation 1, 80'erne (mænd)	-0,21	0,14	-1,5	0,05
Generation 2, 80'erne (mænd)	-0,26	0,16	-1,7	0,10
Generation 1, 00'erne (mænd)	0,07	0,14	0,5	0,62
Generation 2, 00'erne (mænd)	0,04	0,15	0,3	0,77
Generation 3, 00'erne (mænd)	-0,22	0,13	-1,7	0,09

I tabel 6 ses modelopsummeringen for smooth-termerne, som estimerer forskelle gennem trykgruppens grundtoneforløb. Disse resultater fortæller noget om, hvor meget de forskellige faktorer bidrager til den kurvede del af modelleringen af forløbet. Her rapporteres de estimerede frihedsgrader (der bl.a. afgøres af basisfunktionerne), referencefrihedsgrader, resultatet af en *F*-test og dertilhørende *p*-værdi. De estimerede frihedsgrader siger noget om, hvor meget nonlinearitet, faktoren bidrager med: Hvis værdien er tæt på 1, modelleres faktoren lineært, fordi smoothing-parametren er høj, så smooth'en bliver til en lige linje, der kun bruger én frihedsgrad, selvom der indgår flere basisfunktioner i modelleringen. Hvis værdien er over 1, er faktorens bidrag til det samlede estimat nonlinear. Det maksimale antal frihedsgrader begrænses af antallet af basisfunktioner, så der er maksimalt *k*-1 frihedsgrader (dvs.

antallet af knudepunkter minus én, da den første basisfunktion er lig med skæringspunktet).

TABEL 6. MODELLENS ESTIMEREDE SIGNIFIKANS FOR SMOOTH-TERMERNE

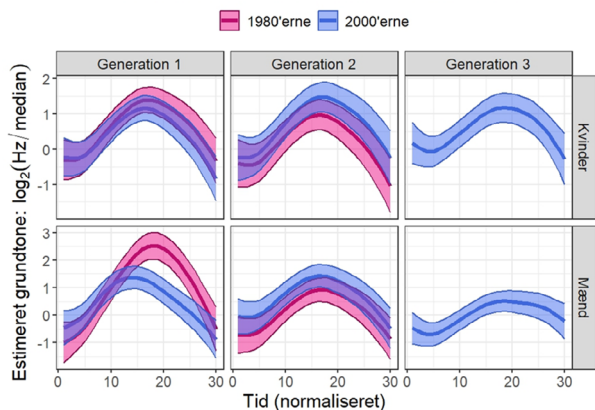
Term	Est. frihedsgrader	Ref. frihedsgrader	<i>F</i> -værdi	<i>p</i> -værdi
Tid	0,01	0,14	0,0	0,975
Taler	298,86	579,00	4,0	< 0,001
Emfase	7,73	8,40	15,1	< 0,001
Intrinsisk F0	4,82	5,88	11,8	< 0,001
2 stavelser	3,52	3,74	12,5	< 0,001
3 stavelser	3,89	3,95	34,4	< 0,001
4 stavelser	3,92	3,97	48,0	< 0,001
5 stavelser	3,90	3,98	38,8	< 0,001
6 stavelser	2,95	4,00	48,4	< 0,001
Generation 1, 80'erne (kvinder)	4,14	4,96	3,9	< 0,001
Generation 2, 80'erne (kvinder)	1,00	1,00	0,3	0,61
Generation 1, 00'erne (kvinder)	3,77	4,47	3,7	< 0,001
Generation 2, 00'erne (kvinder)	1,00	1,00	0,4	0,53
Generation 3, 00'erne (kvinder)	2,19	2,53	1,4	0,19
Generation 1, 80'erne (mænd)	4,02	4,81	10,0	< 0,001
Generation 2, 80'erne (mænd)	0,00	0,00	0,0	0,99
Generation 1, 00'erne (mænd)	4,39	5,12	5,5	< 0,001
Generation 2, 00'erne (mænd)	1,00	1,00	13,2	< 0,001
Generation 3, 00'erne (mænd)	2,62	3,00	2,9	0,03

Samlet set fortæller disse modelopssummeringer os, at også når vi kontrollerer for antallet af stavelser, forskelle i vokalhøjde, graden af prominens og desuden tager højde for den tilfældige variation mellem talerne, er der en effekt både af generation og optagelsesår, dvs. forskelle i både apparent time og real time, på trykgruppemønsteret. Det kan ses ved, at den treleddede faktor generation-optageårti-køn bidrager signifikant til modelleringen af konturerne, jf. *p*-værdierne i de nederste rækker i tabel 6. Forskellene er dog ikke ens for mænd og kvinder. Modelopssummeringerne fortæller os imidlertid ikke, *hvor* i grundtoneforløbet i trykgruppen disse forskelle optræder. Derfor vil

vi se nærmere på en visuel repræsentation af modellens forudsigelser. Bemærk, at såvel parametriske termer som smooth-termer påvirkes af antallet af stavelser i trykgruppen. Det er ikke overraskende, idet der jo foretages lige mange målinger i hver forekomst uanset trykgruppens længde. Men det betyder samtidig, at en visualisering af det tonale trykgruppemønster må præsenteres separat for de forskellige længder. Da datasættet indeholder langt flest observationer af trykgrupper på to og tre stavelser (hhv. 816 og 1072), viser vi nedenfor modellerede grundtonekurver for dem.

I figur 9 viser vi den mest komplekse models estimerer for grundtoneforløbet i tostavede trykgrupper for de tre forskellige generationer i de to årtier fordelt efter talernes køn. Det første, man kan bemærke, er, at kurverne langt fra er forskellige gennem hele forløbet i trykgruppen. Der er overlap af kurverne for 80'erne og 00'erne flere steder, og der er ligheder mellem kurvernes form på tværs af de tre generationer. Men der er også forskelle: Selvom der umiddelbart ikke er de store forskelle på tværs af generation hos kvinderne, er det initiale fald ind i den betonedede stavelse dog noget mere omfangsrigt hos den yngste generation sammenlignet med de to andre, og der er i øvrigt tegn på et sådant fald i alle tre generationer. Desuden er der hos kvinderne en forskel i overordnet gruntoneleje i real-time, tydeligst hos generation 2, men der er ikke nogen tydelig forskel i konturens form. Heller ikke hos mændene er der tegn på en ændring af konturen hos generation 2 i real-time, men det er der til gengæld hos generation 1, som vises i panelet nederst til venstre. I optagelserne fra 80'erne er der kun den allersvageste antydning af et initialt fald, men til gengæld ses en stigning hele vejen gennem den betonedede stavelse. Anderledes ser det ud i optagelserne fra 00'erne, med nogen tendens til et initialt fald og et toppunkt, der nås tidligere i den tostavede trykgruppe. Det initiale fald er endnu tydeligere hos generation 2 og generation 3. Vi bemærker desuden, at der i alle tre generationer og i begge optagelsesårtier stadig ses et finalt fald i trykgruppen, også her hvor de kun indeholder to stavelser. Det minder om det forløb, vi præsenterede i figur 2 ovenfor, dog med den forskel, at det initiale fald bliver mere udtalt i apparent time for både kvinder og mænd og i real time for den ældste generation af mændene.

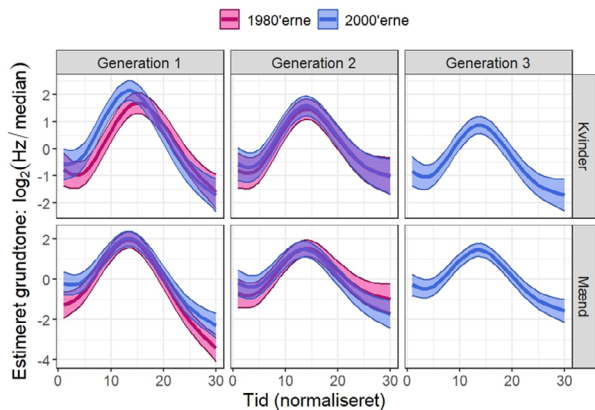
FIGUR 9. PLOT OVER ESTIMEREDE F0-KURVER FOR TOSTAVEDE TRYKGRUPPER EFTER GENERATION, ÅRTI OG KØN



Figuren viser modellens estimerede F0-konturer for hver generation fordelt efter optageår og køn. De lyserøde linjer viser konturerne for optagelserne i 80'erne, mens de blå linjer viser optagelserne fra 00'erne. De lysere bånd angiver 95 %-konfidensintervaller for estimatet. På y-aksen angives det estimerede udsving fra median-F0.

I figur 10 præsenterer vi de estimerede grundtoneforløb i trestavede trykgrupper for de forskellige grupper af talere.

FIGUR 10. PLOT OVER ESTIMEREDE F0-KURVER FOR TRESTAVEDE TRYKGRUPPER EFTER GENERATION, ÅRTI OG KØN



Figuren viser modellens estimerede F0-konturer for hver generation fordelt efter optageår og køn. De lyserøde linjer viser konturerne for optagelserne i 80'erne, mens de blå linjer viser optagelserne fra 00'erne. De lysere bånd angiver 95 %-konfidensintervaller for estimatet. På y-aksen angives det estimerede udsving fra median-F0.

Også for trykgrupper på tre stavelser ser vi det initiale fald for alle slags talere med undtagelse af mændene i den ældste generation i optagelserne fra 1980'erne. Faldet ind i den betonedede stavelse er til gengæld tydeligt for samme gruppe talere, altså mændene i generation 1, i optagelserne fra 00'erne, hvilket indikerer samme ændring i real time som den, vi så i de tostavede trykgrupper. Bemærk, at der også ser ud til at være en real time-forandring i trestavede trykgrupper for kvinder i generation 1, men i den anden retning: Det initiale fald er *mindre* omfangsrigt i optagelserne fra 00'erne end i dem fra 80'erne.

Sammenholdt med konturerne i figur 6 ser det ud til, at der især har været forandringer i trykgruppemønsteret for generation 1 mellem de to optagetidspunkter, dvs. real time-forandringer for den ældste generation. Men hvor vi i analyserne på gruppeniveau som præsenteret i figur 9 og 10 ser størst forskel i real time hos mændene, kan vi med baggrund i de enkelte taleres gennemsnitlige konturer i figur 6 se, at der i høj grad er tale om interindividuel variation inden for grupperne af både mænd og kvinder. Der er flere af mændene, der ikke udviser tegn på forandring, men dem, der gør, ændrer adfærd fra det næstvedske i retning mod det københavnske. Kvinderne i generation 1 ændrer sig også, men det har ikke kraftigt udslag på gruppeniveau. Ser vi på de enkelte kvinder i generation 1, kan vi se, at en af talerne, GEO, havde et københavnsk-lignende trykgruppemønster med et relativt omfangsrigt fald ind i den betonedede stavelse i 1980'erne, men i 00'erne ser hun ud til at have ren stigning gennem den betonedede stavelse. Hun bliver med andre ord mere næstvedsk-lydende, og altså mere som de andre kvinder i samme generation, hvilket kan forklare, hvorfor der på gruppeniveau ikke ses en tydelig tendens til et skift. På samme måde for generation 2 som præsenteret i figur 7 var der interindividuelle forskelle i grupperne af kvinder og mænd med eksempler på enkelte talere, der blev mere næstvedske i 00'erne. Vi har ikke detaljerede oplysninger om de enkelte talere i korpus, der giver mulighed for at forklare de forskellige tendenser til at skifte mod at blive mere eller mindre næstvedsk-lydende gennem livet. Man kunne forestille sig, med udgangspunkt i Monka (2013), at de talere, der havde mere københavnsk-lydende trykgruppemønstre i 1980'erne, har haft arbejde i København, og i det hele taget har orienteret sig mod hovedstaden, men hvorfor skifter nogle af dem så tilbage til

et klart lokalt trykgruppemønster? Det må forblive ved spekulationerne, indtil nøjere analyser af de enkelte taleres baggrund kan inddrages.

Hvor forskellene i real time således i høj grad ser ud til at være individuelle kendetegn for enkelte talere, er der klarere tendenser i apparent time. Der er hovedsageligt tale om et komprimeret tonalt spænd for talerne fra generation 3 sammenlignet med de to ældre generationer for både mænd og kvinder. Der er dog også en klar forskel i stejlheden af det finale fald mellem generation 1 og 2 – det er mere stejlt for de ældste talere, også i optagelserne fra 00'erne. Der er også stor forskel på omfanget af faldet ind i den betonedede stavelse og timingen af toppen af konturen på tværs af de tre generationer. Det er tydeligst for de tostavede trykgrupper, som præsenteret i figur 9. Her varer faldet ind i den betonedede vokal betydeligt længere hos generation 3, og toppen nås først to tredjedele inde i forløbet hos dem, mens den nås allerede omtrent halvvejs inde i forløbet for generation 1 og 2. Ser man på de enkelte talere i generation 3, er der også klare indikationer på en tilnærmelse i retning af det københavnsk-lydende mønster, især hos kvinderne, som altså på dette tidspunkt endnu er teenagere.

8 SÅ ER NÆSTVED BLEVET DET NYE KØBENHAVN?

Vi er stadig i gang med at analysere trykgruppemønstre hos talerne fra København i LANCHART-korpusset, og derfor kan vi ikke direkte sammenligne vores resultater fra Næstved med grundtoneforløb fra København. Vi har således ikke grundlaget for at konkludere definitivt, om næstvedsk er blevet mere københavnsk. Dog kan vi konstatere, at trykgruppemønstret ser ud til at have ændret sig på tværs af generationerne, og at det også har ændret sig fra 80'erne til 00'erne. Vores resultater peger på, at mændene i generation 1 i de gamle optagelser producerer et grundtoneforløb, der også akustisk passer til Grønnums stilisering af trykgruppemønsteret (sammenlign figur 1 med figur 6, 9 og 10). Det tyder på, at det initiale fald i trykgruppen, som vi fremhæver fra hendes målinger, og som også dukker op i vores målinger, er subtilt og ikke lige så fremtrædende som i københavnsk, hvorfor talere fra Næstved og København har lydt forskellige i 1980'erne, og formentlig også i 00'erne (if. Tøndering & Pharao 2020). Dog er det initiale fald på vej til at blive mere fremtrædende, hvilket antyder, at næstvedsk intonation er i færd

med at udvikle sig i retning af københavnsk. Vi observerer desuden, at det initiale fald har været til stede blandt kvinderne i både generation 1 og 2 allerede i optagelserne fra 80'erne, men at det i optagelserne fra 00'erne ser ud til at blive mere tydeligt blandt de yngre talere sammenlignet med de ældre. Det kunne tyde på, at tilpasningen i retning af det københavnske trykgruppemønster begyndte hos kvinderne og gradvist tager til på tværs af generationerne, samtidig med at mændene fra den ældste af de tre generationer kommer med på trenden i 00'erne. Det er dog væsentligt at bemærke, både at det næstvedske trykgruppemønster stadig udviser et fald allerede i løbet af den første posttoniske stavelse, og at der tilsyneladende er betragtelig individuel variation. Samlet set peger dette på, at der nok er forandringer i trykgruppemønsteret hos nogle talere fra Næstved, men ikke at det nødvendigvis vil blive ligesom det københavnske og heller ikke i et omfang, der gør, at man kan sige, at forskellen på de to regiolekter er blevet ophævet. Hvis forandringen mod et anderledes tonalt mønster er gradvist, sådan som vores resultater på tværs af generationer antyder, tilbagesår det at forklare, hvordan sådan en gradvis forandring opfattes af den enkelte sprogbruger, og hvorfor det kunne se ud til, at de enkelte talere, der skifter i løbet af deres levetid, ser ud til at foretage et skift fra et stigende-faldende mønster, som i traditionel næstvedsk regiolekt, til faldende-stigende-faldende. Kan der med andre ord være gradvise ændringer på gruppeniveau som resultat af kategorielle ændringer på individniveau, eller kan forandringen på individniveau også vises at være graduel, hvis man har tilstrækkeligt finkornet tidsopløsning? Det må være op til senere undersøgelser at svare på dette.

Vi har her undersøgt trykgruppemønstrene alene gennem akustisk analyse af grundtoneforløbet. Som vi har nævnt et par gange, har tidligere undersøgelser (fx Tøndering & Pharao 2020) omtalt et auditivt indtryk af mandlige talere fra generation 2, hvor især talerne fra middelklassen i 00'erne lød som københavnske talere, dvs. med et fald ind i den betonedede stavelse, mens mandlige talere i arbejderklassen fra samme generation lød til at have et rent stigende grundtoneforløb gennem den betonedede stavelse. Vores undersøgelse underbygger ikke denne forskel på tværs af socialklasser, men påviser til gengæld en tydelig forskel for mandlige talere fra den ældre generation 1. Det forekommer oplagt at supplere dette studie med et perceptionseksperiment, hvor lyttere skal

kategorisere klip med talere fra forskellige generationer i Næstved og København. Herved vil vi kunne se, om de tendenser til forandringer i trykgruppemønstret, der kan konstateres på akustisk grundlag, også svarer til folks perceptuelle virkelighed. Det vil gøre det muligt at afgøre, om det, de statistiske analyser viser, svarer til folks opfattelse.

TAK

Vi vil gerne takke den anonyme fagfælle, NyS-redaktionen, Nina Grønnum og Frans Gregersen, der alle har læst tidligere versioner af vores manuskript og givet indsigtfulde kritiske og konstruktive kommentarer. Også tak til deltagerne i forskergruppen Sprog & Samfunds mandagsmøder og deltagerne i MAMPF-studiekredsen, der sørgede for, at vi tænkte os om under arbejdet. Og især tak til Emma Vanessa Alice Farr, der arbejdede med segmentering af materialet.

Christophe Zerakitsky Vies, ph.d.-studerende
Oticon og Institut for Nordiske Studier og Sprogvidenskab
Københavns Universitet
czv@hum.ku.dk

Nicolai Pharaos, lektor
Institut for Nordiske Studier og Sprogvidenskab
Københavns Universitet
nicolaip@hum.ku.dk

LITTERATUR

- Arvaniti, A. 2022. The autosegmental-metrical model of intonational phonology. J. Barnes & S. Shattuck-Hufnagel (red.), *Prosodic theory and practice*, 25–63. The MIT Press. DOI: 10.7551/mitpress/10413.001.0001.
- Basbøll, H. 2005. *The phonology of Danish*. Oxford: Oxford University Press.
- Boersma, P. & D. Weenink. 2023. Praat: doing phonetics by computer [Computerprogram]. <http://www.praat.org/>.

- Cavanaugh, J.E. & A.A. Neath. 2019. The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements. *WIREs Computational Statistics* 11(3). e1460. DOI: 10.1002/wics.1460.
- De Looze, C. & D. Hirst. 2014. The OMe (Octave-Median) scale: A natural scale for speech melody. N. Campbell, K. Gibbon & D. Hirst (red.), *Speech prosody* 7, 910–914. ISCA. DOI: 10.21437/SpeechProsody.2014-171.
- Diderichsen, P. 2023. removeboundaries [Praat script].
- Dyhr, N.-J. 1993. An acoustical investigation of the fundamental frequency in Danish spontaneous speech. B. Granström & L. Nord (red.), *Nordic prosody VI*, 23–32. Stockholm: Almqvist & Wiksell.
- Dyhr, N.-J. 1995. The fundamental frequency in Danish spontaneous speech with special reference to syllables boosted for emphasis. J. Rischel & H. Basbøll (red.), *Aspects of Danish prosody*, 49–67. Odense: Odense University Press.
- Gregersen, F. 2009. The data and design of the LANCHART study. *Acta Linguistica Hafniensia* 41(1). 3–29. DOI: 10.1080/03740460903364003.
- Gregersen, F., M. Maegaard & N. Pharaoh. 2014. The Lanchart Corpus. J. Durand, U. Gut & G. Kristoffersen (red.), *The Oxford handbook of corpus phonology*, 534–545. Oxford: Oxford University Press. DOI: 10.1093/oxford-hb/9780199571932.013.007.
- Gregersen, F. & N. Pharaoh. 2016. Lects are perceptually invariant, productively variable: A coherent claim about Danish lects. *Lingua* (Coherence, covariation and bricolage. Various approaches to the systematicity of language variation) 172–173. 26–44. DOI: 10.1016/j.lingua.2015.12.001.
- Grønnum, N. 1992. *The groundworks of Danish intonation*. København: Museum Tusculanum Press.
- Grønnum, N. 2005. *Fonetik og fonologi: almen og dansk*. 3. udg. København: Akademisk Forlag.
- Grønnum, N. 2022. Modeling Danish intonation. J. Barnes & S. Shattuck-Hufnagel (red.), *Prosodic theory and practice*, 85–116. The MIT Press. DOI: 10.7551/mit-press/10413.003.0005.
- Hirst, D. & C. De Looze. 2021. Fundamental frequency and pitch. R.-A. Knight & J. Setter (red.), *The Cambridge handbook of phonetics*, 336–361. Cambridge: Cambridge University Press. DOI: 10.1017/9781108644198.014.
- Kirby, J.P. & D.R. Ladd. 2016. Effects of obstruent voicing on F0: Evidence from “true voicing” languages. *Journal of the Acoustical Society of America* 140. 2400–2411. DOI: 10.1121/1.4962445.

- Kristiansen, T., M. Maegaard & N. Pharao. 2011. Det er intonationen, vi hører det på: Perceptionsstudier i genkendelse af moderne dansk med henholdsvis jysk og københavnsk aksang. T. Arboe & I. Schonderbeck (red.), *Jysk, øsmål, rigsdansk mv.: Studier i dansk sprog med sideblik til nordisk og tysk. Festskrift til Viggo Sørensen og Ove Rasmussen*, 207–224. Aarhus: Peter Skautrup Centret for Jysk Dialektforskning.
- Labov, William. 1972. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania.
- Ladd, D.R. 2008. *Intonational phonology*. 2. udg. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511808814.
- LANCHART. 2010. Kodningsmanual til diskurskontekstanalyse (aka. iiv-analyse). https://dgcss.hum.ku.dk/online-ressourcer/lanchart-korpusset/iiv-kodningsmanual_januar-2011/ (tilgået 3. juli 2024).
- Maegaard, M. m.fl. 2013. Diffusion of language change: Accommodation to a moving target. *Journal of Sociolinguistics* 17(1). 3–36. DOI: 10.1111/josl.12002.
- Monka, M. 2013. *Sted og sprogforandring – En undersøgelse af sprogforandring i virkelig tid hos mobile og bofaste informanter fra Vinderup, Odder og Tinglev*. Ph.d.-afhandling. Københavns Universitet. https://theses.hal.science/tel-00843522/file/Ph.d._2013_Monka.pdf
- Pharao, N. 2024a. *slet_alt_andet_end_ortografi_og_interaktionsstruktur* [Praat script].
- Pharao, N. 2024b. *monologExtractorTOT* [Praat script].
- Pichora-Fuller, M.K. m.fl. 2016. Importance of F0 for predicting vocal emotion categorization. *The Journal of the Acoustical Society of America* 140(4), 3401. DOI: 10.1121/1.4970917.
- R Core Team. 2022. R: A language and environment for statistical computing [Computerprogram]. Wien, Austria. <http://www.R-project.org/>
- Rij, J. van, R.H. Baayen & H. van Rijn. 2022. *itsadug: Interpreting time series and autocorrelated data using GAMMs*. R Package version 2.4.1.
- Sankoff, G. 2006. Age: Apparent time and real time. B. Keith (red.), *Elsevier encyclopedia of language & linguistics* (2. udg.), 110–116. Elsevier. DOI: 10.1016/B0-08-044854-2/01479-6.
- Schachtenhaufen, R. 2022. *Ny dansk fonetik*. Hellerup: BoD – Books on Demand.
- Scheuer, J. 2019. Tonegang i fynsk. Akustisk analyse af dialektoptagelser. *Danske Talesprog* 19. 131–152.
- Simpson, G.L. 2024. *Graceful ggplot-based graphics and other functions for GAMs fitted using mgcv*. R package version 0.9.0. <https://gavinsimpson.github.io/gratia/>

- Sóskuthy, M. 2017. Generalised additive mixed models for dynamic analysis in linguistics: A practical introduction. *arXiv preprint*. 1703.05339. <http://arxiv.org/abs/1703.05339> (tilgået 11. juni 2024).
- Sóskuthy, M. 2021. Evaluating generalised additive mixed modelling strategies for dynamic speech analysis. *Journal of Phonetics* 84, 101017. DOI: 10.1016/j.wocn.2020.101017.
- Thorsen, N.G. 1979. Interpreting raw fundamental-frequency tracings of Danish. *Phonetica* 36(1). 57–78. DOI: 10.1159/000259946.
- Thorsen, N.G. 1984. Variability and invariance in Danish stress group patterns. *Phonetica* 41(2). 88–102. DOI: 10.1159/000261714.
- Thorsen, N.G. 1988. Stress group patterns, focus signalling and sentence intonation in two regional Danish standard languages: Aalborg and Næstved. *Annual Report of the Institute of Phonetics University of Copenhagen* 22. 145–195. DOI: 10.7146/aripuc.v22i.131896.
- Thorsen, N.G. 1989. Stress group patterns, sentence accents and sentence intonation in southern Jutland (Sønderborg and Tønder) – with a view to German. *Annual Report of the Institute of Phonetics University of Copenhagen* 23. 1–85. DOI: 10.7146/aripuc.v23i.131902.
- Torre, P. & J.A. Barlow. 2009. Age-related changes in acoustic characteristics of adult speech. *Journal of Communication Disorders* 42(5). 324–333. DOI: 10.1016/j.jcomdis.2009.03.001.
- Tøndering, J. 2008. *Skitser af prosodi i spontant dansk*. Ph.d.-afhandling. Københavns Universitet. https://www.researchgate.net/profile/John-Tondering/publication/268345710_Skitser_af_prosodi_i_spontant_dansk/links/54e4c8410cf29865c335179f/Skitser-af-prosodi-i-spontant-dansk.pdf
- Tøndering, J. 2010a. Prosodiske fraser og syntaktisk struktur i spontan tale. *NyS, Nydanske Sprogstudier* (39), 166–198. DOI: 10.7146/nys.v39i39.13538.
- Tøndering, J. 2010b. Praat for begyndere. *NyS, Nydanske Sprogstudier* 39. 199–222. DOI: 10.7146/nys.v39i39.13540.
- Tøndering, J. & N. Pharao. 2020. Recognising regional varieties of Danish. *Linguistics Vanguard*. DOI: 10.1515/lingvan-2019-0020.
- Vies, C.Z. 2021. Prosodisk demarkering og domæneinitial styrkelse i dansk. *Journal of Language Works – Sprogvidenskabeligt Studentertidsskrift* 6(1), 5–21.

- Wieling, M. 2018. Analyzing dynamic phonetic data using generalized additive mixed modeling: A tutorial focusing on articulatory differences between L1 and L2 speakers of English. *Journal of Phonetics* 70, 86–116. DOI: 10.1016/j.wocn.2018.03.002.
- Winter, B. & M. Wieling. 2016. How to analyze linguistic change using mixed models, Growth curve analysis and Generalized Additive Modeling. *Journal of Language Evolution* 1(1), 7–18. DOI: 10.1093/jole/lzv003.
- Wood, S.N. 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society Series B: Statistical Methodology* 73(1), 3–36. DOI: 10.1111/j.1467-9868.2010.00749.x.
- Wood, S.N. 2017. *Generalized additive models: An introduction with R*. 2. udg. New York: Chapman and Hall/CRC. DOI: 10.1201/9781315370279.
- Young, N. J. 2024. DanFA 4.0 – Forced Alignment of Danish, version 4.0. www.autophon.se.
- Young, N.J. & K. Anikwe. 2024. Autophon – Automatic phonetic annotation of Nordic languages [Computerprogram]. <https://autophon.org/> (tilgæet 14. juni 2024).

APPENDIKS

Oversigt over samtlige informanter med angivelse af antallet af forekomster af trykgrupper i korpus pr. taler fordelt på trykgrupper af forskellig længde.

Generation	Køn	Social- klasse	Årti	Taler	Antal stavelser	n
1	f	AK	1980	AJO	2	6
1	f	AK	1980	AJO	3	10
1	f	AK	1980	AJO	4	2
1	f	AK	1980	AJO	5	1
1	f	AK	1980	ALA	2	3
1	f	AK	1980	ALA	3	3
1	f	AK	1980	ALA	4	4
1	f	AK	1980	HST	2	4
1	f	AK	1980	HST	3	4
1	f	AK	1980	HST	4	4
1	f	AK	1980	LHA	2	5
1	f	AK	1980	LHA	3	5
1	f	AK	1980	LHA	4	1
1	f	AK	2000	AFL	2	9
1	f	AK	2000	AFL	3	8
1	f	AK	2000	AFL	4	1
1	f	AK	2000	AJO	2	3
1	f	AK	2000	AJO	3	10
1	f	AK	2000	AJO	4	1
1	f	AK	2000	ALA	2	7
1	f	AK	2000	ALA	3	7
1	f	AK	2000	HST	2	10
1	f	AK	2000	HST	3	21
1	f	AK	2000	HST	4	5
1	f	AK	2000	LHA	2	10
1	f	AK	2000	LHA	3	23
1	f	AK	2000	LHA	4	6
1	f	MK	1980	BKO	2	8
1	f	MK	1980	BKO	3	13

Generation	Køn	Social- klasse	Årti	Taler	Antal stavelser	n
1	f	MK	1980	BKO	6	1
1	f	MK	1980	GAN	2	1
1	f	MK	1980	GAN	3	9
1	f	MK	1980	GAN	4	1
1	f	MK	1980	GEO	2	2
1	f	MK	1980	GEO	3	3
1	f	MK	1980	GEO	4	4
1	f	MK	1980	HKA	2	2
1	f	MK	1980	HKA	3	6
1	f	MK	1980	HKA	4	4
1	f	MK	1980	LNL	2	1
1	f	MK	1980	LNL	3	4
1	f	MK	1980	LNL	5	1
1	f	MK	1980	MLA	2	4
1	f	MK	1980	MLA	3	6
1	f	MK	1980	MLA	4	6
1	f	MK	1980	MLA	5	2
1	f	MK	1980	MLS	2	2
1	f	MK	1980	MLS	3	11
1	f	MK	1980	MLS	4	2
1	f	MK	2000	BKO	2	8
1	f	MK	2000	BKO	3	18
1	f	MK	2000	BKO	4	4
1	f	MK	2000	GAN	2	2
1	f	MK	2000	GAN	3	4
1	f	MK	2000	GAN	4	3
1	f	MK	2000	GEO	2	2
1	f	MK	2000	GEO	3	7
1	f	MK	2000	GEO	4	4
1	f	MK	2000	HKA	2	5
1	f	MK	2000	HKA	3	4
1	f	MK	2000	HKA	4	1
1	f	MK	2000	LNL	2	1
1	f	MK	2000	LNL	3	5

Generation	Køn	Social- klasse	Årti	Taler	Antal stavelser	n
1	f	MK	2000	MLA	2	7
1	f	MK	2000	MLA	3	11
1	f	MK	2000	MLA	4	5
1	f	MK	2000	MLA	5	1
1	f	MK	2000	MLS	2	7
1	f	MK	2000	MLS	3	5
1	f	MK	2000	MLS	4	3
1	m	AK	1980	CNL	3	12
1	m	AK	1980	CNL	4	1
1	m	AK	1980	LJE	2	5
1	m	AK	1980	LJE	3	9
1	m	AK	1980	LJE	4	3
1	m	AK	1980	TRA	2	13
1	m	AK	1980	TRA	3	8
1	m	AK	1980	TRA	4	1
1	m	AK	2000	BHS	2	15
1	m	AK	2000	BHS	3	12
1	m	AK	2000	BHS	4	10
1	m	AK	2000	BHS	5	3
1	m	AK	2000	CNL	4	4
1	m	AK	2000	CNL	5	2
1	m	AK	2000	FJN	2	13
1	m	AK	2000	FJN	3	17
1	m	AK	2000	FJN	4	4
1	m	AK	2000	FJN	6	1
1	m	AK	2000	LJE	2	12
1	m	AK	2000	LJE	3	15
1	m	AK	2000	TRA	2	9
1	m	AK	2000	TRA	3	11
1	m	AK	2000	TRA	4	4
1	m	AK	2000	TRA	5	1
1	m	MK	1980	BGO	2	3
1	m	MK	1980	BGO	3	4
1	m	MK	1980	BGO	4	3

Generation	Køn	Social- klasse	Årti	Taler	Antal stavelser	n
1	m	MK	1980	BGO	5	1
1	m	MK	1980	ENI	2	1
1	m	MK	1980	ENI	3	8
1	m	MK	1980	ENI	4	2
1	m	MK	1980	ENI	6	1
1	m	MK	1980	FHN	2	1
1	m	MK	1980	FHN	3	5
1	m	MK	1980	FHN	5	1
1	m	MK	1980	HJL	2	20
1	m	MK	1980	HJL	3	10
1	m	MK	1980	HJL	4	7
1	m	MK	1980	HJL	5	1
1	m	MK	1980	JKR	3	1
1	m	MK	1980	JKR	4	2
1	m	MK	1980	SKA	2	4
1	m	MK	1980	SKA	3	9
1	m	MK	1980	SKA	4	1
1	m	MK	2000	BGO	2	15
1	m	MK	2000	BGO	3	9
1	m	MK	2000	BGO	4	10
1	m	MK	2000	BGO	5	4
1	m	MK	2000	BGO	6	2
1	m	MK	2000	ENI	2	15
1	m	MK	2000	ENI	3	22
1	m	MK	2000	ENI	4	23
1	m	MK	2000	ENI	5	4
1	m	MK	2000	ENI	6	3
1	m	MK	2000	FHN	2	3
1	m	MK	2000	FHN	3	10
1	m	MK	2000	FHN	4	7
1	m	MK	2000	FHN	5	4
1	m	MK	2000	FHN	6	1
1	m	MK	2000	HJL	2	7
1	m	MK	2000	HJL	3	19

Generation	Køn	Social- klasse	Årti	Taler	Antal stavelser	n
1	m	MK	2000	HJL	4	10
1	m	MK	2000	HJL	5	9
1	m	MK	2000	HJL	6	2
1	m	MK	2000	JKR	2	5
1	m	MK	2000	JKR	3	13
1	m	MK	2000	JKR	4	8
1	m	MK	2000	JKR	5	2
1	m	MK	2000	SKA	2	10
1	m	MK	2000	SKA	3	21
1	m	MK	2000	SKA	4	9
1	m	MK	2000	SKA	5	3
2	f	AK	1980	ALY	2	6
2	f	AK	1980	ALY	3	4
2	f	AK	1980	ALY	4	1
2	f	AK	1980	CJA	2	2
2	f	AK	1980	CJA	3	3
2	f	AK	2000	ALY	2	17
2	f	AK	2000	ALY	3	22
2	f	AK	2000	ALY	4	9
2	f	AK	2000	ALY	5	2
2	f	AK	2000	CJA	2	8
2	f	AK	2000	CJA	3	18
2	f	AK	2000	CJA	4	2
2	f	AK	2000	CJA	5	1
2	f	MK	1980	CHN	2	3
2	f	MK	1980	CHN	3	4
2	f	MK	1980	CHN	4	3
2	f	MK	1980	CHN	6	1
2	f	MK	1980	COL	2	4
2	f	MK	1980	COL	3	7
2	f	MK	1980	COL	4	1
2	f	MK	1980	CRI	2	2
2	f	MK	1980	CRI	3	5
2	f	MK	1980	CRI	4	2

Generation	Køn	Social- klasse	Årti	Taler	Antal stavelser	n
2	f	MK	1980	CSK	2	1
2	f	MK	1980	CSK	3	7
2	f	MK	1980	KJS	2	2
2	f	MK	1980	KJS	3	5
2	f	MK	1980	MLG	2	7
2	f	MK	1980	MLG	3	7
2	f	MK	2000	CHN	2	4
2	f	MK	2000	CHN	3	30
2	f	MK	2000	CHN	4	5
2	f	MK	2000	CHN	5	1
2	f	MK	2000	COL	2	6
2	f	MK	2000	COL	3	22
2	f	MK	2000	COL	4	6
2	f	MK	2000	COL	5	1
2	f	MK	2000	CRI	2	5
2	f	MK	2000	CRI	3	19
2	f	MK	2000	CRI	4	6
2	f	MK	2000	CRI	5	2
2	f	MK	2000	CSK	2	5
2	f	MK	2000	CSK	3	12
2	f	MK	2000	CSK	4	4
2	f	MK	2000	CSK	5	2
2	f	MK	2000	KJS	2	5
2	f	MK	2000	KJS	3	13
2	f	MK	2000	KJS	4	4
2	f	MK	2000	MLG	2	2
2	f	MK	2000	MLG	3	27
2	f	MK	2000	MLG	4	1
2	f	MK	2000	MLG	5	1
2	f	MK	2000	MLG	6	2
2	m	AK	1980	PRB	2	15
2	m	AK	1980	PRB	3	5
2	m	AK	1980	PRB	4	4
2	m	AK	1980	PRB	5	1

Generation	Køn	Social- klasse	Årti	Taler	Antal stavelser	n
2	m	AK	1980	RJN	2	7
2	m	AK	1980	RJN	3	4
2	m	AK	1980	RJN	4	2
2	m	AK	2000	PRB	2	21
2	m	AK	2000	PRB	3	22
2	m	AK	2000	PRB	4	14
2	m	AK	2000	PRB	5	4
2	m	AK	2000	RJN	2	8
2	m	AK	2000	RJN	3	23
2	m	AK	2000	RJN	4	8
2	m	AK	2000	RJN	5	3
2	m	AK	2000	THA	2	10
2	m	AK	2000	THA	3	11
2	m	AK	2000	THA	4	6
2	m	AK	2000	THA	5	1
2	m	MK	1980	JAN	2	10
2	m	MK	1980	JAN	3	5
2	m	MK	1980	JAN	4	4
2	m	MK	1980	JAN	5	1
2	m	MK	1980	JBB	2	8
2	m	MK	1980	JBB	3	6
2	m	MK	1980	JBB	4	3
2	m	MK	1980	MBH	2	3
2	m	MK	1980	MBH	3	7
2	m	MK	1980	MBH	4	1
2	m	MK	1980	MJU	2	6
2	m	MK	1980	MJU	3	1
2	m	MK	1980	MJU	4	1
2	m	MK	2000	JAN	2	23
2	m	MK	2000	JAN	3	20
2	m	MK	2000	JAN	4	10
2	m	MK	2000	JAN	5	2
2	m	MK	2000	JBB	2	17
2	m	MK	2000	JBB	3	23

Generation	Køn	Social- klasse	Årti	Taler	Antal stavelser	n
2	m	MK	2000	JBB	4	4
2	m	MK	2000	JBB	5	3
2	m	MK	2000	JBB	6	3
2	m	MK	2000	MBH	2	32
2	m	MK	2000	MBH	3	23
2	m	MK	2000	MBH	4	13
2	m	MK	2000	MBH	5	3
2	m	MK	2000	MBH	6	1
2	m	MK	2000	MJU	2	19
2	m	MK	2000	MJU	3	29
2	m	MK	2000	MJU	4	15
2	m	MK	2000	MJU	5	7
2	m	MK	2000	MJU	6	3
3	f	AK	2000	MHV	2	4
3	f	AK	2000	MHV	3	4
3	f	AK	2000	MHV	4	1
3	f	AK	2000	NMN	2	5
3	f	AK	2000	NMN	3	3
3	f	AK	2000	NSA	2	8
3	f	AK	2000	NSA	3	3
3	f	AK	2000	NTH	2	2
3	f	AK	2000	NTH	3	3
3	f	AK	2000	NTH	4	1
3	f	AK	2000	SBO	2	11
3	f	AK	2000	SBO	3	6
3	f	AK	2000	SBO	4	3
3	f	MK	2000	ASL	2	3
3	f	MK	2000	ASL	3	7
3	f	MK	2000	ASL	4	1
3	f	MK	2000	ASM	2	8
3	f	MK	2000	ASM	3	10
3	f	MK	2000	ASM	4	3
3	f	MK	2000	CIF	2	6
3	f	MK	2000	CIF	3	4

Generation	Køn	Social- klasse	Årti	Taler	Antal stavelser	n
3	f	MK	2000	CIF	4	2
3	f	MK	2000	EMU	2	3
3	f	MK	2000	EMU	3	9
3	f	MK	2000	EMU	4	2
3	f	MK	2000	EMU	5	2
3	f	MK	2000	ITJ	2	5
3	f	MK	2000	ITJ	3	7
3	f	MK	2000	ITJ	4	4
3	f	MK	2000	KAU	2	6
3	f	MK	2000	KAU	3	5
3	f	MK	2000	KML	2	7
3	f	MK	2000	KML	3	6
3	f	MK	2000	LSG	2	1
3	f	MK	2000	LSG	3	2
3	f	MK	2000	LSG	4	1
3	f	MK	2000	MJG	2	8
3	f	MK	2000	MJG	3	1
3	f	MK	2000	MMS	2	5
3	f	MK	2000	MMS	3	2
3	f	MK	2000	MMS	4	1
3	f	MK	2000	PJH	2	4
3	f	MK	2000	PJH	3	7
3	f	MK	2000	PJH	4	2
3	m	AK	2000	DDG	2	25
3	m	AK	2000	DDG	3	17
3	m	AK	2000	DDG	4	4
3	m	AK	2000	FFA	2	8
3	m	AK	2000	FFA	3	10
3	m	AK	2000	JRO	2	2
3	m	AK	2000	JRO	3	6
3	m	AK	2000	MMM	2	7
3	m	AK	2000	MMM	3	6
3	m	AK	2000	MMM	4	8
3	m	AK	2000	NLO	2	4

Generation	Køn	Social- klasse	Årti	Taler	Antal stavelser	n
3	m	AK	2000	NLO	3	5
3	m	AK	2000	NLO	4	2
3	m	AK	2000	RKR	2	5
3	m	AK	2000	RKR	3	6
3	m	MK	2000	DBF	2	14
3	m	MK	2000	DBF	3	9
3	m	MK	2000	DBF	4	1
3	m	MK	2000	DBF	5	2
3	m	MK	2000	DRA	2	10
3	m	MK	2000	DRA	3	10
3	m	MK	2000	MIN	2	5
3	m	MK	2000	MIN	3	3
3	m	MK	2000	MIN	5	1
3	m	MK	2000	MPP	2	15
3	m	MK	2000	MPP	3	8
3	m	MK	2000	NCN	2	12
3	m	MK	2000	NCN	3	1
3	m	MK	2000	NCN	4	1
3	m	MK	2000	PML	2	15
3	m	MK	2000	PML	3	7
3	m	MK	2000	PML	4	7
3	m	MK	2000	PML	5	1
3	m	MK	2000	RNL	2	8
3	m	MK	2000	RNL	3	3
3	m	MK	2000	RNL	4	2