



Universiteit
Leiden
The Netherlands

**Telefoniske båndpasfiltres indflydelse på
talergenkendelse: høje frekvenser og dansk /t/
Hansen, R.P.**

Citation

Hansen, R. P. (2019). Telefoniske båndpasfiltres indflydelse på talergenkendelse: høje frekvenser og dansk /t/. *Nydanske Sprogstudier*, 54, 129-155. Retrieved from <https://hdl.handle.net/1887/70786>

Version: Not Applicable (or Unknown)
License: [Leiden University Non-exclusive license](#)
Downloaded from: <https://hdl.handle.net/1887/70786>

Note: To cite this publication please use the final published version (if applicable).

Telefoniske båndpasfiltres indflydelse på talergenkendelse: høje frekvenser og dansk /t/

RASMUS PUGGAARD

ABSTRACT

Denne artikel er en undersøgelse af hvordan telefoniske båndpasfiltre påvirker talesignalet med mulige implikationer for retsfonetikken. Der fokuseres særligt på tre forskellige filtre: Det filter der anvendes i fastnetforbindelser, AMR's filter, der benyttes i det meste mobiltelefoni, og som ved dårlig dækning leverer et dårligere signal end fastnetforbindelser, og VoLTEs filter, der giver det mest moderne telefonsignal, og som benyttes i nyere smartphones. Artiklen viser hvordan variation i udtalen af dansk /t/ kan bruges til at identificere en specifik taler, f.eks. i retsfonetisk sammenhæng. Tre forskellige kilder til /t/-variation præsenteres: dialektal variation, variation i udtalen af fremmedsprog og idiolektal variation blandt talere af rigsdansk. For alle disse variationskilder vises det hvordan de forskellige telefoniske båndpasfiltre påvirker talesignalet, og særligt hvordan fastnet- og AMR-forbindelser udviser meget af variationen fordi disse forbindelser ikke transmitterer særligt høje frekvenser. De parametre som en retsfonetiker vil kigge på i en optagelse af høj kvalitet, vil ofte ikke være tilgængelige i en optaget telefonsamtale. Artiklen foreslår derfor nogle andre parametre i talesignalet som kan være værd at kigge på hvis man leder efter variation i /t/-udtale i en telefonoptagelse.

EMNEORD: retsfonetik, talergenkendelse, telefoni, sociofonetik, accent

1 INTRODUKTION

Når talesignalet transmitteres telefonisk, gennemgår det en vis komprimering vha. det der kaldes telefoniske båndpasfiltre. Komprimeringen foregår på den måde at de laveste og højeste frekvenser af talesignalet skæres væk. Det varierer meget på tværs af de nuværende telefoniske systemer hvor hård komprimeringen er, men det er velkendt at den akustiske information fra talesignalet påvirkes af båndpasfiltrene. Det

er generelt mere velbeskrevet hvordan talesignalet påvirkes af at lavere frekvenser skæres væk end højere frekvenser. Men denne artikel vil beskrive nogle konsekvenser af bortfiltrering af høje frekvenser, for de høje frekvenser rammes temmelig hårdt af telefoniske båndpasfiltre, på trods af at der er meget information på de højere frekvenser der er vigtig for menneskelig tale. Mennesker er nemlig meget mere følsomme overfor lyd ved relativt lave frekvenser end ved højere frekvenser. Et almindeligt hørende barn kan høre lyde der strækker sig fra 20 til 20.000 Hz, men følsomheden overfor forskelle i det akustiske signal er bedst omkring 4.000 Hz (Howard & Angus 2009: 89). Det ”koster” lige meget at transmittere lyd fra f.eks. 3.000-4.000 Hz som fra 8.000-9.000 Hz, men det menneskelige øre er bedre til at opfatte signaler fra 3.000-4.000 Hz.

Den akustiske information der relaterer sig til vokaler, findes i reglen på lave frekvenser, og det er båndpasfiltrenes påvirkning på disse der er bedst beskrevet i den videnskabelige litteratur. Der er skrevet forholdsvist lidt om påvirkningen på konsonanter. Denne artikel er et lille forsøg på at lave om på det idet den fokuserer på en enkelt konsonant, nemlig det danske /t/. Nærmere bestemt ser artiklen på hvordan /t/ er påvirket af telefoniske båndpasfiltre, og hvad dette kan indebære for retsfonetik. Konsonanten /t/ er velegnet til dette idet den på rigsdansk realiseres som affrikeret, hvilket leder til en høj grad af akustisk aktivitet på netop de høje frekvenser. Denne aktivitet ligner til dels den man ser i /s/. Udover de akustiske aspekter er /t/ interessant af sociolingvistiske og sproglæringsmæssige årsager: Den er en kendt kilde til inter- og intradialektal variation (se bl.a. Nielsen 1984: 149, Espegaard 1996: 262, Petersen 2009 og Puggaard 2017, under udg. a); Grønnum (2009: 51) og Puggaard (2015, under udg. b) skriver at /t/ varierer betragteligt blandt talere af rigsdansk; og førstesprogstalere af rigsdansk har en tendens til at overføre den affrikerede variant til andre sprog hvilket betyder at konsonanten er en god indikator for en dansk accent (Puggaard 2015, under udg. b).

Denne artikel viser at telefoniske båndpasfiltre fjerner megen vigtig akustisk information ved /t/, uanset filtrets type, hvilket vil sløre eller formindske den variation i det akustiske signal som en eventuel talergenkendelse ville skulle basere sig på. Sommetider er der nok in-

formation tilbage på de højere frekvenser til at variation stadig er hørbart, sommetider ikke. En retsfonetiker der er på udkig efter variation i /t/-udtale, vil være nødt til at overveje andre parametre i en telefonoptagelse end i en optagelse af høj kvalitet. Nogle sådanne parametre foreslås i artiklen.

Artiklen indeholder i andet afsnit en kort gennemgang af den relevante videnskabelige litteratur om talergenkendelse og en mere grundig gennemgang af de telefoniske båndpasfiltre der er i omløb. I afsnit 3 gives der en introduktion til /t/ og andre danske obstruenter der kan påvirkes af filtrene. I fjerde afsnit præsenteres tre akustiske pilotstudier af båndpasfiltres indflydelse på /t/ i forskellige kontekster, og i femte afsnit diskuteres resultaterne inklusive de fejlkilder der måtte være, og inklusive den indflydelse de kan have for retsfonetikken.

2 RETSLINGVISTISK TALERGENKENDELSE

2.1 Talergenkendelse generelt

Den nok hyppigste anvendelse af lingvistik i retsøjemed er i sager der involver identifikation af en taler på baggrund af en optagelse (Foulkes og Barron 2000). Sagerne er primært af to slags: 1) naiv talergenkendelse, hvor et vidne skal identificere en taler ud fra hukommelsen (Nolan 1983, 1997), og 2) teknisk talergenkendelse, hvor en lingvist har adgang til optagelser af tale. Nærværende studie berører primært teknisk talergenkendelse, men det er potentielt relevant for begge slags.

Naiv talergenkendelse i retsøjemed udføres typisk vha. ørevidnekonfrontationer, også kendt som stemmeparader. Disse minder om øjenvidnekonfrontationer og bruges når vidnet ikke har set gerningsmanden (se f.eks. Broeders og van Amelsmoort 1999).

Teknisk talergenkendelse blev notorisk kendt i den sidste halvdel af det tyvende århundrede da International Association of Voice Identification påstod at de kunne identificere talere ved at studere deres 'stemmeaftryk' (egentlig blot spektrogrammer), som de påstod var lige så unikke som fingeraftryk (Hollien 1990: 224). Der var ikke meget empirisk evidens der understøttede denne kontroversielle påstand, som sidenhen da også er blevet anfægtet af mange forskere (f.eks. Bolt et al. 1970, Hollien 1974, 1990: 210ff., Nolan 1983: 198ff. og Meuwly 2003),

og som der i dag er klar konsensus imod blandt retslingvister (Jessen 2008: 706).

Selvom spektrogrammer ikke kan bruges til at identificere en specifik taler, er der ingen tvivl om at talesignalet bærer information om dets taler, hvilket tydeligt ses ved at lægfolk generelt er ret gode til at identificere bekendte på baggrund af deres stemmer (f.eks. Foulkes og Barron 2000). En stemmes særlige kendetegn kan både skyldes indlærte vaner og individuelt varierende anatomi, og vores viden om forholdet mellem indlært og anatomisk variation er stadig begrænset. Mennesker er i stand til en enorm grad af bevidst og ubevidst variation af deres stemme, så den fysiske forskel på talere og hvordan denne påvirker talergenkendelse, kan kun forklare en relativt lille del af den akustiske forskel talere imellem. På trods af dette er der dog forskellige aspekter af talen der har vist sig brugbare til taleridentifikation da de er forholdsvis stabile for den enkelte taler. Disse inkluderer f.eks. gennemsnitlig grundtone (F0) og gennemsnitlig frekvens af anden formant (F2) i den lave ende af det akustiske spektrum (Nolan 1997). Nasale konsonanter har også vist sig at være brugbare til talergenkendelse, eftersom mennesker er mindre i stand til bevidst at modificere næsehulen end resten af vokalapparatet (f.eks. Wolf 1972). Det er også påvist af Künzel (1997) at talehastighed er rimeligt konstant for den enkelte taler. Generelt er det påvist at talere der falder udenfor normalen på et givent parameter, er lettere at identificere (Foulkes og Barron 2000). Det vises nedenfor hvordan variation i /t/-udtale også er værd at kigge på.

2.2 Telefoners indflydelse på talesignalet

Telefoner er meget ofte involverede i talergenkendelse; mere end 95 % af sager med ukendte taleprøver i Tyskland var i slut-90'erne optaget over telefonen (Künzel 1997), og dette tal ses ofte som repræsentativt. Talesignalet ændres eller forværres på flere måder når det transmitteres gennem en telefon, f.eks. gennem 1) telefonstemmen, 2) aktiv støjreduktion og 3) telefoniske båndpasfiltre.

Telefonstemmen er et særligt register der benyttes af talere når de taler i telefon. Registret associeres med formel stil og lav talehastighed (Wells 1982: 28) og en stigning i F0 som resultat af højere tale (Hirson et al. 1995). Det er også blevet påvist at telefonstemmen resulterer i en

stigning i første formant (Künzel 2001; Byrne og Foulkes 2004). Konsonanter kan også blive påvirket, hvilket bliver diskuteret yderligere i afsnit 5.

Aktive støjreduktionssystemer i mobiltelefoner påvirker højst sandsynligt højere frekvenser i talesignalet. De kan til gengæld ikke simuleres da detaljerne om hvordan de virker, er beskyttet information, og forskellige mobiltelefonproducenter har formentlig forskellige metoder til at udføre aktiv støjreduktion. Problemet bliver derfor ikke yderligere behandlet i denne artikel.

Denne artikel behandler telefoniske båndpasfiltres indflydelse på talesignalet. En ting som højst sandsynligt vil ændre sig hurtigt i de kommende år, er hvilken kvalitet der kan forventes af en telefonoptagelse. Indtil for relativt få år siden blev langt størstedelen af telefonopkald udført ved brug af fastnettet. Pga. det båndpasfilter der benyttes på fastnettet, udsendes talesignalet i en frekvensskala på mellem ca. 300 og 3.400 Hz (Foulkes og Barron 2000), hvorved der fjernes vigtige dele af talesignalet både i de lave og de øvre frekvenser. Nogle af disse er velbeskrevne i den retslingvistiske litteratur, og andre kan udtrækkes af fonetiklitteraturen. Künzel (2001) viser at første formant (F1) for alle vokaler udover /a/ på tysk påvirkes af den nedre frekvensgrænse. De fleste voksne taleres F0 vil også til en vis grad blive påvirket (Titze 1994: 170). Derudover påvirkes tredje og fjerde formant (F3, F4) også af den øvre frekvensgrænse, og information fra disse formanter er meget vigtig for talergenkendelse (Nolan 1983). Den øvre grænse kan også antages at neutralisere meget af den akustiske forskel mellem ustemte frikativer, hvor den primære akustiske aktivitet kan ligge et langt stykke over 3.400 Hz (f.eks. Gordon et al. 2002).

På trods af at mobiltelefoner overvejende har erstattet fastnettelefoner, er der udført relativt lidt forskning i hvordan mobiltelefonsignaler – særligt moderne signaler – påvirker talergenkendelse. Det akustiske signal som sendes og modtages af mobiltelefoner, går igennem en digital komprimeringsproces. Denne komprimeringsproces er af forskellig art, afhængigt af hvilken type mobiltelefon og hvilken type dækning der er tale om. Komprimeringsmetoderne kaldes kodekser, og de benytter hver især forskellige filtre. De filtre der ud over filtret i fastnetforbindelser er relevante i denne artikel, er filtrene der indgår i AMR's

smalbåndskodeks og VoLTEs bredbåndskodeks. I det følgende beskriver jeg først disse to filtre nærmere, og derpå illustrerer jeg hvilken effekt filtrene i de tre typer forbindelser har på fem forskellige frikativer. Flere nylige studier i talergenkendelse tager stadig udgangspunkt i fastnetsignalet (f.eks. McDougall et al. 2015) fordi fastnettet altid benytter samme filter, mens de fleste mobiltelefoner stadig primært benytter det temmelig ustabile AMR-smalbåndskodeks, som har forskellige typer filtre alt afhængig af dækningen. Det er dog problematisk at basere moderne forskning på fastnetsignalet da det overvejende er blevet erstattet af mobiltelefoni.

AMR-smalbåndskodekset blev standardiseret til det tidlige Global System Mobile Communication mobilnetværk (GSM) og blev brugt hele vejen op til tredjegerationsnetværket (3G; Guillemin og Watson 2006). Filtret som benyttes af AMR, filtrerer grundlæggende de samme frekvenser væk som fastnetfiltret, men bevarer lidt flere af de nedre frekvenser (Bessette et al. 2002). Til gengæld er de trådløse forbindelser som benytter sig af AMR, markant langsommere end fastnetforbindelser, hvilket kan påvirke lyd kvaliteten på uforudsigelige måder; signaler kan ændre bithastighed hvert 40. msec, og ved lave bithastigheder kan den øvre frekvensgrænse falde helt ned til 2.800 Hz (Guillemin og Watson 2008; Öhman et al. 2010). Alt i alt vil mobiltelefonssignaler som benytter AMR, derfor ofte have lavere kvalitet end fastnetsignaler, og talergenkendelse vil være negativt påvirket i endnu højere grad end ved fastnetsignaler (Byrne og Foulkes 2004; Guillemin og Watson 2008). Når der senere i denne artikel refereres til AMR's filter, menes der den hårdest mulige komprimering med den smalleste båndbredde, dvs. 400-2.800 Hz.

Voice over LTE (VoLTE) er en nylig udvikling i telefoni der højst sandsynligt inden for de næste få år vil kunne afhjælpe problemet med den lavere kvalitet. VoLTE er navnet på den komprimeringsmetode der benyttes af 3G's afløser, Long Term Evolution (LTE eller 4G; Sauter 2014). Filtret der benyttes af VoLTE, hedder AMR-WB, men i det følgende vil jeg for at undgå forveksling referere til det som 'VoLTEs filter'. VoLTEs filter dækker en meget bredere båndbredde – dvs. en større frekvensskala – end AMR's filter idet det spænder over 50-7.000 Hz (Nguyen et al. 2016). Denne brede båndbredde kan løse mange af

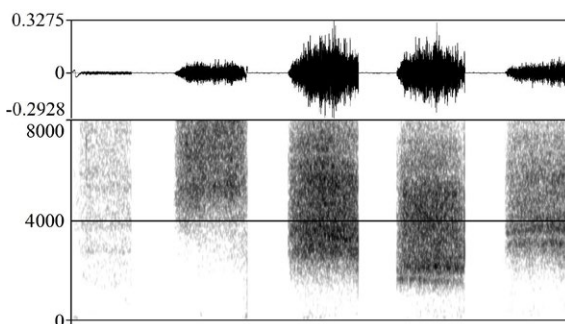
de ovennævnte problemer: Det udvidede signal på de nedre frekvenser vil resultere i at F0 og F1 ikke længere er påvirkede af filtret. Det udvidede signal på de øvre frekvenser vil resultere i at heller ikke F3 og F4 længere er påvirkede, og på trods af at de højeste frekvenser i frikativt generelt ligger over 7.000 Hz (Stevens 1960), ligger den primære aktivitet for alle de ustemte frikativt målt af Gordon et al. (2002) på under 6.000 Hz, så der vil som regel være nok brugbar information under 7.000 Hz til at alle fonemiske distinktioner bevares (Shadle og Mair 1996). Dette betyder at VoLTE vil være i stand til at inkludere det meste eller al fonemisk taleinformation, men at noget talerspecifikt og sociofonetisk information i den høje ende af spektret stadig vil være tabt. For eksempel rapporterer Pharao et al. (2014: 9) at en variant af /s/ [s+], som på dansk associeres med homoseksuelle mænd, har sin primære akustiske aktivitet ved omkring 9.000 Hz. Fonemisk information kan dog også gå tabt på sprog med flere koronale frikativt, hvor den primære forskel kan ligge over 7.000 Hz, men selv her vil der som regel være nok brugbar information under 7.000 Hz til at de fonemiske distinktioner bevares (Stevens 1960; Shadle og Mair 1996).¹

2.3 Demonstration af båndpasfiltrenes indflydelse

De følgende figurer demonstrerer hvordan forskellige filteres øvre frekvensgrænser kan påvirke talesignalet. Bemærk at figurerne ikke viser optagelser fra egentlige telefonsignaler. Effekten af de telefoniske båndpasfiltre er i stedet simuleret vha. metoden beskrevet i afsnit 4.1.

1 Efterhånden som VoLTE-teknologien udbredes i mobildækningen, vil lyd kvaliteten generelt stige for telefonopkald, men på nuværende tidspunkt er den stadig utilgængelig i det meste af Danmark. De fleste nyere smartphones har VoLTE-support, men kun ét dansk teleselskab, Telenor, tilbyder VoLTE-dækning i øjeblikket, og kun i dele af landet hvor selskabet har tilstrækkelig LTE-dækning, jf. <https://www.telenor.dk/kundeservice/drift-og-dakning/dakning/om-dakningen/>.

FIGUR 1. OSCILLOGRAM OG SPEKTROGRAM AF DE KORONALE FRIKATIVER [θ s ʃ ʒ ç] I HENHOLD TIL IPA, INTET FILTER

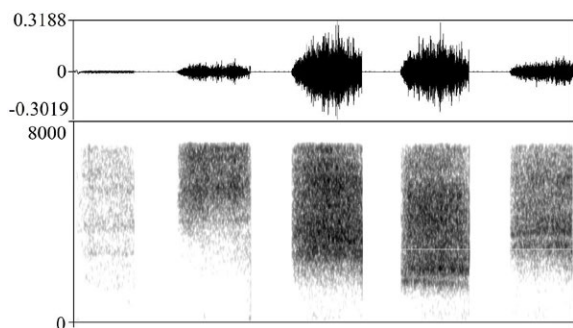


Figur 1 viser det fonetiske alfabets (IPA) ustemte koronale frikativer [θ s ʃ ʒ ç] uden for kontekst og uden noget filter². Det er nogenlunde ligetil at skelne mellem de forskellige lyde på baggrund af spektrogrammerne:

- [θ] har lav intensitetsstøj spredt i hele spektret, jf. den svage sværtning der ses i figuren.
- [s] har høj intensitetsstøj fra ca. 4.500 Hz og opefter som topper ved ca. 5.300 Hz, jf. sværtningen øverst i figuren.
- [ʃ] har høj intensitetsstøj fra ca. 2.500-6.500 Hz som topper ved ca. 3.600 Hz, jf. sværtningen som er mest prominent midt i figuren.
- [ʒ] har høj intensitetsstøj fra ca. 1.500-5.500 Hz som topper ved ca. 2.100 Hz, jf. sværtningen som er mest prominent nederst i figuren.
- [ç] har et smallere bånd af høj intensitetsstøj fra ca. 3.000-4.500 Hz som topper ved ca. 3.700 Hz. Støjen er generelt mindre intens end i de øvrige koronale frikativer.

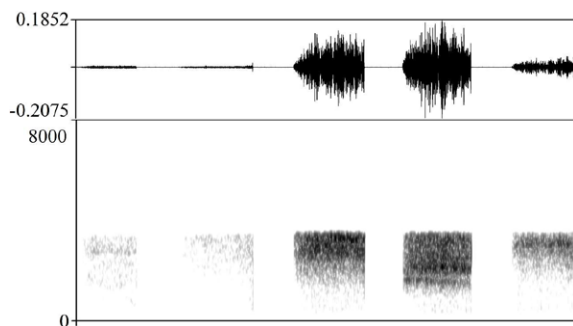
² Optagelserne som benyttes her, er indtalt af Peter Isotale, og de kan tilgås på <http://www.ipachart.com/>.

FIGUR 2. OSCILLOGRAM OG SPEKTROGRAM AF DE KORONALE FRIKATIVER [θ s f ʃ ç] I HENHOLD TIL IPA, VOLTES FILTER



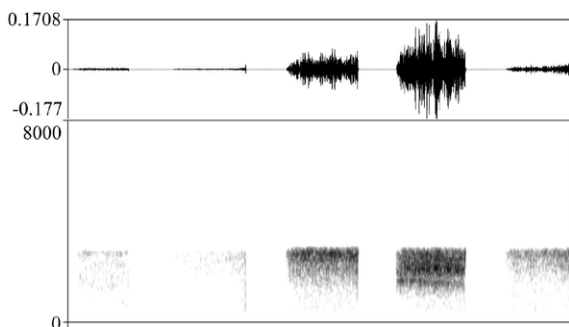
Figur 2 viser de samme lyde som i figur 1 efter simulering af båndpasfiltret benyttet af VoLTE (50-7.000 Hz). Det er stadig ligetil at skelne mellem lydene, og ingen fonemisk information er gået tabt.

FIGUR 3. OSCILLOGRAM OG SPEKTROGRAM AF DE KORONALE FRIKATIVER [θ s f ʃ ç] I HENHOLD TIL IPA, FASTNETTETS FILTER



Figur 3 viser de samme lyde efter simulering af båndpasfiltret benyttet i fastnetopkald (300-3.400 Hz). Det er stadig ligetil at skelne mellem [ʃ ʒ ç], men meget akustisk information er gået tabt. Forskellen mellem [θ s] er de facto neutraliseret; den eneste forskel mellem dem er at den eksisterende støj i [θ] er en anelse mere intens, hvilket kan ses på oscillogrammet. Dette er dog ikke tydeligt hørbart.

FIGUR 4. OSCILLOGRAM OG SPEKTROGRAM AF DE KORONALE FRIKATIVER [θ s ʃ s̺] I HENHOLD TIL IPA, AMR'S FILTER



Figur 4 viser de samme lyde efter simulering af den lavest mulige lyd-kvalitet i AMR-kodekst (400-2.800 Hz). Det er stadig muligt at skelne mellem [ʃ s̺], men det meste af den akustiske information i [ʃ] er tabt, sammenlignet med figur 2. Forskellen mellem [θ s̺] er næsten komplet neutraliseret; lidt af frikationsstøjen i [s̺] er intakt, men det er næsten ikke hørbart, og forskellen i intensitet mellem [θ s̺] er endnu mindre end i figur 3 foroven.

3 EN KORT INTRODUKTION TIL DANSKE OBSTRUENTER

3.1 Danske obstruenter

Alt afhængigt af ens analyse er der otte eller ni fonemiske obstruenter på dansk /b d g p t k f s (ç)/³, fonetisk [b̥ d̥ ɡ̥ p^h t^s k^h f s (ç)] i simpelt onset, dvs. stavelsesinitialt og udenfor konsonantklynger (Basbøll 2005: 60ff.; Grønnum 2009: 39). Forskellen mellem de to sæt plosiver regnes for at være en forskel i aspiration (Grønnum: 2009: 107). I Lisker og Abramsons (1964) terminologi kan man sige at /b d g/ har kort opløsning, mens /p t k/ har lang opløsning. I medial position udtales alle plosiver i reglen som mildt stemte (Fischer-Jørgensen og Hutter 1981). Ifølge Grønnums analyse (2009: 318ff.) realiseres /p t k/ som uaspirerede stavelsesfinalt, mens /b d g/ realiseres som halv-

3 /ç/ analyseres som et fonem af Basbøll (2005: 61), men analyseres som en fonologisk sekvens af /sj/ af Grønnum (2009: 306).

vokaler i denne position. Den følgende diskussion vil kun dække fonemernes varianter i simpelt onset.

De akustiske karakteristika ved /b d g/ burde ikke blive påvirket af de smalle båndpasfiltres øvre frekvensgrænser. I forhold til /p k/ er det ikke usandsynligt at disse plosivers opløsningsfaser indeholder værdifuld information over de smalle båndpasfiltres øvre frekvensgrænser – dette kunne meget vel være tilfældet for /k/, som kan have en vis grad af velær affrikation eller palatal opløsning.

/f s ɸ/ vil derimod i høj grad påvirkes af smalle filtre, og /s ɸ/ vil sandsynligvis også påvirkes af brede filtre. Meget af den relevante akustiske aktivitet vil formentlig bevares for /f/, og det er sandsynligt at nok aktivitet også bevares for /ɸ/ til at identifikation er ligetil. Det er til gengæld sandsynligt at det meste af den akustiske aktivitet i /s/ filtreres væk. Der er mange indikationer på at /s/ varierer blandt talere; bl.a. viser Gordon et al. (2002) at der er stor forskel fra sprog til sprog og fra taler til taler på hvor den primære akustiske aktivitet i /s/ er centreret. Al denne variation foregår over den øvre frekvensgrænse af fastnetfiltret. Grønnum (2009: 306) skriver også at der er sociolektal variation i udtalen af /ɸ/, og meget af denne variation ligger antageligt også over den øvre frekvensgrænse der filtreres bort i smalle båndpasfiltre.

/t/ vil sandsynligvis også blive påvirket af både smalle og brede filtre. /t/ er fonologisk kategoriseret som et aspireret stop på dansk, men på rigsdansk og i mange dialekter er det realiseret som affrikeret [tʰ]. Basbøll (2005: 60) angiver den fine transskription [tʰs̺] for rigsdansk /t/, og angiver den altså som en fonetisk affrikat. I mit arbejde med /t/-udtaler har jeg dog fundet at affrikationen efter pustet har en primær koncentration af energi ved lavere frekvenser end /s/. Derudover er affrikationen kortere end hvad ville forventes for /s/, og medmindre /t/ forekommer foran høje vokaler, efterfølges affrikationen som regel af et pust af almindelig aspiration – dvs. glottal friktion – før onset af den følgende vokal. Ikke desto mindre er opløsningen af /t/ på rigsdansk overvejende defineret af akustisk energi over 3.400 Hz, og der er også ofte relevant information over 7.000 Hz, så /t/ vil helt sikkert blive påvirket af båndpasfiltrene.⁴

4 Sekvensen /tj/ analyseres i reglen ikke som et fonem, men realiseres normalt som et affrikat [tʃ] (Grønnum 2009: 148). Et studie af hvordan /tj/ påvirkes af smalle båndpasfiltre, særligt

Det er sandsynligt at både smalle og brede båndpasfilteres øvre frekvensgrænser påvirker flere af de danske obstruenter med en mulig relevans for talergenkendelse. Disse er alle værd at studere, men da denne artikel udelukkende fokuserer på /t/, vil der i det næste afsnit kun blive introduceret nogle af de faktorer der gør /t/ særligt relevant når det kommer til talergenkendelse.

3.2 /t/ i danske dialekter

En fonetisk variant af /t/, kendt i folkemunde som det 'tørre t', nævnes af og til i litteraturen (f.eks. Nielsen 1984: 149 og Petersen 2009). Varianten er generelt forbundet med nordjysk dialekt og associeres med en udtale der er længere fremme i munden, mangler affrikation og har en ret svag aspiration. Der er ingen detaljerede beskrivelser af /t/'s artikulation i dialektlitteraturen, men i et studie over den geografiske spredning af denne variant har jeg fundet at ikkeaffrikeret /t/ er meget mere udbredt i jyske dialekter end på rigsdansk, og at det ikke begrænses til nordjyske dialekter (Puggaard 2017, under udg. a). Jeg fandt at der i høj grad er en forbindelse mellem /t/-udtale og dialektområde, og at der lader til at være et kontinuum fra nord mod syd mht. grad af aspiration og affrikeret /t/-opløsning. Denne aspiration og affrikation kan ikke nødvendigvis relateres til variation i Voice Onset Time (VOT): Både talere af nordjysk og sønderjysk har relativt kort gennemsnits-VOT i /t/-udtale på ca. 40 ms, mens talere af midtjysk har en længere gennemsnits-VOT i /t/-udtale på ca. 60 ms⁵. Variationen kan snarere relateres til hvor markeret affrikeret udtale er på de forskellige dialekter; på nordjysk vil et langt /t/ med VOT på 70 ms i en trykstærkt stavelse højst sandsynligt ikke være affrikeret. På sønderjysk vil et umarkeret /t/ på 50 ms i en trykssvagt stavelse derimod højst sandsynligt være affrikeret. Affrikeret udtale af /t/ er altså ikke

med reference til udtalen af /t/, ville være meget interessant, men er antageligt ikke direkte relevant for talergenkendelse.

5 Dette ligger stadig markant under gennemsnittet for /t/ på rigsdansk, som er blevet målt til ca. 85 ms (Mortensen og Tøndering 2013), 93 ms (Puggaard 2015, under udg. b), og sågar 140 ms (Garibaldi og Bohn 2015). Den store forskel skyldes antageligt forskellige typer data, hvor Mortensen og Tønderings data er mest sammenligneligt med dialektoptagelserne benyttet af Puggaard (2017, under udg. a).

en effekt af lang VOT på jyske dialekter, men nærmere et resultat af emfase, tryk og talerens dialektområde.

I Puggaard (2017, under udg. a) fandt jeg desuden en tendens til intradialektal variation i /t/-udtale. Mens nogle talere af midt-østjysk meget ofte brugte det tørre *t*, brugte andre det næsten aldrig. Der blev i høj grad fundet dialekt-specifik /t/-udtale, men udtalen var stadig i nogen grad idiosynkratisk indenfor dialekterne. Selvom de traditionelle dialekter ikke er specielt udbredte længere (Kristiansen 1998, 2003), er det stadig mit indtryk at det tørre *t* er et vigtigt nordjysk dialektræk, og at der er en høj grad af inter- og intradialektal variation i /t/-udtale blandt talere af jysk. Dette betyder at /t/-udtale bærer information der potentielt kan have betydning for udpegningen af en talers dialektale tilhørsforhold og for identifikation af talere indenfor kendte regioner. Selvom dette kun er bekræftet for talere af jysk, må det forventes at en lignende grad af talerspecifik /t/-udtale kan findes i resten af Danmark også (jf. afsnit 4.4). Det er meget sandsynligt at den akustiske information ved /t/-udtale påvirkes af smalle båndpasfiltre, hvorfor muligheden for taleridentifikation besværliggøres.

3.3 /t/ i accenter

Der er en tendens til at den høje grad af /t/-affrikation udvides til udtalen af fremmedsprog blandt talere af rigsdansk. Puggaard (2015, under udg. b) undersøger både denne effekt blandt danskeres udtale af kinesisk og i et mindre pilotstudie også danskeres udtale af engelsk. Affrikation i /t/-udtale blandt førstesprogstalere af engelsk er ikke velbeskrevet. Det forekommer i en vis grad, men ikke i samme grad som på dansk. VOT på engelsk /t/ er desuden kortere end på dansk; en gennemsnitlig varighed på 63 ms rapporteres af Docherty (1992: 116). I pilotstudiet fandt jeg at selv meget kyndige danske engelsktalere har affrikeret /t/-udtale i næsten lige så høj grad som når de taler dansk. Graden af affrikation i /t/-udtale kan altså i høj grad bruges til at identificere en dansk accent, og det er sandsynligt at denne information vil forsvinde eller i hvert fald blive påvirket af telefoniske båndpasfiltre.

4 UNDERSØGELSE: BÅNDPASFILTRES EFFEKT PÅ /t/

Ovenfor har jeg fremstillet grundlaget for tre testbare hypoteser: 1) at jysk variation i /t/-udtale i høj grad vil påvirkes af telefoniske båndpasfiltre, 2) at dansk udtale af /t/ i fremmedsprog også vil blive påvirket, og 3) at talerspecifik udtale af /t/ på rigsdansk også vil blive påvirket. De tre hypoteser testes nedenfor.

4.1 Metode

Lydoptagelserne som vil blive analyseret i det følgende, er blevet manipuleret vha. softwaren Praat (Boersma og Weenink 2016). Stavelser med /t/ i simpelt onsets er isolerede, og vha. den indbyggede Hann-båndpasfilterfunktion bliver de tre telefoniske filtre simulerede: VoLTEs filter (50-7.000 Hz), fastnettets filter (300-3.400 Hz) og den hårdest mulige komprimering gennem AMR-kodekset (400-2.800 Hz), som præsenteret i afsnit 2.2. Frekvensudjævning er sat til 100 Hz. Stavelserne er dernæst visuelt og auditorisk inspiceret i Praat.

4.2 Case 1: jyske varianter af /t/

Som diskuteret i afsnit 3.2 er der en både idiolektal og dialektal distinktion mellem [t^h] og [t^s] blandt talere af jyske dialekter. I det følgende bliver det testet hvorvidt denne distinktion påvirkes af de forskellige telefoniske båndpasfiltre.

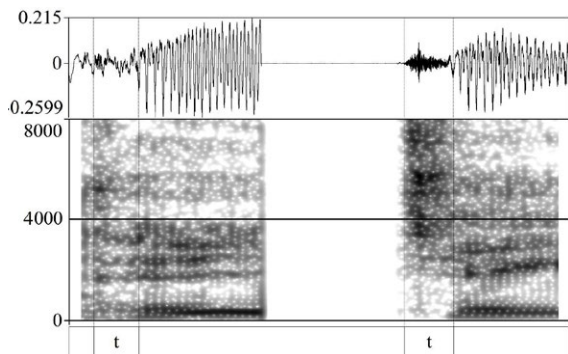
De data som benyttes i case 1, kommer fra et korpus indsamlet af Peter Skautrup Centret for Jysk Dialektforskning; en beskrivelse af korpusset og dets formål kan findes i Andersen (1981). Optagelserne er oprindeligt optaget på bånd og er nu opbevaret digitalt som MPEG1-filer. To optagelser er benyttet her:

- N: en mandlig taler fra Vrensted, Nordjylland, født 1889, optaget 1983.
- S: en mandlig taler fra Henne, Sønderjylland, født 1902, optaget 1975.

I stavelserne undersøgt her siger N *ti*, og S siger *tīd*. Disse er valgt fordi den fonetiske kontekst [_i] jævnligt udløser affrikeret opløsning. N's udtale er rent aspireret, og S' udtale er rent affrikeret, som forventet på

baggrund af deres dialektområder. De to /t/'ers VOT er respektivt 42 ms og 43 ms lange, så udtaleforskellene er ikke et resultat af forskel i VOT (jf. afsnit 3.2).

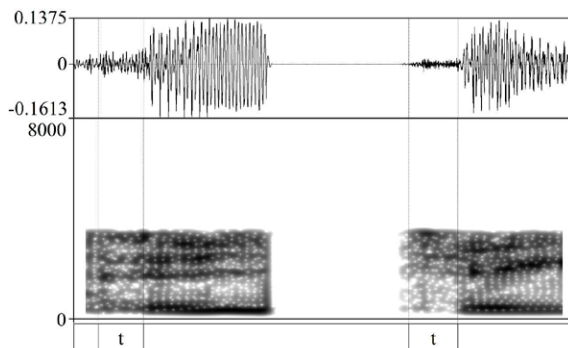
FIGUR 5. JYSK VARIATION HOS N, TIL VENSTRE, OG S, TIL HØJRE, INTET FILTER



Figur 5 viser stavelserne uden pålagte filtre. Forskellene i udtale er tydelige at se ud fra spektrogrammerne og oscillogrammerne. Aspirationsstøjen hos N er jævnt fordelt udover spektret, og den mest bemærkelsesværdige koncentration af støj ligger ved en relativt lav frekvens omkring 1.600 Hz (jf. det sværtede afsnit nederst til venstre i spektrogrammet for N). Dette er blot et resultat af præstemte vokalformanter. Affrikationsstøjen hos S er koncentreret i den højere ende af spektret, primært omkring 5.700 Hz, jf. det sværtede afsnit over midterlinjen.

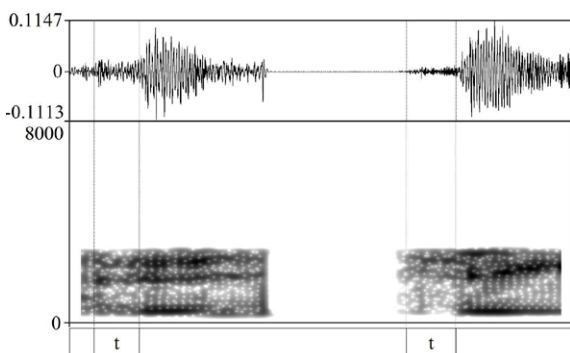
Intet distinktivt forsvinder når VoLTEs filter pålægges, selvom det er tydeligt hørbart at akustisk information er forsvundet fra S' optagelse. Det simulerede VoLTE-filter er derfor ikke vist her.

FIGUR 6. JYSK VARIATION HOS N, TIL VENSTRE, OG S, TIL HØJRE, FASTNETTETS FILTER



Figur 6 viser stavelserne med fastnettets filter simuleret. Det er stadig muligt at skelne mellem de to; fordi aspirationen i N's stavelse er jævnt fordelt udover spektret, er der bevaret mere støj end i S' stavelse. Den primære koncentration af støj er filtreret ud hos N, men den akustiske effekt af affrikationen ses stadig ved lavere frekvenser. Det samme billede viser sig ved AMR's filter (figur 7), selvom affrikationsstøjen i S' stavelse er næsten helt slettet. De to udtaler er dog stadig til at skelne fra hinanden da N's stavelse er hørbart aspireret, mens opløsningen af S' stavelse er tilnærmelsesvist lydløs.

FIGUR 7. JYSK VARIATION HOS N, TIL VENSTRE, OG S, TIL HØJRE, AMR'S FILTER



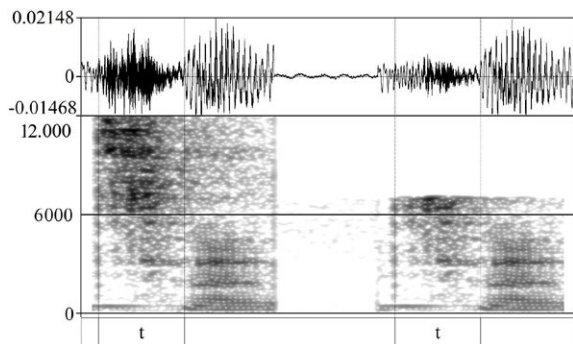
Det må konkluderes at den jyske distinktion mellem $[t^h]$ og $[t^s]$ ikke fuldstændig neutraliseres når de øvre frekvenser filtreres væk. I lavkvalitetsopkald med f.eks. AMR's filter kan affrikeret /t/ skelnes pga. en iøjnefaldende mangel på akustisk støj under opløsningsfasen af plosiven. Som forventet er begge typer udtale – men særligt affrikeret udtale – i høj grad påvirket af filtrene, og de parametre der kan benyttes til at skelne de to, er anderledes. Det bør nævnes at kun et klart aspireret og et klart affrikeret /t/ er taget i betragtning her, og det sande billede er noget mere kompliceret, med mange kontinuerlige trin (Puggaard 2017, under udg. a). Forskellige udtaler langs dette kontinuum vil muligvis blive påvirket i højere grad af filtrene.

4.3 Case 2: dansk accent

Som diskuteret i afsnit 3.3 har danskere en tendens til at udvide affrikeret udtale af /t/ til andre sprog med aspirerede /t/-fonemer. Her undersøges hvordan telefoniske båndpasfiltre påvirker dette accentelement.

De benyttede data kommer fra pilotstudiet nævnt ovenfor, der tester udtalen af engelsk /t/ blandt danske talere (Puggaard 2015, under udg. b). I dette studie blev 25 universitetsstuderende optaget med en Zoom APQ3HD-mikrofon. Alle de studerende læste tolv maskerede sætninger højt hvor /t/ forekom i en række forskellige kontekster. Hvorvidt og i hvilken grad båndpasfiltrene påvirkede /t/-udtalen, varierede meget fra taler til taler. Under alle filtre blev relevant akustisk information slettet, men for de fleste mandlige talere og nogle af de kvindelige talere var affrikationen stadig akustisk målbar og hørbar selv med AMR's filter. Dette skyldtes at affrikationen påvirkede både relativt lave og høje frekvenser. For nogle talere påvirkede affrikationen dog kun frekvenser over 3.400 Hz, hvilket betyder at den ikke var akustisk målbar eller hørbar med fastnetfiltret eller AMR-filtret. Sådan et tilfælde bliver undersøgt nedenfor i figur 9 og 10. Først ser vi på effekten af VoLTES filter på /t/-udtalen hos en 23-årig kvinde der ytrer stavelsen *ten* 'ti'.

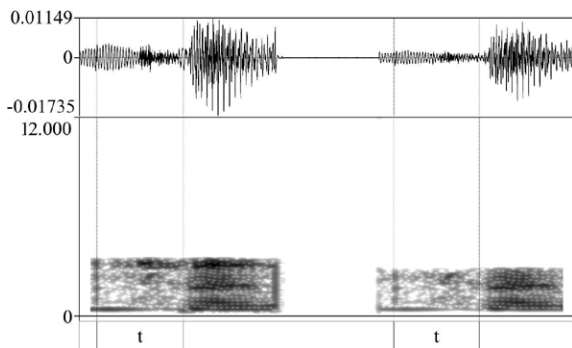
FIGUR 8. ENGLSK MED ACCENT, FØRST UDEN FILTER, DERNÆST PÅLAGT VOLTES FILTER



Figur 8 viser stavelsen uden pålagt filter og med VoLTES filter simuleret. Bemærk at spektrogrammet her viser 0-12.000 Hz da det meste af affrikationen ligger over 8.000 Hz. Som det ses, ligger det meste af affrikationen over 8.000 Hz. Den er stadig klart synlig (og hørbar) efter

at VoLTEs filter er pålagt, men det er også tydeligt både i spektrogrammet og oscillogrammet at det er blevet meget mindre prominent.

FIGUR 9. ENGLISK MED ACCENT MED HHV. FASTNETTETS OG AMR'S FILTRE



Figur 9 viser stavelsen med fastnettets og med AMR's filter simuleret. Affrikationen er helt væk i begge figurer, og forskellen i affrikation i /t/-opløsning på andetsprogsaccenter kan altså for nogle talere neutraliseres fuldstændig hvis lydssignalet er filtreret gennem fastnettets eller AMR's filtre. Dette vil dog ikke være tilfældet for alle talere; nogle gange vil der være affrikationsstøj nok tilbage ved lavere frekvenser til at det stadig er hørbart gennem fastnettets og AMR's filtre.

4.4 Case 3: rigsdansk

Jeg nævnte i afsnit 3.2 at det måske ville være muligt at finde visse personspecifikke elementer i /t/-udtale på rigsdansk såvel som på dialektalt dansk. Det testes her hvorvidt der er tegn på en sådan variation, og hvorvidt den er synlig efter at telefoniske filtre pålægges.

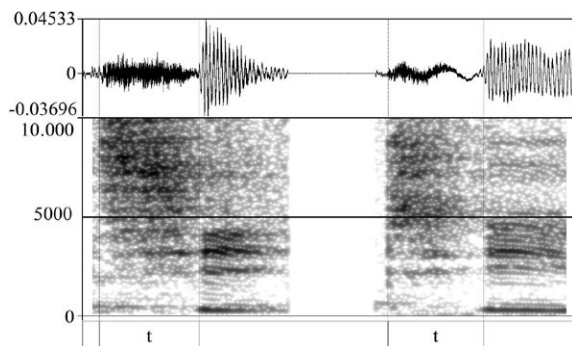
De data som benyttes i case 3, blev indsamlet på samme tid som dem til case 2; de samme deltagere læste også 12 maskerede danske sætninger højt med samme optageudstyr.

Et hurtigt kig på /t/-udtalen hos nogle af de forskellige talere viser at der er en høj grad af variation med henblik på placeringen af affrikation i spektret, og hvor i spektret affrikationsstøjen er mest intens. Som det også var tilfældet i case 2, varierer det meget fra taler til taler hvor meget af denne variation der påvirkes af de telefoniske båndpasfiltre. Der er en tendens til at VoLTE-filtret sletter meget akustisk in-

formation, men som regel er der nok affrikationsstøj tilbage på lavere frekvenser til at individuel variation bevares. Dette er noget mindre sikkert når det kommer til de smallere filtre. Det er ikke usædvanligt at der kun i ringe grad – eller overhovedet ikke – er affrikationsstøj under 3.400 Hz, mens der kan være meget variation i realiseringen over 3.400 Hz. Det er dog heller ikke usædvanligt at der er nok affrikationsstøj under 3.400 Hz til at individuel variation kan bevares.

I det nedenstående vil jeg analysere en case hvor to taleres /t/-udtale udelukkende kan skelnes ved frekvenser over 3.400 Hz for at vise hvordan en sådan variation påvirkes af båndpasfiltre. Talerne, A og B, er begge kvinder på hhv. 23 og 22 år. I den benyttede optagelse udtaler de begge stavelen /ti/ som del af samme sætning. /t/-udtalerne har sammenlignelige VOT'er på ca. 100 ms.

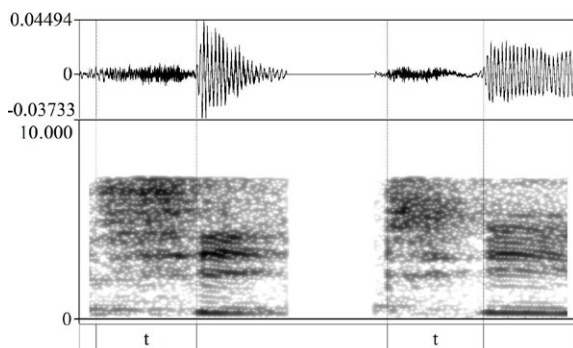
FIGUR 10. TO TALERE AF RIGSDANSK, A, TIL VENSTRE, OG B, TIL HØJRE, INTET FILTER, STAVELSEN /ti/



Figur 10 viser stavelserne uden pålagt filter. Der er tydelig variation: Det meste af affrikationsstøjen hos A ligger over 7.000 Hz og fortsætter godt over 10.000 Hz. Affrikationen er spredt i opløsningen og fortsætter næsten indtil vokalens begyndelse. Dette kan ses i spektrogrammet ved at sværtningen er nogenlunde konstant i løbet af konsonantens udtale. Støjen er tydeligst omkring midten af opløsningen. Dette tyder på en udtale med en rimeligt stram konstriktion mellem tungen og alveolærranden der først opløses meget kort tid inden vokalens begyndelse. Denne type affrikation minder om nogle /s/-udtaler og er meget auditivt prominent. Meget af affrikationen hos B er cen-

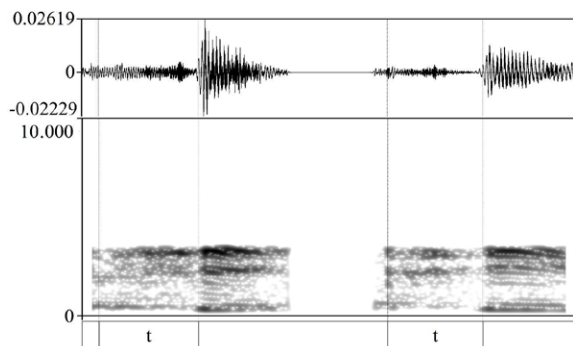
treret omkring 4.000-6.500 Hz, og der er svagere aktivitet over 9.000 Hz. Affrikationen er mest prominent til at starte med, og de sidste ca. 30 ms før vokalens begyndelse er der ikke affrikation men blot glottal aspiration. Dette ses ved at sværtningen i spektrogrammet er tydeligst til at starte med og derefter gradvist bliver svagere. Dette tyder på en udtale med en mindre stram konstriktion mellem tungen og alveolærranden. Isoleret set minder udtalen her ikke meget om /s/, da /s/-udtale som regel ikke opløses lige så gradvist. Præstemte foranter er desuden tydelige tidligt i /t/-opløsningen hos B.

FIGUR 11. TO TALERE AF RIGSDANSK, A OG B, VOLTES FILTER, STAVELSEN /ti/



Figur 11 viser stavelserne med VoLTES filter simuleret. A er mest påvirket; det er tydeligt på oscillogrammet at affrikationen er blevet meget mindre intens, og meget information er blevet filtreret væk. Variationen er dog stadig intakt, og det er stadig tydeligt at A's affrikation er mest prominent ved højere frekvenser.

FIGUR 12. TO TALERE AF RIGSDANSK, A OG B, FASTNETFILTER, STAVELSEN /ti/



Figur 12 viser de samme to stavelser efter at det simulerede fastnetfilter er pålagt. Al variation der var resultatet af affrikation, er neutraliseret. Men der er stadig variation i og med at formanterne er tydelige tidligere hos B, hvilket igen kan ses ved at den øvre sværtning sidst i plosiven er noget mere prominent hos A. Denne variation er i sig selv langt mindre tydelig (og hørbar) end variationen i realiseringen af affrikationen, der ses i figur 10. Det simulerede AMR-filter (ikke vist) giver et lignende resultat.

Det kan altså konkluderes at realiseringen af affrikation for nogle talere varierer ved så høj frekvens at smalle båndpasfiltre vil neutralisere det. I disse tilfælde kan der være andre parametre der kan afsløre tallet, såsom begyndelsestidspunktet for præstemte formanter, hvilket også muligvis kan relateres til graden af affrikation. Dette må undersøges nærmere i fremtidige undersøgelser.

5 DISKUSSION OG KONKLUSION

De optagelser som denne artikel har analyseret, er digitalt manipulerede og ikke faktisk indtalt i en telefon. Dette var et forsøg på at løse nogle af de mange problemer der er ved at skaffe gode optagelser af egentlige telefonsignaler, men manipulationen begrænser også i nogen grad hvor bredt resultaterne fra denne artikel kan generaliseres. For det første er telefoniske båndpasfiltre ikke den eneste måde hvorpå lydsignalet bliver påvirket af telefonisk overførsel, og der er også andre måder hvorpå /t/-udtale kan blive påvirket. Dette inkluderer som tidligere diskuteret aktiv støjreduktion, hvilket kan påvirke højfrekvensfænomener som affrikation selv i stille omgivelser, men særligt i støjende omgivelser; den varierende bithastighed der benyttes i AMR-forbindelser, kan desuden lede til mikroforsinkelser i signalet der kan påvirke VOT-varigheden, et aspekt jeg ikke har taget op i denne artikel. Den ovenstående analyse har vist at /t/-affrikation er temmelig akustisk prominent, men det er sandsynligt at blandingen af støjende omgivelser og et telefonisk båndpasfilter vil resultere i at endnu mere akustisk information bortfiltreres end det ses ovenfor.

Fænomenet 'telefonstemme' blev præsenteret i afsnit 2.2. Det er ikke klart hvordan dette særlige stilleje kan påvirke /t/-udtale, men

en effekt kan ikke udelukkes. Som nævnt associerer Wells (1982) 'telefonstemme' med formel stil, mens Hirson et al. (1995) associerer det med volumenmæssigt høj udtale. /t/ er affrikeret på rigsdansk, og derfor kan telefonstemme-fænomenet resultere i at talere der ikke sædvanligvis har affrikeret /t/-udtale, bruger det alligevel når de taler i telefon. Fremtidige studier må vise hvorvidt hævet lydstyrke påvirker /t/-udtale.

I denne artikel er det blevet påvist at VoLTEs filter repræsenterer en klar forbedring i lyd kvalitet i forhold til de tidligere benyttede filtre, og at dette medfører en forbedret mulighed for at finde information der er relevant for taleridentifikation. Der er ofte relevant akustisk information over 7.000 Hz i obstruenter, men analysen af figurerne 8 og 11 viste at det ikke er tilfældet at der *kun* er relevant information over 7.000 Hz. Hvis der er variation over 7.000 Hz, vil der også være variation under. Det er til gengæld usandsynligt at der vil være god VoLTE-dækning i hele landet lige foreløbigt.

Båndpasfiltre brugt til fastnettet og andre slags mobiltelefoni påvirker obstruenter (og lydsignalet generelt) i langt højere grad end VoLTEs filtre. Denne artikel har ikke fundet at AMR-filtre skaber problemer der ikke i forvejen fandtes i fastnetfiltrerne; problemerne bliver blot en anelse værre. I nogle tilfælde kan et fastnetfilter komplet neutralisere forskellen mellem to taleres /t/-udtale. Det er dog mere sandsynligt at et fastnetfilter vil neutralisere det meste af den relevante variation, men efterlade nok variation til at en talergenkendelse kan baseres på det. For et sådant retsfonetisk perspektiv er det vigtigt at gøre sig overvejelser som f.eks.:

Er der hørbar aspirationsstøj umiddelbart efter pustet? Hvis det er tilfældet, er /t/'et ikke affrikeret. Hvis opløsningen er iøjnefaldende stille, er den højst sandsynligt affrikeret over 3.400 Hz (jf. f.eks. figur 9 og 12).

Hvis der er bevaret en grad af affrikationsstøj, hvilke frekvenser begynder den så ved? Hvis støjen begynder ved relativt lave frekvenser, fortsætter den nok heller ikke op til særligt høje frekvenser. Hvis al affrikation er filtreret væk, er det sandsynligt at den begynder ved høje frekvenser og derfor også fortsætter op til meget høje frekvenser (jf. f.eks. figur 7).

Er formanter synlige før stemtheden begynder – og hvis ja, hvornår starter de? Præstemte formanter kan indikere at affrikation ikke er særligt intens, og at de har en tendens til at starte efter at affrikationens intensitet er toppet (jf. f.eks. figur 12).

Slutter plosiven på et prominent pust? I så fald kunne det tyde på at affrikationen ikke er specielt prominent. Dette kræver dog yderligere studier (jf. f.eks. figur 12).

Denne artikel har haft et temmelig smalt omfang og har udelukkende behandlet udtalen af /t/ foran en høj, urundet vokal hos enkelte udvalgte talere. Resultaterne viser til gengæld klart hvor stor en indflydelse båndpasfiltrene kan have på de høje frekvenser – hermed har resultaterne ikke blot relevans for dansk retsfonetik, men retsfonetisk praksis generelt. En retsfonetikers mål er at identificere en taler på baggrund af talerens stemmes særlige kendetegn, men mange af disse særlige kendetegn slettes når de udsættes for telefoniske båndpasfiltre, hvorved retsfonetikerens arbejdsvilkår for dataanalyse forringes. Denne artikel har både vist hvordan specifik /t/-udtale kan være et sådant særligt kendetegn, og hvad omfanget af båndpasfiltrernes betydning er for netop dette kendetegn. Dermed har artiklen påpeget væsentlige forhold ved retsfonetikken.

TAK

Tak til Mette Hjortshøj Sørensen for kommentarer og vejledning på en tidligere udgave af denne artikel. Tak desuden til den anonyme fagfællebedømmer og redaktionen ved NyS for brugbare kommentarer som har hjulpet til at gøre artiklens budskab meget mere klart.

Rasmus Puggaard
Leiden University Centre for Linguistics
rasmuspuggaard@gmail.com

LITTERATUR

- Andersen, T.A. 1981. Dialektbånd og databehandling. *Ord & Sag* 1. 11-18.
- Basbøll, H. 2005. *The phonology of Danish*. Oxford: Oxford University Press.
- Bessette, B. m.fl. 2002. The Adaptive Multi-Rate Wideband speech codec (AMR-WB). *IEEE Transactions of Speech and Audio Processing* 10(8). 620-636. DOI: 10.1109/TSA.2002.804299.
- Boersma, P. & D. Weenink. 2016. *Praat. Doing phonetics by computer*. Version 6.0.21.
- Bolt, R.H. m.fl. 1970. Speaker identification by speech spectrograms. A scientists' view of its reliability for legal purposes. *The Journal of the Acoustical Society of America* 47(2:2). 597-612. DOI: 10.1121/1.1911935.
- Broeders, A.P.A. & A.G. van Amelsvoort. 1999. Lineup construction for forensic ear-witness identification. A practical approach. J.J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, A.C. Bailey (red.), *Proceedings of the XIVth International Congress of Phonetic Sciences*, 1373-1376. University of California, Berkeley.
- Byrne, C. & P. Foulkes. 2004. The 'mobile phone effect' on vowel formants. *The International Journal of Speech, Language and the Law* 11(1). 83-102. DOI: 10.1558/ijssl.v11i1.83.
- Docherty, G.J. 1992. *The timing of voicing in British English obstruents*. Berlin: Foris Publications.
- Espegaard, A. 1996. *Nogle nordjyske mål I. Sproghistorie og dialektgeografi*. Hjørring: Forlaget Vendsyssel.
- Fischer-Jørgensen, E. & B. Hutters. 1981. Aspirated stop consonants before low vowels. A problem of delimitation. *Annual Report of the Institute of Phonetics, University of Copenhagen* 15. 77-102.
- Foulkes, P. & A. Barron. 2000. Telephone speaker recognition amongst members of a close social network. *Forensic Linguistics* 7(2). 180-198. DOI: 10.1558/sll.2000.7.2.180.
- Garibaldi, C.L. & O. Bohn. 2015. Phonetic similarity predicts ultimate attainment quite well: the case of Danish /i, y, u/ and /d, t/ for native speakers of English and of Spanish. *Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow, 10-14 August 2015*. The University of Glasgow.
- Gordon, M., P. Barthmaier & K. Sands. 2002. A cross-linguistic acoustic study of voiceless fricatives. *Journal of the International Phonetic Association* 32(2). 141-174. DOI: 10.1017/S0025100302001020.
- Grønnum, N. 2009. *Fonetik og fonologi. Almen og dansk*, 3. udgave. København: Akademisk Forlag.

- Guillemin, B.J. & C.I. Watson. 2006. Impact of the GSM AMR speech codec on formant information important to forensic speaker identification. P. Warren & C.I. Watson (red.), *Proceedings of the 11th Australian International Conference on Speech Science and Technology*, 483-488. Auckland: University of Auckland.
- Guillemin, B.J. & C.I. Watson. 2008. Impact of the GSM mobile phone network on the speech signal: some preliminary findings. *The International Journal of Speech, Language and the Law* 15(2). 193-218. DOI: 10.1558/ijssl.v15i2.193.
- Hirson, A., P. French & D. Howard. 1995. Speech fundamental frequency over the telephone and face-to-face. Some implications for forensic phonetics. J.W. Lewis (red.), *Studies in general and English phonetics. Essays in the honor of Professor J.D. O'Connor*, 230-254. London: Routledge.
- Hollien, H. 1974. Peculiar case of "voiceprints". *Journal of the Acoustical Society of America* 56(1). 210-213. DOI: 10.1121/1.1903257.
- Hollien, H. 1990. *The acoustics of crime. The new science of forensic phonetics*. New York: Plenum Press.
- Howard, D.M. & J.A.S. Angus. 2009. *Acoustics and psychoacoustics*. Oxford: Focal Press.
- Jessen, M. 2008. Forensic phonetics. *Language and Linguistics Compass* 2(4). 671-711. DOI: 10.1111/j.1749-818X.2008.00066.x.
- Kristiansen, T. 1998. The role of standard ideology in the disappearance of the traditional Danish dialects. *Folia Linguistica* 32(1-2). 115-129. DOI: 10.1515/flin.1998.32.1-2.115.
- Kristiansen, T. 2003. Language attitudes and language politics in Denmark. *International Journal of the Sociology of Language* 159. 57-71. DOI: 10.1515/ijsl.2003.009.
- Künzel, H.J. 1997. Some general phonetic and forensic aspects of speaking tempo. *Forensic Linguistics* 4(1). 48-83. DOI: 10.1558/ijssl.v4i1.48.
- Künzel, H.J. 2001. Beware of the 'telephone effect': the influence of telephone transmission on the measurement of formant frequencies. *Forensic Linguistics* 8(1). 80-99. DOI: 10.1558/ijssl.v8i1.80.
- Lisker, L. & A.S. Abramson. 1964. A cross-language study of voicing in initial stops. Acoustical measurements. *Word* 20(3). 384-422. DOI: 10.1080/00437956.1964.11659830.
- McDougall, K., F. Nolan & T. Hudson. 2015. Telephone transmission and earwitnesses. Performance on voice parades controlled for voice similarity. *Phonetica* 72. 257-272. DOI: 10.1159/000439385.
- Meuwly, D. 2003. Le mythe de "l'empreinte vocale". *Revue internationale de criminologie et de police technique et scientifique* 56(2). 219-236.

- Mortensen, J. & J. Tøndering. 2013. The effect of vowel height on Voice Onset Time in stop consonants in CV sequences in spontaneous Danish. R. Eklund (red.), *Proceedings of Fonetik 2013. The XXVth Annual Phonetics Meeting, 12-13 June 2013, Linköping University, Linköping, Sweden*, 49-52. Linköping: Linköpings Universitet.
- Nguyen, D.H., H. Nguyen & É. Renault. 2016. Predicting VoLTE quality using Random Neural Network. *International Journal of Applied Information Systems* 11(3). 1-5. DOI: 10.5120/ijais2016451587.
- Nielsen, B.J. 1984. *Bidrag til vendelbomålets fonologi. Del I: Dialekten i Tornby*. København: Akademisk Forlag.
- Nolan, F. 1983. *The phonetic bases of speaker recognition*. Cambridge: Cambridge University Press.
- Nolan, F. 1997. Speaker recognition and forensic phonetics. W.J. Hardcastle & J. Laver (red.), *The handbook of phonetic sciences*, 744-767. Oxford: Blackwell Publishing.
- Petersen, N.R. 2009. Aspirater. *Den store danske*. København: Gyldendal. Online encyklopædi, <http://denstoredanske.dk>.
- Pharao, N. m.fl. 2014. Indexical meanings of [s+] among Copenhagen youth. Social perception of a phonetic variant in different prosodic contexts. *Language in Society* 43(1). 1-31. DOI: 10.1017/S0047404513000857.
- Puggaard, R. 2015. How do Danish L2 learners produce the distinction between *tai-cai-zai*? Præsenteret ved *International Symposium on Chinese Theoretical and Applied Linguistics*. 10.-12. december, Newcastle University.
- Puggaard, R. 2017. Realizations of /t/ in Jutlandic dialects of Danish. Præsenteret ved *4th International Scientific Conference. Contemporary Research in Phonetics and Phonology: Methods, Aspects and Problems*. 11.-12. maj, Latvian Academy of Sciences, Riga.
- Puggaard, R. Under udg. a. Realizations of /t/ in Jutlandic dialects of Danish. *Linguistica Lettica* 26.
- Puggaard, R. Under udg. b. The productive acquisition of dental obstruents by Danish learners of Chinese. Chen D. & Guo Y. (red.), (Titel ikke fastlagt). Newcastle upon Tyne: Cambridge Scholars Publishing.
- Sauter, M. 2014. *From GSM to LTE-Advanced. An introduction to mobile networks and mobile broadband*. Chichester: Wiley.
- Shadle, C.H. & S.J. Mair. 1996. Quantifying spectral characteristics of fricatives. *Fourth International Conference on Spoken Languages Processing*, 1521-1524. Philadelphia, October 3-6, 1996.

- Stevens, P. 1960. Spectra of fricative noise in human speech. *Language and Speech* 3(1). 32-49. DOI: 10.1177/002383096000300105.
- Titze, I.R. 1994. *Principles of voice production*. New Jersey: Prentice Hall.
- Wells, J.C. 1982. *Accents of English 1. An introduction*. Cambridge: Cambridge University Press.
- Wolf, J.J. 1972. Efficient acoustic parameters for speaker recognition. *Journal of the Acoustical Society of America* 51(6:2). 2044-2056. DOI: 10.1121/1.1913065.
- Öhman, L., A. Eriksson & P.A. Granhag. 2010. Mobile phone quality vs. direct quality. How the presentation format affects earwitness identification accuracy. *The European Journal of Psychology Applied to Legal Context* 2(2). 161-182.