

# En analyse af realisationer af fonemet [A:] i udvalgte regioner og aldersgrupper i SpeechDat 2

Jacob Teckemeier og Kim Hornbech

Aalborg Universitet

E-post: [j\\_teckemeier@hotmail.com](mailto:j_teckemeier@hotmail.com), [k\\_hornbech@hotmail.com](mailto:k_hornbech@hotmail.com)

## Abstract

In our BA dissertation, we dealt with the question whether regional pronunciation variations of the phoneme [ A: ]<sup>1</sup> could be detected by means of an F'-analysis in the Danish version of the speech data base SpeechDat 2. We chose to analyse limited segments of SpeechDat 2 which were selected under consideration of gender, age and region. F'-analysis was run on the segments and we additionally examined the results using various statistic implements. Our investigation displayed patterns in the pronunciation which however were so vague that they could merely be described as tendencies. Upon this conclusion we proposed certain methodical adjustments as well as elaboration on the speech data base SpeechDat 2.

## Indledning

Nærværende artikel er et sammendrag af vores bachelorprojekt, der dannede baggrund for vores oplæg til DALF's årsmøde 2000. Formålet med vores projekt var ved hjælp af en F'-analyse at fastslå, hvorvidt det var muligt at efterspore tendenser i udtalen af fonemet [A:] i udvalgte regioner og aldersgrupper i taledatabasen SpeechDat 2. At projektet lod sig gennemføre, kan i høj grad tilskrives den store imødekommenhed Center for Personkommunikation udviste over for vores ide. Center for Personkommunikation stillede således undersøgel-

sesmaterialet SpeechDat 2 til vores fulde disposition og var desuden behjælpelig med implementeringen af de tekniske dele af analysen.

## **1. Materiale**

### **1.1. SpeechDat 2**

Taledatabasen SpeechDat 2 er blevet til på et initiativ fra EU og skal anvendes til at udvikle stemmestyrede servicefunktioner over det faste telefonnet.

Den danske version af SpeechDat 2 er den største taledatabase af sin art, da dens korpus er indtalt af 4000 talere der hver har bidraget med 43 ytringer. Registreringen af materialet foregik ved, at talerne ringede ind og oplæste ytringerne fra et fortrykt ark. Der forekom dog også spontane ytringer, da talerne f.eks. skulle oplyse deres geografiske placering. Bidragene blev elektronisk registreret og efterfølgende kontrolleret for at sikre en ensartet kvalitet i hele databasen. Under opbyggelsen af SpeechDat 2 har en række strenge kriterier været medvirkende til at sikre, at databasen fik nøjagtig det udseende og indhold, man ønskede. Det var således krav om ligelig kønsfordeling, ligesom antallet af medvirkende talere var i proportionelt forhold til befolkningstallet i de regioner, man ud fra dialektale hensyn havde inddelt landet i.

## **2 Metode**

### **2.1. Begrundelse for [A:]**

Når vi valgte kun at beskæftige os med et enkelt fonem, skyldtes det, at vi dels var nødt til at forholde os til grænserne for omfanget af et bachelorprojekt, dels at vi med det fonem, vi havde valgt at undersøge kunne have god grund til at tro, at det kunne reflektere eventuelle udtaleforskelle.

---

<sup>1</sup> In this paper,<sup>43</sup> we make use of the Danish phonetic alphabet Sampa.

Valget af [A:] var motiveret af den fluktuerende udtale a-lydene traditionelt har haft på dansk. Otto Jespersen var en af det første til at påpege en tendens til at den korte a-lyd blev fremskudt. Frem mod århundredskiftet blev denne fremskydning mere og mere udbredt for sluttelig at blive den dominerende udtale. Der var dog flere undtagelser hvor den lange a-lyd anvendtes foran alveolærlyde og nul for eksempel i ord som ”datter” og ”alter”.

Den typiske lange realisation af bogstavet *a* var i begyndelsen af det nittende århundrede den tilbagetrukne og åbne *a*-lyd. Denne realisation var dominerende indtil 1875, hvorefter tendensen blev, at lyden blev fremskudt og lukket. Imidlertid skete der også en modsatrettet udvikling, idet det tilbagetrukne og åbne [a:] hvor det efterfølges af et [R] smeltede sammen med dette og blev til [A:]. Resultatet af dette kan i dag spores i far, skare og spar.

Da den sproglige udvikling naturligvis sker i forskellige tempi og varierer fra landsdel til landsdel, var det vores antagelse at man ved en F'-analyse måtte kunne klarlægge eventuelle mønstre i variationen af udtalen af [A:].

(Brink, Lars, et.al., 1974)

## 2.2. Udvalgelse af undersøgelsesgrupper

Det første udvælgelseskriterium, vi blev stillet over for, var, om undersøgelsen skulle omfatte begge køn. Imidlertid valgte vi at koncentrere os om mandestemmerne, da den tekniske analyse viste, at disse ud fra en akustisk vurdering viste sig at være mere velegnede end kvindestemmerne.

Da de mandlige deltagere aldersmæssigt spredte sig fra børn til oldinge var vi nødsaget til endvidere at afgøre, hvorvidt vi ønskede at inddrage hele aldersspændet i vores undersøgelser. Vi endte med at lægge os fast på tre aldersgrupper, der byggede på den inddeling, der var foretaget i databasen. De tre grupper bestod af 16-30-årige, 31-45-årige og 46-60-årige. Årsagen til, at vi satte den aldersmæssige grænse ved 16 år, var at personer under 16 endnu ikke har fuldt

udviklede taleorganer, hvorfor den lydlige kvalitet kan være svingende. Endvidere satte vi den øvre grænse ved 60 år, da ældres taleorganer ændres hvorfor den lydlige kvalitet igen kan være varierende.

Det sidste kriterium, vi opstillede, var af geografisk og dermed dialektal karakter. I SpeechDat 2 er landet inddelt i elleve regioner, men af hensyn til det maksimale omfang af opgaven kunne en undersøgelse af alle disse ikke udføres. Vi begrænsede analysen til at omfatte tre af de regioner, der var afstukket i databasen. Det blev regionerne Nordjylland, Nordfyn og Nordsjælland, vi sluttelig lagde os fast på. At valget faldt på disse tre regioner skyldtes, at de indbyrdes fremstod så dialektalt forskellige, at de virkede velvalgte for at kunne besvare vores problemformulering. Da den ovenstående selektering var foretaget, nåede vi frem til et undersøgelsesmateriale med en sammensætning som vist i nedenstående skema.

	Nordfyn	Nordjylland	Nordsjælland
	Mænd	Mænd	Mænd
16 – 30 år	21	39	64
31 – 45 år	32	47	65
46 – 60 år	15	39	68
I alt	68	125	197

### 2.3. F'-analysen

Et uundværligt redskab til at gennemføre vores undersøgelser har været analysemetoden F'. Med F'-analysen har det været muligt at udregne formantværdierne og efterfølgende indsætte dem i et koordinatsystem. Vi har valgt at indsætte F1-værdien på den ene akse og et gennemsnit af F2 og F3-værdierne på den anden akse og på den måde få mulighed for at sammenligne to taleres udtale af et fonem. Til analysen er programmet Waves+ Product release 5,3, Entropic Inc. blevet anvendt.

## **2.4. Kontrol af analysemateriale**

Inden de udvalgte fonemer blev omsat til  $F'$ -plot, blev de kontrolleret auditivt såvel som visuelt. Kontrollen foregik ved, at vi gennemlyttede de sætninger, der senere skulle analyseres for at sikre, at lydsiden stemte overens med transskriptionen. Samtidig kunne vi frasortere sætninger med diegetisk støj eller andre elementer, der kunne virke forstyrrende på analysen.

Vi kontrollerede endvidere visuelt, at der ikke var sammenfald i formantkurverne, som blev estimeret for hvert enkelt fonem. Forekom der sammenfald, betød det, at programmet havde fejlestimeret formantværdierne, og den pågældende analyse blev kasseret.

## **3 Resultaterne af den teknisk analyse**

### **3.1. $F'$ -plottene**

De fremanalyserede  $F'$ -plot blev placeret i tre separate koordinatsystemer for at få et samlet billede af et eventuelt mønster i hver aldersgruppe. Herunder ses således det koordinatsystem, der viser  $F'$ -plottene for gruppe et, altså de 16 til 30-årige. De bogstaver der markerer plottene refererer til de udvalgte regioner. Bogstavet j betyder Nordjylland, f Nordfyn og z Nordsjælland. Som det fremgår af nedenstående koordinatsystem ligger plottene temmelig spredt. Plottene for  $F_1$  ligger således fra 350-950 Hz mens de for  $F_2'$  ligger fra 950-2250 Hz. Vi har dog afgrænset et kerneområde, der for  $F_1$  cirka ligger mellem 525 og 700 Hz og for  $F_2'$  mellem 1700 og 2000 Hz. Af pladshensyn har vi ikke her valgt at medtage de to resterende koordinatsystemer for gruppe to og tre, men vil naturligvis angive spredning såvel som kerneområder.

For gruppe to, der bestod af 31 til 45-årige, fordelte plottene sig fra 250 til 900 Hz på  $F_1$ -aksen og fra 1100 til 2200 Hz på  $F_2'$ -aksen. Kerneområdet for gruppe

to blev fra cirka 500 til 700 Hz på F1-aksen og fra cirka 1700 til 2000 Hz på F2'-aksen.

I gruppe tre, der udgjordes af de 46 til 60årige, var plottene fordelt mellem 275 og 900 Hz på F1-aksen og 950 og 2300 Hz på F2'-aksen. Kerneområdet var mellem cirka 600 og 800 Hz på F1-aksen og 1700 og 1900 Hz på F2'-aksen.

Når man sammenligner de tre gruppers kerneområderne, ses et vist sammenfald, hvilket dog tydeligst gør sig gældende for gruppe et og to. Ved at behandle de værdier, F'-plottene byggede på, efter en række statistiske metoder var det dog muligt at sammenligne værdierne på en mere tydelig og præcis måde, hvilket kan ses i det følgende.

### **3.2. Middelværdier, standardafvigelser og T-tests**

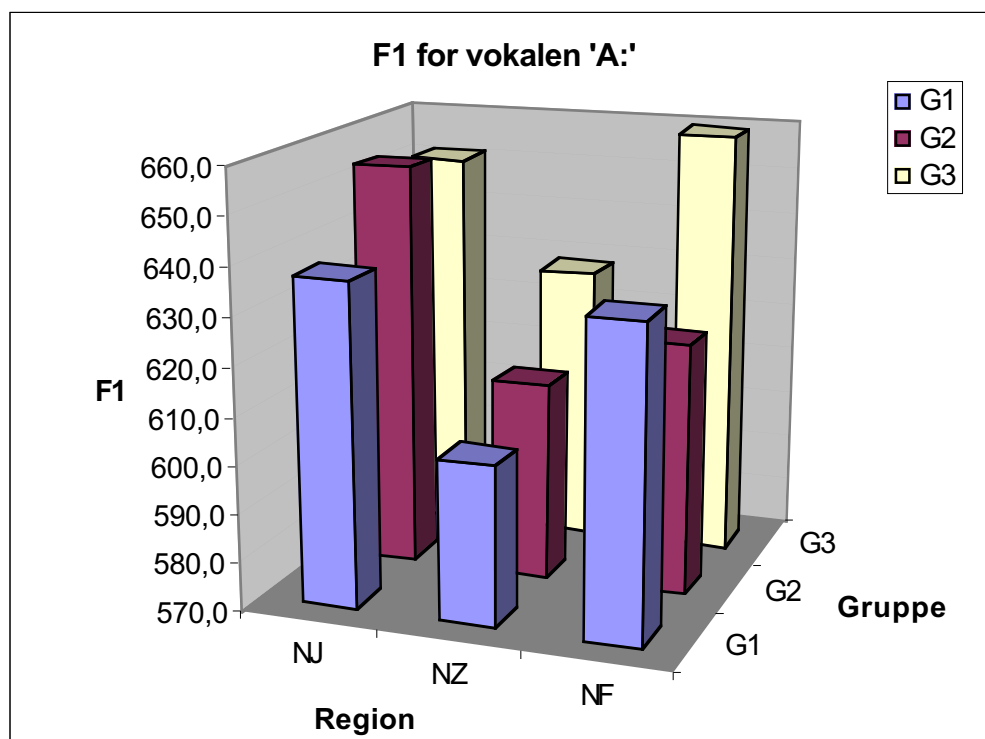
Middelværdien er udregnet ved at addere værdierne for samtlige forekomster på hver akse og dividere dem med antallet af forekomster. Herved har det været muligt at se tendenser i realisationerne i grupperne, men denne vurdering bør dog ikke foretages uden slutteligt at sammenholde middelværdierne med standardafvigelserne.

Standardafvigelsen er et udtryk for hvor stor spredningen er mellem de værdier, som middelværdien er udregnet på baggrund af. Dermed får man et udtryk for pålideligheden af det billede, middelværdien giver af realisationerne.

Den sidste metode, vi har benyttet os af for at vurdere resultaterne af F'-analysen, er en T-test; en beregning af sandsynligheden for sammenfald af formantværdierne i to grupper, hvorfor man også kan vurdere sandsynligheden for sammenfald i realisationen af [A:] i de forskellige regioner.

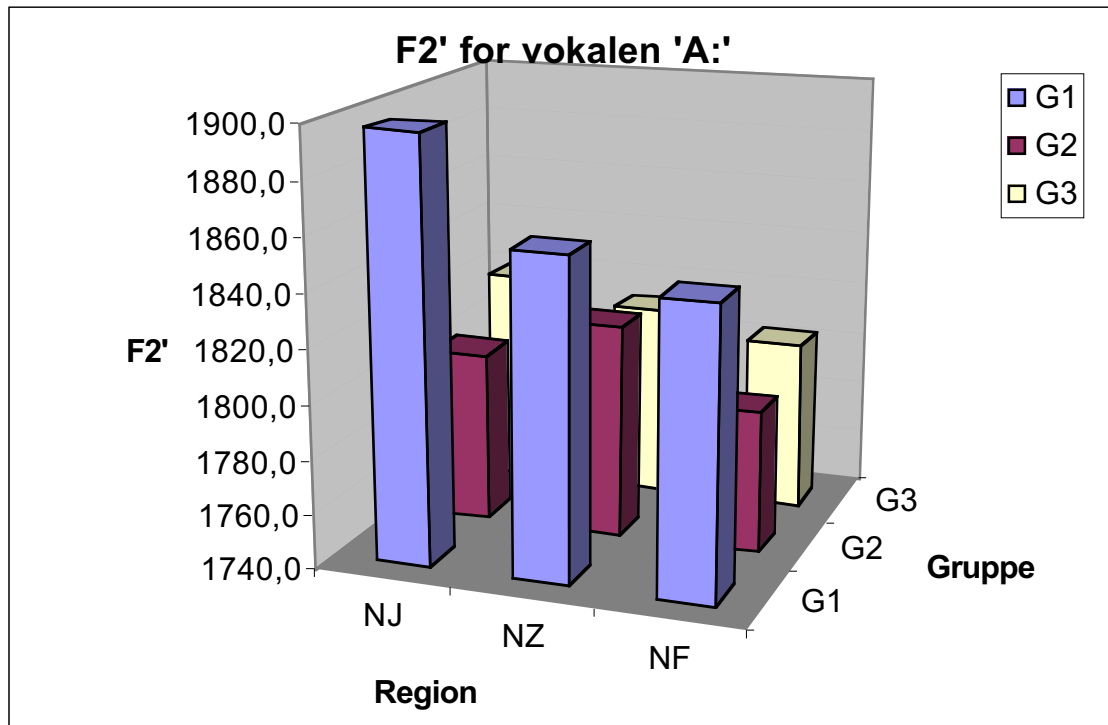
På de næste sider ses to tredimensionale søjlediagrammer, der viser middelværdierne for henholdsvis F1 og F2' på tværs af regionerne og aldersgrupperne. Skemaerne, der findes under søjlediagrammerne, indeholder standardafvigelserne for F1 og F2' i de tre regioner og aldersgrupper.

En analyse af realisationer af fonemet [A:] i udvalgte regioner og aldersgrupper



Standardafvigelse

F1	NJ	NZ	NF
G1	122,6	109,3	117,6
G2	121,1	129,8	90,5
G3	127,7	121,6	92,2



Standardafvigelse

F2'	NJ	NZ	NF
G1	119,2	150,4	148,3
G2	134,5	140,5	96,7
G3	181,8	129,3	110,1

T-test

		F1	F2'
G1	Z – F	29%	77%
	J – F	94%	18%
	J – Z	17%	20%
G2	Z – F	67%	34%
	J – F	21%	71%
	J – Z	8%	56%
G3	Z – F	35%	81%
	J – F	82%	73%
	J – Z	38%	78%



Ved at betragte diagrammerne med F'-plottene for alle tre aldersgrupper var det muligt at bestemme kerneområderne omtrentligt. Søjlediagrammerne over middelværdierne viser separat visse tendenser, selv om det er svært at påpege nogen tydelig relation mellem F1- og F2'-middelværdierne. Inden for F1-middelværdierne var der tre tendenser. Den tydeligste tendens var, at de nordsjællandske middelværdier var lavest i samtlige aldersgrupper. Endvidere havde Nordjylland generelt høje middelværdier. Denne tendens blev dog forstyrret af Nordfyns gruppe 3, der havde den højeste middelværdi i F1. Desuden lå middelværdierne for gruppe 3 generelt højere end de øvrige. Et billede, der kun blev ødelagt af Nordjyllands gruppe 2, hvis middelværdi lå få Hz over samme regions gruppe 3.

Standardafvigelserne for F1-værdierne er generelt høje, især de nordjyske samt Nordsjællands gruppe to og tre. Dette betyder ikke nødvendigvis, at de nævnte grupperes F1 generelt er lave på trods af den lave middelværdi. Omvendt er standardafvigelserne for Nordfyns gruppe 2 og 3 forholdsvis lave, hvilket betyder, at middelværdierne giver et relativt pålideligt billede af realisationerne.

For F2' kan der ligeledes påvises visse tendenser inden for materialet. Middelværdierne for gruppe 1 lå markant højere end de øvrige F2'-middelværdier, og der var endvidere tendens til, at middelværdierne for gruppe 2 og 3 ikke afveg ret meget fra hinanden. Standardafvigelserne for F2' var generelt højere end standardafvigelserne for F1. På trods af dette forhold bør man anse tendenserne i F2' for at være mere pålidelige end tendenserne for F1, da standardafvigelsen også bør ses i forhold til den numeriske værdi den refererer til.

Ved at sammenligne standardafvigelserne for F1 og F2' var det muligt at påpege en klar tendens i materialet. Nordfyns gruppe 2 og 3 udmærkede sig i denne sammenhæng ved at have de laveste standardafvigelser overhovedet i både F1 og F2'. Derfor må disse to grupper betragtes som de relativt mest ensartede med hensyn til realisationen af [A:].

I T-testene fik vi et billede af sandsynligheden for sammenfald i realisationerne. Det er dog svært at se et gennemgående mønster i sandsynligheden for sammenfald mellem regionerne, når man sammenholder alle tre aldersgrupper. Der kan alligevel spores enkelte tendenser i resultaterne. Imellem Nordjylland og Nordsjælland er der således generelt en lav sandsynlighed for sammenfald i alle tre grupper. Det skal dog bemærkes at sandsynligheden for sammenfald er relativt høj i F2' mellem Nordsjælland og Nordjylland i gruppe 3. Dette ligger i forlængelse af en anden tendens i T-testene, nemlig at sandsynligheden for sammenfald er jævnt stigende fra gruppe 1 til gruppe 3.

Man må på baggrund af den tekniske analyse konkludere, at de forskellige behandlinger af data og sammenligningerne, der er blevet foretaget på baggrund af disse, har gjort det muligt at påvise en række tendenser i materialet. Disse tendenser skal dog tages med forbehold, ligesom man må konstatere, at det ikke var muligt at se tydelige endsige entydige regionale eller aldersbestemte mønstre i materialet.

#### **4. Mulige fejlkilder**

Vi har i analysen konstateret visse tendenser, men disse, såvel som analysematerialet som sådan, bør dog vurderes ud fra visse forhold, der har kunnet tænkes at påvirke deres udseende. Som tidligere nævnt har vi i udgangspunktet forsøgt at eliminere eller mindske de faktorer, som materialet har kunnet være under indflydelse af.

Når resultatet alligevel blev så disparat som det er fremgået af analysen, bør man overveje, hvilke andre faktorer der har kunnet tænkes at have indvirket på resultatet.

Det ville umiddelbart være nærliggende at antage, at en del af spredningen kan tillægges fejl i analyseprogrammet. Gennem prøveanalyser af materialet har det dog vist sig at alvorlige fejl i søgealgoritmen tydeligt ses i F'-diagrammet.

Endvidere var det muligt at konstatere et vist fællestræk i grupperingen af plot for alle tre grupper samt at bestemme kerneområder for disse der desuden ligger i relativ nærhed af det estimat for [A:], som blandt andre Grønnum og Thorsen angiver.

En faktor i materialet, som ligeledes kan have været udslagsgivende for formantværdierne, er stød. Der er i transskriptionen af de i SpeechDat 2 optagne sætninger ikke taget højde for stød. Denne mangel blev først klar, efter analysen var foretaget, hvorfor det ikke var muligt efterfølgende at kompensere for dette. Det har derfor heller ikke været muligt at vurdere, hvor stor indvirkning dette forhold har haft på analyseresultaterne.

Derudover kan resultatet være blevet forstyrret af det forhold, at de medvirkende talere selv har skullet indplacere sig i forhold til deres egen vurdering af deres udtales regionale tilhørsforhold. Man kunne dermed forestille sig, at en del af talerne hverken har haft det fornødne kendskab til sine egne eller regionernes dialektale karakteristika.

Et andet forhold ved indsamling af data til SpeechDat 2, som må forventes at have præget udtalen, er, at hovedparten af de ord og sætninger som personerne har indtalt, var fortrykte på det ark, de fik tilsendt. Visse talere har derfor utvivlsomt forsøgt at tillempe en udtale, som tydeligvis ikke har været dem naturlig, men svarede til en opfattelse af rigsdansk.

En anden årsag til, at regionernes [A:]-realisationer fremtræder så uensartede, kunne være, at der i SpeechDat 2 ikke er taget forbehold for sociolektale variationer. Man kunne med andre ord forestille sig, at der i materialet findes realisationsmønstre for sociale grupper inden for eller på tværs af områderne, som kunne få regionen til at fremstå endnu mere uhomogen. Desuden må man formode, der inden for de respektive regioner kan findes dialektale realisationsvariationer, som ligeledes vil forstyrre det samlede billede for udtale. Dette kunne pege i ret-

ning af, at regionerne måske ikke i alle henseender var hensigtsmæssigt afgrænsede.

## 5. Konklusion

Vi har ved hjælp af den foregående analyse været i stand til at påvise visse tendenser i materialet. Når man betragter de udledte tendenser samlet, kan de i visse tilfælde underbygges af hinanden og i andre tilfælde forringes. I F1-middelværdierne påpegede vi tendenser til, at de nordsjællandske værdier lå lavere og de nordjyske generelt højere. Dette underbygges til dels af T-testene, der viser, at der netop mellem Nordjylland og Nordsjælland er en lav sandsynlighed for sammenfald i realisationerne. Man kunne derved umiddelbart forledes til at antage, at der mellem disse to regioner var et tydeligt mønster i realisationerne inden for disse. Ved at sammenholde resultatet med standardafvigelserne viser det sig, at den regionale forskel ikke kan siges at bygge på en klart regional realisationsforskel.

Indenfor F2'-middelværdierne påviste vi en aldersmæssig tendens i materialet. Gruppe 1 lå højest og gruppe 2 og 3 tæt på hinanden. Dette kunne umiddelbart tyde på, at man kunne efterspore aldersbestemte fællestræk i realisationen af [A:] for F2'. Denne antagelse kunne til en vis grad underbygges af T-testene, idet vi i disse påviste en tendens til at sandsynligheden for sammenfald var stigende i forhold til gruppens aldersniveau. Man kunne gå så langt som til at fremsætte en hypotese om, at når gruppe 1 har de højeste F2'-middelværdier, kunne det være et tegn på en aldersbestemt udvikling og ændring af realisationen af [A:] i F2', hvilket igen kunne underbygges af at de øvrige to gruppers middelværdier lå så tæt. Tendensen til en sådan udvikling må dog alene på baggrund af standardafvigelserne betegnes som, om ikke fejlagtig, så i hvert fald yderst vag.

Den eneste tendens i analysen der peger i retning af en vis grad af regional ensartethed i realisationen af [A:] er at Nordfyns gruppe 2 og 3 har de laveste standardafvigelser i både F1 og F2'. Denne tendens må dog igen sammenholdes med det faktum, at standardafvigelserne generelt er høje, hvorfor det i det hele taget er vanskeligt at tale om regional homogenitet i realisationen, når man ser på standardafvigelserne.

Ved at gennemføre analysen har vi dog kunnet besvare vores problemformulering, da det har været muligt at påvise regionale og aldersbestemte tendenser i materialet.

Resultaterne er dog som nævnt præget af en vis usikkerhed, hvorfor det ikke er muligt at konkludere, at der tegner sig et overbevisende billede af regionale eller aldersmæssige mønstre i realisationen af fonemet [A:]. Det ville dog være rimeligt at antage, at det havde været muligt at få et andet og måske tydeligere analyseresultat, hvis der havde været taget højde for de eventuelle fejlkilder og mangler, der er blevet anført ovenfor.

## 6. Litteratur

- Brink, Lars, et al. 1974. *Udtaleforskelle i Danmark*. København:Gjellerup
- Grønnum, Nina, Oluf Thorsen. 1991. *Fonetik for sprogstuderende*. København:Københavns Universitet, Institut for Almen og Anvendt Sprogvidenskab.
- Lindberg, Børge. 1998. SpeechDat Danish FDB 4000 Speaker Database for Fixed Telephone Network. *EU-Project SpeechDat, LE2-4001, Version 1.0*.