



NORGE

(12) PATENT

(19) NO

(11) 311471

(13) B1

(51) Int Cl<sup>7</sup> G 10 L 19/00

## Patentstyret

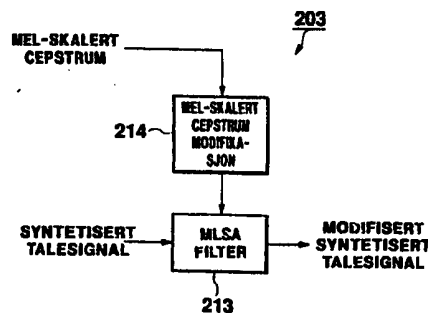
|                   |            |                                      |                           |
|-------------------|------------|--------------------------------------|---------------------------|
| (21) Søknadsnr    | 19961894   | (86) Int. inng. dag og søknadsnummer |                           |
| (22) Inng. dag    | 1996.05.10 | (85) Videreføringssdag               |                           |
| (24) Løpedag      | 1996.05.10 | (30) Prioritet                       | 1995.05.12, JP, 114752/95 |
| (41) Alm. tilgj.  | 1996.11.13 |                                      |                           |
| (45) Meddelt dato | 2001.11.26 |                                      |                           |

(71) Patenthaver Mitsubishi Denki KK,  
2-3, Marunouchi 2-chome, Chiyoda-ku, Tokyo 100-8310, JP  
(72) Oppfinner Hirohisa Tasaki, Kamakura-shi, Kanagawa, JP  
(74) Fullmektig Bryns Zacco AS, 0106 Oslo

(54) Benevnelse **Filter for modifikasjon eller forbedring av tale og forskjellig utstyr, systemer og fremgangsmåte som gjør bruk av dette**

(56) Anførte publikasjoner Ingen

(57) Sammendrag Et talemodifikasjons- eller forbedringsfilter 203 og anordning, system og fremgangsmåte som gjør bruk av dette. Syntetiserte talesignaler blir filtrert for å frembringe modifiserte syntetiserte talesignaler. Fra spektralinformasjonen som er representert som en flerdimensjonal vektor blir det bestemt en filterkoeffisient for å sikre at formantkarakteristikkene i de modifiserte syntetiserte talesignaler blir forbedret i sammenligning med de for det syntetiserte talesignal og i overensstemmelse med spektralinformasjonen. Spektralinformasjonen kan være en hvilken som helst av LSP-informasjon, PARCOR-informasjon og LAR-informasjon. En grad av frihet når det gjelder formgiving av talemodifikasjonsfilteret som benyttes for hørselsmessig undertrykkelse av kvantiseringsstøy som finnes i de syntetiserte talesignaler blir dermed høyere og fører til forbedring av oppfatningen av de syntetiserte talesignaler. En god formantforbedrende virkning kan oppnås uten å tillate noe merkbart nivå av forvrengninger å danne seg innenfor et område av tillatte spektralgradienter.



Foreliggende oppfinnelse vedrører et filter, en talesyntetiserende anordning, et talelagrende/overførende system og en fremgangsmåte til talemодifikasjon.

Oppfinnelsen finner hovedsakelig anvendelse ved overføring eller lagring av taleinformasjon ved hjelp av koder som har et lavere informasjonsinnhold enn de inngående talesignaler. På grunnlag av de innkommende talesignaler kan det utledes parametere som angir signalenes karakteristikk, overføring eller lagring av de utledede parametere og syntetisering av de opprinnelige talesignaler på grunnlag de overførte eller lagrede parametere. Ved bruk av talemодifikasjonsfilter vil kvantiserende, hørbar støy som opptrer i de syntetiserte talesignaler bli undertrykket. Ved forbedring av kvaliteten på signalet kan for eksempel tale gjøres lettere forståelig. Oppfinnelsen er særlig knyttet til taleforbedring som er egnet til å forbedre forståeligheten av tale i signalet som har forvrengninger som skyldes analog overføring eller det signal som mottas av personer med høreapparat og som er egnet til forbedring av klarheten i den tale som overføres eller skal komme som utgang fra en høyttaler.

En oppbygning av et taleanalyse/syntesystem er vist som eksempel på Fig. 28. Systemet i dette skjema omfatter en analyseenhet 100 og en syntetiseringsenhet 200. Analyseenheten 100 innbefatter en analysator 101 og en koder 102, mens synteseenheten 200 innbefatter en dekode 201 og en syntetisator 202. Ved noen anvendelser er enhetene 100 og 200 knyttet til hverandre gjennom kommunikasjonskanaler der en enhet som regel kan ligge i avstand fra den andre. I andre anvendelser overfører enheten 100 informasjon gjennom lagringsmedier til enheten 200, der de to enheter kan utgjøre en enkel anordning eller være to adskilte anordninger. Fra inngangstalesignalene som kommer fra en bruker vil analysatoren 101 utlede en parametergruppe som innbefatter spektralinformasjon som viser karakteristikkene ved de innkommende talesignaler. Den utledede parametergruppe blir kodet med koderen 102 og mates gjennom kommunikasjonskanalene eller lagringsmediene til syntetiseringsenheten 200 der den kodede parametergruppe blir dekodet av dekode 201. Syntetisatoren 202 tjener til å syntetisere talesignaler på grunnlag av parametergruppen som er dekodet på denne måte. En fordel med systemet som har denne oppbygning ligger i det lavere informasjons-innhold i de overførte eller lagrede signaler. Dette kan føres tilbake til det faktum at de overførte eller lagrede signaler, dvs. den kodede parametergruppe har et lavere informasjonsinnhold sammenlignet med de innkommende talesignaler.

En variant av syntetiseringsenheten 200 er vist på fig. 29. Denne variant omfatter videre et sluttfilter 203 som tjener til å behandle de talesignaler som utledes fra syntetisatoren 202 (i det følgende betegnet som syntetiserte talesignaler) med en på forhånd bestemt modifikasjonsprosess på grunnlag av den dekoderte parametergruppe, for dermed å frembringe modifiserte talesignaler (i det følgende betegnet som modifiserte syntetiserte talesignaler). Sluttfilteret 203 benyttes på noen anvendelsesområder for å undertrykke den kvantiserende hørbare støy som finnes i de syntetiserte talesignaler, men på andre anvendelsesområder benyttes det til å forbedre den subjektive kvalitet som for eksempel oppfatningen av tale. I den følgende beskrivelse vil sluttfilteret av denne type bli betegnet som et talemодifikasjonsfilter eller et taleforbedringsfilter. Syntetiseringsenheten 200 som er utstyrt med et slikt filter 203 er egnet til bruk i et stemmekoder/dekoder system eller et stemmegjenkjennelse system og stemmereagerende system.

En rekke forskjellige filtere er tilgjengelige som filteret 203. Fremfor alt har et filter av en type som forbedrer formantkarakteristikkene den fordel at det er overveiende effektivt til å undertrykke kvantiserende støy og ved forbedring av den subjektive kvalitet. Tidligere kjente referanser som beskriver et slikt filter innbefatter for eksempel:

Japansk utlagt patentpublikasjon nr. Sho64-13200 (i det følgende betegnet som referanse 1);

Japansk utlagt patentpublikasjon nr. Hei5-500573 (i det følgende betegnet som referanse 2);

Japansk utlagt patentpublikasjon nr. Hei2-82710 (i det følgende betegnet som referanse 3); og

"Speech Coding System Based on Adaptive Mel-Cepstral Analysis for Noisy Channel" Proceeding of Spring Meeting of Acoustical Society of Japan, bind 1, side 257-258 (1994. 3) (i det følgende betegnet som referanse 4).

Filtrene som er omhandlet i referansene 1 og 2 blir begge benyttet som talemодifikasjonsfilteret 203 i syntetiseringsenheten 200 som mottar lineære forutsigelseskoder (LPC) som den ovenfor beskrevne kodede parametergruppe fra analyseenheten 100. Et filter som omhandlet i referanse 3 blir benyttet som det talemодifiserende filter 203 i syntetiseringsenheten 200 som mottar

autokorrelasjonskoeffisienter som den ovenfor beskrevne kodede parametergruppe fra analyseenheten 100. Sluttelig blir et filter som omhandlet i referanse 4 benyttet som talemodifikasjonsfilteret 203 i syntetiseringsenheten 200 som mottar mel-skalert kepstrum eller mel-kepstrum som den ovenfor beskrevne parametergruppe fra analyseenheten 100.

Fig. 29 viser en skjematisk oppbygning av det filter som er beskrevet i referanse 1. Dette filter 203 mottar dekoderte LPC fra dekoderen 201 i tillegg til de syntetiserte talesignaler som mates fra syntetisatoren 202. LPC som det her vises til betyr  $\alpha$ -parametere som fremkommer ved lineær forutsigelseskode som skal utføres av analysator 101 som er vist på fig. 28. Den lineære forutsigelseskode er en fremgangsmåte til, på grunnlag av utvalgsbehandlede verdier av inngangstalesignalets bølgeformer og i henhold til den lineære forutsigelsesfremgangsmåte, å bestemme  $\alpha$ -parametrene eller filterkoeffisientene for filtere av f.eks. størrelsesordener åtte til tolv ved modellering av en menneskelig stemmemekanisme.

Filteret 203 som er vist på fig. 30 innbefatter et filter 204 for filtrering av syntetiserte talesignaler til frembringelse av semi-modifiserte syntetiserte talesignaler og et filter 205 for filtrering av de semi-modifiserte syntetiske talesignaler til frembringelse av modifiserte syntetiserte talesignaler der begge filtere 204 og 205 gjør bruk av  $\alpha$ -parametere som deres filterkoeffisienter. Det skal påpekes at  $\alpha$ -parameteren som benyttes i filteret 204 ikke er  $\alpha$ -parameter  $\alpha_i$  (der  $i = 1, 2, \dots, p$ ; der  $p$  er en forutsigelsesorden) som mates fra dekoderen 201, men  $\alpha_i = \alpha_i / \nu^{-i}$  som fremkommer ved modifisering av  $\alpha$ -parameteren  $\alpha_i$  med en modifisert koeffisient  $\nu$ . På samme måte er  $\alpha$ -parameteren til bruk i filteret 205 parameteren  $\alpha_2 = \alpha_i / \eta^{-i}$  som fremkommer ved modifisering av  $\alpha$ -parameteren  $\alpha_i$  med en modifisert koeffisient  $\eta$ . Prosessen for modifisering av  $\alpha$ -parameteren  $\alpha_i$  med de modifiserte koeffisienter  $\nu$  og  $\eta$  utføres av henholdsvis LPC modifikasjonsseksjonene 206 og 207.

Det antas nå at filtrene 204 og 205 henholdsvis danner en nevner og en teller for en overføringsfunksjon  $H(z)$  for transformering av de syntetiserte talesignaler til de modifiserte syntetiserte talesignaler. Med andre ord kan man la filtrene 204 og 205 være et LPC-filter resp. et inverst LPC-filter. Videre kan det forutsettes at filtrering som benytter  $\alpha$ -parameteren  $\alpha_i$  som filterkoeffisientene er gitt som:

$$A(z) = \sum_{i=0}^p (\alpha_i z^{-i}) \quad \dots(1)$$

der  $z$  er en  $z$  transformasjonsoperator. Siden filterkoeffisientene som benyttes i filterne 204 og 205 henholdsvis er  $\alpha_{1i} = \alpha_i / v^{-i}$  og  $\alpha_{2i} = \alpha_i / \eta^{-i}$  som beskrevet ovenfor, blir overføringsfunksjonene for filterne 204 og 205 henholdsvis representert i form av  $1/A$   
 5  $(z/v)$  og  $A(z/\eta)$ . Derfor kan overføringsfunksjonen for transformering av de syntetiserte talesignaler til modifiserte syntetiserte talesignaler uttrykkes som:

$$H(z) = A(z/\eta) / A(z/v) \quad \dots(2)$$

10 Fig. 31 viser skjematisk en filteroppbygning som er beskrevet i referanse 2. I dette filter 203 blir  $\alpha_{1i}$  som frembringes i LPC modifikasjonsseksjonen 206 transformert av LPC/ACC transformseksjonen 208 fra et LPC verdiområde til et autokorrelasjonsverdiområde og blir underkastet en båndbreddeutvidelse innenfor autokorrelasjonsverdiområdet av en ACC modifikasjonsseksjon 209 og blir i henhold til Levinsons  
 15 rekursjon transformert av en ACC/LPC transformseksjon 210 fra autokorrelasjonsverdiområdet til LPC verdiområdet. Filteret 205 mottar  $\alpha_{2i}$  som fremkommer på denne måte. Selv om LPC modifikasjonsseksjonen 207 som er vist på fig. 30 er fjernet på dette skjema, antyder referanse 2 også en utformning som innbefatter LPC modifikasjonsseksjonen 207 hvis utgang  $\alpha_{2i}$  igjen blir modifisert av LPC/ACC transformseksjonen  
 20 208, ACC modifikasjonsseksjonen 209 og ACC/LPC transformseksjonen 210.

Fig. 32 viser skjematisk en utformning av et filter som er beskrevet i referanse 3. Dette filter 203 er slik utført at det har ACC/LPC transformseksjoner 211 og 212 i tillegg til utformningen i referanse 1. ACC/LPC transformseksjonen 211 mottar autokorrelasjons-  
 25 konstanter som spektral informasjon innbefattet i den dekodete parametergruppe og transformerer deretter de mottatte autokorrelasjonskonstanter fra autokorrelasjonsverdiområdet til LPC verdiområdet. ACC/LPC transformseksjonen 212 mottar en del av ordenen  $m$  ( $m < p$ ) eller mindre av autokorrelasjonskonstantene som skal godtas av ACC/LPC transformseksjonen 211 og transformerer deretter de mottatte  
 30 autokorrelasjonskonstanter fra autokorrelasjonsverdiområdet til LPC verdiområdet. LPC modifikasjonsseksjonene 206 og 207 modifiserer  $\alpha$ -parametere som er utledet fra ACC/LPC transformseksjoner 211 og 212 på samme måte som i referanse 1. Det skal påpekes at autokorrelasjonskonstantene som skal innmates til denne utformning kan være de som er blitt dekodet i dekoderen 201 (dvs. autokorrelasjonskonstanter som  
 35 fremkommer ved beregning i analysatoren 101 og etter koding i koderen 102) eller kan være de som er blitt beregnet av dekoderen 201 eller syntetisatoren 202 på grunnlag av en annen type spektrale parametere som er dekodet i dekoderen 201.

Fig. 33 til fig. 35 representerer log-potens mot frekvensspektrum karakteristika for talemodifikasjons- (eller forbedrings-) filterne som er beskrevet i referansene 1 til 3. I disse diagrammer representerer A til D henholdsvis karakteristika for syntetisatoren

5 202, karakteristika for filteret 204, inverse karakteristika for filteret 205 og overføringsfunksjonen  $H(z)$ . For eksempel, på fig. 30 og 33, vil A representere  $1/A(z)$ ; B representere  $1/A(z/\nu)$ ; C representere  $1/A(z/\eta)$ ; og D representere  $H(z) = A(z/\eta) / A(z/\nu)$ . Det fremgår av uttrykk (2) som gjelder referanse 1 og også fra fig. 33 til 35 som gjelder referansene 1 til 3 at filteret 204 virker som et filter som forbedrer

10 formantene for spektrum av de syntetiserte talesignaler og undertrykker bølgedaler i dette spektrum, mens filteret 205 virker som et filter som eliminerer en spektralgradient som innføres av filteret 204. Det er antatt at graden av forbedring og undertrykkelse i filteret 204 vil øke tilsvarende når  $\nu$  blir større, og at dette vil avta når  $\nu$  blir mindre. Det antas i referanse 1 at  $\eta$  og  $\nu$  tilfredsstillers  $0 \leq \eta \leq \nu < 1$ . Fig. 33 representerer et eksempel med  $\nu = 0,8$ ,  $\eta = 0,5$ ; Fig. 34 representerer et eksempel på en prosess med

15 utvidelse av båndbredden gjennom et 1200 Hz etterslepvindu med  $\nu = 0,8$ ; og fig. 35 representerer et eksempel med  $p = 10$ ,  $m = 4$ ,  $\nu = 0,95$ ,  $\eta = 0,95$ .

Som det skulle fremgå av sammenligningen mellom fig. 33 og 34 eller fra

20 sammenligningen mellom fig. 33 og 35, vil talemodifikasjons- (eller forbedrings-) filteret i referansene 2 og 3 være i stand til å forhøye virkningen ved eliminering av spektralgradienten ved bruk av filteret 205 sammenlignet med det filter som er beskrevet i referanse 1. Dette betyr at den teknikk som er beskrevet i referanse 1 ikke vil sette filteret 205 helt i stand til å utligne spektralgradienten som er innført av filteret

25 204. Siden spektralgradienten varierer i tidens løp, vil det dessuten være vanskelig for en fast høyfrekvens spektrumforbedrende prosess å utligne spektralgradienten, noe som vil føre til en variasjon i klarhet med tiden. På den annen side vil de teknikker som er beskrevet i referansene 2 og 3 gjøre det mulig å øke virkningen av forbedringen av toppbølgedal strukturen for spekteret for å gjøre den spektrale gradient flattere. Dette vil føre

30 til at forringelsen i klarhet og naturlighet motvirkes av filteret 203.

Det skal påpekes at de teknikker som er beskrevet i referansene 2 og 3 i et henseende er en forbedring sammenlignet med den teknikk som er beskrevet i referanse 1, men i et annet henseende er mindreverdige i forhold til denne. Selv om det kan avhenge av

35 oppbygningen av analyseenheten 100 eller den fremgangsmåte som systemet er tilpasset, har for eksempel den teknikk som er beskrevet i referanse 2 en mangel ved at de resulterende modifiserte syntetiserte talesignaler ofte innebærer unike

forvrengninger. Dette skyldes det faktum at en overordentlig kraftig glatteprosess for spekteret utføres i autokorrelasjonens verdiområde med det resultat at spekteret blir merkbart forvrengt i nærheten av de sterke formanter. Dette kan resultere i at de modifiserte syntetiske talesignaler har en mindreverdige kvalitet sammenlignet med den teknikk som er beskrevet i referanse 1. Når det gjelder den teknikk som er beskrevet i referanse 3, vil denne på grunn av reduksjonen i filterordenen i autokorrelasjonens verdiområde ofte lide av mangler ved at posisjonene for formantene blir forskjøvet i en større utstrekning eller at flere formanter blir integrert til en. En slik ustabil spektral variasjon vil være årsak til forvrengninger i de modifiserte syntetiserte talesignaler. Fra en sammenligning mellom karakteristikene B og C som er angitt i fig. 35 for eksempel vil man se at et fenomen oppstår der en formant som har den laveste frekvens blant formantene i B føres til en lavere frekvens i C og et fenomen med integrasjon av to formanter på midten. Videre vil den signifikante formantforskyvning på grunn av disse årsaker kunne opptre eller kanskje ikke opptre med tid, med den følge at den resulterende modifiserte syntetiserte tale vil fluktuere unaturlig.

De teknikker som er beskrevet i referansene 1 til 3 medfører også et felles problem med lav frihetsgrad i utformningen (frihet når det gjelder operasjon og styring av karakteristikene). Når det gjelder den teknikk som er beskrevet i for eksempel referanse 1 vil det være vanskelig å forandre karakteristikene for filteret 203 i en større utstrekning ved bare å variere  $\nu$  og  $\eta$  innen et område der problemene med den spektrale gradient og dens variasjon med tid ikke blir så markert. Når det gjelder den teknikk som er beskrevet i referanse 2, vil, hvis større variable områder er stilt for  $\nu$  og etterslepwindusfrekvensen for å øke den formantforbedrende virkning ved filteret 204, de ovenfor beskrevne forvrengninger, dvs. de forvrengninger som skyldes prosessen med glatting av spekteret i autokorrelasjonens verdiområde, bli mer fremtredende. Derfor må de variable områder for  $\nu$  og etterslepwindusfrekvensen bli begrenset, noe som gjør det umulig i større grad å endre karakteristikene for filteret 203. Når det gjelder den teknikk som er beskrevet i referanse 3, vil friheten for karakteristikene naturligvis bli senket siden det her gjøres bruk av filterordenen som dets styrevariabel, som er en endelig integralverdi.

Fig. 36 viser skjematisk en oppbygning av det talemодifiserende (eller forbedrende) filter 203 som er beskrevet i referanse 4. Filteret 203 i dette skjema skiller seg vesentlig fra de ovenfor beskrevne tidligere kjente teknikker ved at det mottar mel-skalert kepstrum som spektralinformasjon innbefattet i den dekodete parametergruppe fra dekoderen 201 og ved at det transformerer syntetiserte talesignaler til modifiserte

syntetiserte talesignaler ved filtrering, der det som dets filterkoeffisient benyttes det modifiserte mel-skalerte kepstrum som oppnås ved modifisering av det mel-skalerte inngangs-kepstrum. Dette betyr at syntetiserte talesignaler blir filtrert med et filter 213 der det som filterkoeffisienter brukes de mel-skalerte kepstrum som er frembragt av en  
5 mel-skalert kepstrum-modifiserende seksjon 214. Mer bestemt erstatter den mel-skalerte kepstrum-modifiserende seksjon 214 den førsteordenskomponent av det mel-skalerte inngangs-kepstrum med 0 og multipliserer de andre komponenter med  $\beta$  for dermed å frembringe et modifisert mel-skalert kepstrum. Filteret 213 gjør bruk av dette modifiserte mel-skalerte kepstrum som sin filterkoeffisient for å filtrere de syntetiserte  
10 talesignaler og leverer frembragte signaler som sin utgang i form av modifiserte syntetiserte talesignaler. Det skal nevnes at filteret 213 blir betegnet som et mel-skalert log-spektraltilnærmende (MLSA) filter siden det gjør bruk av det modifiserte mel-skalerte kepstrum som sin filterkoeffisient.

15 Uttrykket mel-skalert kepstrum som her er benyttet betyr en parameter som er beregnet av analysatoren 101 ved ortogonal transformasjon av log-spekteret for inngående talesignaler. Det ville i alminnelighet være umulig med teknikkene i referansene 1 til 3 å kunne anvende disse slik de står i et system der taleinformasjonen blir transformert i mel-skalert kepstrum for overføring eller lagring. Dette betyr at overføring av kepstrum-  
20 parametere så som mel-skalert kepstrum til LPC verdiområdet ville føre til betydelig forvrengning av spektralgeometrien, noe som ville nødvendiggjøre beregning av LPC ved re-analyse av de syntetiserte talesignaler. Selv om den på denne måte beregnede LPC inneholder forvrengninger i forhold til den LPC som fremkommer ved analyse av den opprinnelige tale, vil det i tillegg ikke sikre noen gode talemodifikasjons-  
25 karakteristikk. Tvert imot vil fremgangsmåten i referanse 4 være i stand til å unngå at disse forvrengninger oppstår.

Omvendt betyr dette at den teknikk som er beskrevet i referanse 4 vil stå overfor et problem som skyldes dårlige forbindelsesmuligheter, med andre ord ville det være  
30 umulig å anvende denne teknikk på systemer som er beregnet på å syntetisere talesignaler ved bruk av en parametergruppe som er en annen enn kepstrum-parametere. Typiske slike systemer er for eksempel de som benytter parametergrupper som for eksempel LPC, LSP (linjespektrumpar) og PARCOR (partielle autokorrelasjonskoeffisienter). Dette problemet er alvorlig siden LPC, LSP og PARCOR ofte blir  
35 benyttet til talekoding/dekoding. Hvis et talemodifikasjonsfilter som benytter mel-skalert kepstrum som sin filterkoeffisient blir tatt med i syntetiseringsenheten 200 som mottar LPC med en av parametrene, vil spektralgeometrien bli forvrent med

transformasjon fra LPC verdiområdet til det mel-skalerte kepstrum-verdiområde som tidligere beskrevet. Det er naturlig at denne forvrengning kan elimineres i en viss grad ved igjen å beregne det mel-skalerte kepstrum ved re-analyse av de syntetiske talesignaler. Selv om det mel-skalerte kepstrum er blitt beregnet på denne måte, vil det  
5 imidlertid fremdeles inneholde fler forvrengninger sammenlignet med det mel-skalerte kepstrum som vil bli utledet fra den opprinnelige tale. Dermed kan det ikke ventes særlig gode talemodifikasjons-karakteristikker.

Et første formål med foreliggende oppfinnelse er å komme frem til et talemodifikasjons-  
10 (eller forbedrings-, et uttrykk som vil bli utelatt senere) filter som sikrer en god formantforbedrende virkning innenfor et område med tillatte spektralgradienter. Et annet formål med foreliggende oppfinnelse er å komme frem til et talemodifikasjonsfilter som sikrer en god formantforbedrende virkning uten å føre til noe merkbart nivå og forvrengning i formantstrukturen. Et tredje formål med foreliggende oppfinnelse er å  
15 komme frem til et talemodifikasjonsfilter som er i stand til å implementere den samme formantforbedrende virkning som tidligere kjent ved bruk av et mindre antall bestanddeler enn det som er tidligere kjent. Et fjerde formål med foreliggende oppfinnelse er å komme frem til et talemodifikasjonsfilter som muliggjør selektiv utførelse av styringen av klarhet, reduksjon av prosessprosedyrene, forbedring i oppfatning, etc. Et femte  
20 formål med foreliggende oppfinnelse er å unngå behovet for stabilitetens sikkerhet i verdiområdet hvis natur er forskjellig fra det verdiområde som den spektrale inngangsinformasjon hører til, og dermed komme frem til et talemodifikasjonsfilter med større frihetsgrad når det gjelder utformning. Et sjette formål med foreliggende oppfinnelse er å komme frem til et talemodifikasjonsfilter som er egnet for en syntetiseringsenhet som  
25 mottar LSP, PARCOR, LAR (log-områdeforhold), etc., som spektralinformasjon fra analyseenhetens side. Et syvende formål med foreliggende oppfinnelse er å komme frem til et talemodifikasjonsfilter som etter inngang med LSP, PARCOR, LAR, etc. som spektralinformasjon sikrer en god forbindelsesmulighet uten behov for noen spektral re-analyse eller parametertransformering. Det er et åttende formål med foreliggende  
30 oppfinnelse å implementere et talesyntetiserende system ved bruk av det talemodifiserende filter som er i stand til å oppfylle de ovenstående første til syvende formål.

De for oppfinnelsen kjennetegnende trekk fremgår av de vedlagte patentkrav, samt av  
35 den nå etterfølgende beskrivelse.

Ved foreliggende oppfinnelse blir syntetiserte talesignaler filtrert gjennom en overføringsfunksjon som er fastlagt med en filterkoeffisient for å frembringe modifiserte syntetiserte talesignaler. Denne filterkoeffisient blir frembragt på grunnlag av spektralinformasjon som er representert i form av en flerdimensjonal vektor og hører til et på 5 forhånd bestemt verdiområde og angår inngangstalesignalene på en slik måte at formantkarakteristikker for de modifiserte syntetiserte talesignaler blir forbedret ifølge den ovenstående spektralinformasjon og i sammenligning med de for de syntetiserte talesignaler. Som tilgjengelig spektralinformasjon er en hvilken som helst av LSP-informasjon, PARCOR-informasjon og LAR-informasjon. På grunn av de spesielle 10 trekk ved LSP-informasjonen, PARCOR-informasjonen og LAR-informasjonen kan operasjonene til frembringelse av filterkoeffisientene foregå som operasjoner med en slik natur at aritmetikk som er knyttet til de individuelle dimensjoner blir avhengig av aritmetikk som er knyttet de gjenværende dimensjoner. Når LSP-, PARCOR- eller LAR-informasjon benyttes til å frembringe filterkoeffisienter, kan filterstabiliteten 15 sikres uten at koeffisientene transformeres fra LSP, PARCOR eller LAR verdiområdet til et annet verdiområde. Det skal påpekes at i det filter som for eksempel ved bruk av filterkoeffisientene som er frembragt fra LPC-informasjonen er det nødvendig å transformere filterkoeffisientene fra LPC verdiområdet til et annet verdiområde for å få til stabilitet for filteret. Som følge av dette er det lettere å utforme talemodifikasjons- 20 prosessen eller filteret uten å innføre ustabilitet i dette enn tilfellet er med tidligere kjente utførelser som benytter filterkoeffisienter frembragt fra LPC-informasjonen. I tillegg vil anvendelse av dette trekk på systemer for overføring eller lagring av LSP-informasjon, PARCOR-informasjon eller LAR-informasjon ikke ha behov for noen spektrum re-analyse eller parametertransformasjon hvorved en god tilkoblingsmulighet 25 kan sikres.

Filtreringen ifølge foreliggende oppfinnelse kan utføres innenfor et hvilket som helst av LPC verdiområde, LSP verdiområde og PARCOR verdiområde. Med andre ord kan filterkoeffisientene tilhøre et hvilket som helst av LPC verdiområde, LSP verdiområde 30 og PARCOR verdiområde. Ifølge et andre aspekt ved oppfinnelsen blir spektralinformasjon først modifisert i et verdiområde som den tilhører for å frembringe modifisert spektralinformasjon og den modifiserte spektralinformasjon blir så transformert fra dette verdiområde til LPC verdiområde for å frembringe filterkoeffisienter, og de filterkoeffisienter som fremkommer på denne måte blir benyttet for 35 filtrering i LPC verdiområdet. Siden en rekke modifiserte koeffisienter kan anvendes for denne modifikasjon, vil dette gjøre det mulig å modulere filterkoeffisientsyntesen med større frihet enn tidligere kjente utgaver, i overensstemmelse med de filtrerings-

karakteristikk (syntetiserte talesignalers modifikasjonkarakteristika) som brukere krever.

Ifølge et ytterligere aspekt ved oppfinnelsen blir spektralinformasjonen modifisert slik at toppene på formantene i de modifiserte syntetiserte talesignaler blir redusert. Derfor vil dette gjøre det mulig å komme frem til en god formantforbedrende virkning innenfor et område som kan godtas for spektralgradienter og å komme frem til en god formantforbedrende virkning uten at det oppstår noe følbart nivå på forvrengninger i formantstrukturen.

10

Som en første modifikasjon kan det tenkes en løsning der spektralinformasjonen som gjelder inngangstalesignalene og referanseinformasjonen som hører til samme verdiområde blir proporsjonalt delt ifølge den modifiserte koeffisient. Denne fremgangsmåte kan anvendes når spektralinformasjonen er LSP-informasjon. Avhengig av fremgangsmåtene for oppstilling av denne referanseinformasjon ville denne fremgangsmåte gjøre det mulig å utføre de følgende modifikasjoner, for eksempel: en modifikasjon for innføring av en fast spektralgradient i de modifiserte syntetiserte talesignaler; en modifikasjon for å innføre en spektrumgradient som viser gjennomsnittlig støyspektrum for de modifiserte syntetiserte talesignaler (dvs. en modifikasjon for svak forbedring av et annet talespektrum enn støyspekteret); og en modifikasjon for å innføre i de modifiserte syntetiserte talesignaler en spektrumgradient som viser en historie som den spektrale informasjonen har gjennomløpt så langt (dvs. en modifikasjon for forbedring av størrelsen på variasjonen i talespekteret). Dette vil gjøre det mulig å utøve kontroll av klarhet, reduksjon i behandlingsprosedyrene for informasjon, og forbedring i oppfatningsmuligheten. Denne fremgangsmåte lar også filteret videre utvikle karakteristikkene for de andre sekundære filtreringsprosesser (for eksempel en fast høyfrekvent forbedringsprosess).

Tenkelig som en andre modifikasjon er en fremgangsmåte der for hver av en rekke dimensjoner som utgjør spektralinformasjonen for inngangstalesignalene blir denne spektralinformasjonen multiplisert med en modifisert koeffisient eller med potensen for den modifiserte koeffisient. Denne fremgangsmåte kan brukes når spektralinformasjonen er enten PARCOR-informasjon eller LAR-informasjon. Denne fremgangsmåte sikrer også noen av de virkninger som er angitt ovenfor, for eksempel reduksjon av prosessen, forbedret oppfatningsmulighet, etc. Det skal påpekes at når spektralinformasjonen er PARCOR-informasjon, benyttes fremgangsmåten med

35

multiplikasjon av spektralinformasjonen med potensen for den modifiserte koeffisient og denne potens er avhengig av dimensjonen på spektralinformasjonen.

En tenkelig tredje modifikasjon er en fremgangsmåte hvori avstander utvides mellom sammenstående dimensjoner blant en rekke dimensjoner som representerer den spektralinformasjon som gjelder inngangstalesignalene. Mer bestemt, når en avstand mellom sammenstående dimensjoner er mindre enn en referanseavstand, blir avstanden utvidet forbi referanseavstanden og deretter blir avstanden likt krympet i forhold til alle dimensjonene for dermed å sikre at bredden på spektralinformasjonen i sin helhet vil falle sammen med bredden før utvidelsen. Denne fremgangsmåte kan benyttes når spektralinformasjonen er LSP-informasjon. Fremgangsmåten gjør det mulig å modifisere spektralinformasjonen slik at spekteret for de modifiserte syntetiserte talesignaler blir flatet ut og sikrer noen av de virkninger som er angitt ovenfor, for eksempel redusert prosess, forbedret oppfatningsmulighet, etc. når det gjelder glatting av spektralgradienten. I tillegg oppnås reduksjon av den prosess eller de komponenter som gjelder den første og den andre fremgangsmåte.

Det er også tenkelig at nevnte første og tredje modifikasjoner kan kombineres med hverandre etter valg, eller som alternativ kan begge bli benyttet ved samvirkning. Når det gjelder fordelene ved hver fremgangsmåte i forhold til de to andre fremgangsmåter og forskjellene mellom de tre fremgangsmåter, skulle dette fremgå for fagfolk på området fra den følgende beskrivelse av utførelsesformer.

De første til tredje modifikasjoner kan utformes slik: for det første stilles det opp en omsettingstabell som lagrer spektralinformasjon om inngangstalesignalene i korrelasjon med modifisert spektralinformasjon og det frembringes modifisert spektralinformasjon som reaksjon på tilførsel av den spektrale informasjon; og for det annet frembringes det et nevralt nett som ved innlæring har oppnådd evne til å transformere spektralinformasjon til modifisert spektralinformasjon for dermed å kunne frembringe den modifiserte spektralinformasjon ved tilførsel av spektralinformasjonen om innmatede talesignaler. Det er å foretrekke at omsettingstabellen og nevraltettet anordnes for hver av en rekke kategorier som ikke overlapper hverandre og som er fremkommet ved klassifisering av de verdiområder som den spektrale informasjon om de innmatede talesignalene tilhører, eller at de benyttes mens deres aktiviteter veksles ved veksling av koeffisientene for hver kategori. Dette ville gjøre det mulig å anordne en adaptiv styring gjennom kategoridelingen og redusere forvrengninger ved kategoriernes grenser. Det

ville også være mulig å benytte en hvilken som helst fremgangsmåte til modifikasjon i tillegg til de første til tredje fremgangsmåter for hver kategori.

Ifølge et fjerde aspekt ved oppfinnelsen der filtrering utføres i et hvilket som helst av  
5 LSP verdiområde og PARCOR verdiområde blir den spektrale informasjon om inngangstalesignalene modifisert i et verdiområde som den hører til og den resulterende modifiserte spektrale informasjon benyttes som en filterkoeffisient. Dette vil oppheve behovet for overføring av verdiområder som er knyttet til den modifiserte spektral-  
informasjon, noe som gjør det mulig å oppnå stort sett samme formantforbedrende  
10 virkning som tidligere er oppnådd, men med et mindre antall bestanddeler enn det som er tidligere kjent.

I henhold til et femte aspekt ved oppfinnelsen blir filtrering utført slik at formantene i de modifiserte syntetiserte talesignaler blir ytterligere forbedret sammenlignet med  
15 formantene i de syntetiserte talesignaler. I henhold til et sjette aspekt ved oppfinnelsen blir den spektrale gradient som skal innføres i nevnte modifiserte syntetiserte talesignaler undertrykket.

Ifølge et syvende aspekt ved oppfinnelsen blir syntetiserte talesignaler frembragt på  
20 grunnlag av spektralinformasjon som er representert som en flerdimensjonal vektor og tilhører et på forhånd bestemt verdiområde og gjelder de innmatede talesignalene, og deretter blir den prosess som inngår i de ovenfor beskrevne trekk utført på grunnlag av spektralinformasjonen. Ifølge et åttende aspekt ved oppfinnelsen frembringes det syntetiserte talesignaler på basis av en første spektralinformasjon som er representert  
25 som en flerdimensjonal vektor og tilhører et på forhånd bestemt verdiområde og gjelder inngangstalesignalene, og den første spektralinformasjon blir transformert til en andre spektralinformasjon som hører til et annet verdiområde enn det verdiområdet som den første spektralinformasjon har tilhørt så langt, og deretter blir de prosesser som inngår i de ovenfor beskrevne trekk utført på grunnlag av den andre spektralinformasjon. Ifølge  
30 et niende aspekt ved oppfinnelsen frembringes det syntetiserte talesignaler på grunnlag av den første spektralinformasjon som gjelder inngangstalesignaler og tilhører et på forhånd bestemt verdiområde og er representert som en flerdimensjonal vektor, og de syntetiserte talesignaler blir analysert for å frembringe en andre spektralinformasjon, og deretter blir de prosesser som inngår i de ovenfor beskrevne trekk utført på grunnlag av  
35 den andre spektralinformasjon. Ifølge et tiende aspekt ved oppfinnelsen blir det før prosessene som inngår i nevnte syvende til niende aspekt spektralinformasjonen eller

den første spektralinformasjon frembragt ved analyse av inngangstalesignalene, og spektralinformasjonen eller den første spektralinformasjon blir lagret eller overført.

Fig. 1 og fig. 2 er blokkskjemaer som hver viser en konfigurasjon for et  
5 talemодifikasjonsfilter ifølge en LSP-basert utførelse blant foretrukne utførelser av foreliggende oppfinnelse;

Fig. 3 er et blokkskjema som, som eksempel viser en konfigurasjon for et  
taleanalyse/syntesystem;

10

Fig. 4 er et blokkskjema som viser et eksempel på en fremgangsmåte for LSP-  
modifikasjon;

Fig. 5 er et forklarende skjema for en fremgangsmåte til frembringelse av modifisert  
15 LSP ved en proporsjonal deling;

Fig. 6 og fig. 7 er blokkskjemaer som hver viser et eksempel på fremgangsmåten til  
LSP-modifikasjon;

Fig. 8 er en grafisk gjengivelse av log-potens i forhold til  
20 frekvensspektrumkarakteristika for den LSP-baserte utførelse blant de foretrukne utførelser av foreliggende oppfinnelse, hvilke karakteristika er fremkommet i eksemplet der det benyttes en fremgangsmåte til frembringelse av modifisert LSP ved proporsjonal deling i den konfigurasjon som er vist på fig. 1;

25

Fig. 9 er et blokkskjema som viser et eksempel på fremgangsmåten til LSP-  
modifikasjon;

Fig. 10 er en grafisk gjengivelse av log-potens i forhold til  
30 frekvensspektrumkarakteristika for den LSP-baserte utførelse blant de foretrukne utførelser av foreliggende oppfinnelse, hvilke karakteristika er fremkommet ved bruk av en fremgangsmåte til frembringelse av modifiserte LSP ved utvidelse av avstanden mellom sammenstående dimensjoner i konfigurasjonen på fig. 2;

Fig. 11, fig. 12, fig. 13, fig. 14, fig. 15 og fig. 16 er blokkskjemaer som hver viser et  
35 eksempel på fremgangsmåten til LSP-modifikasjon;

Fig. 17 og fig. 18 er blokkskjemaer der hver viser en konfigurasjon for et talemodifikasjonsfilter i overensstemmelse med en utformning som utfører filtrering innen LSP verdiområde blant de foretrukne utførelser av foreliggende oppfinnelse;

- 5 Fig. 19 er et blokkskjema som viser konfigurasjonen av et talemodifikasjonsfilter ifølge en PARCOR-basert utførelse blant de foretrukne utførelser av oppfinnelsen;

Fig. 20 er en grafisk gjengivelse av log-potens i forhold til frekvensspektrumkarakteristika for den PARCOR-baserte utførelsen blant de foretrukne  
10 utførelser av foreliggende oppfinnelse;

Fig. 21 og fig. 22 er blokkskjemaer som hver viser en konfigurasjon for et talemodifikasjonsfilter ifølge en utformning som utfører filtrering i PARCOR verdiområde blant de foretrukne utførelser av foreliggende oppfinnelse;

15

Fig. 23 er et blokkskjema som viser en konfigurasjon for et talemodifikasjonsfilter i henhold til en LAR-basert utførelse blant de foretrukne utførelser av foreliggende oppfinnelse;

- 20 Fig. 24 er en grafisk gjengivelse av log-potens i forhold til frekvensspektrumkarakteristika for den LAR-baserte utførelse blant de foretrukne utførelser av foreliggende oppfinnelse;

Fig. 25 og fig. 26 er blokkskjemaer som hver viser en konfigurasjon av et  
25 talemodifikasjonsfilter ifølge en utformning som utfører filtrering i et LAR verdiområde eller et PARCOR verdiområde blant de foretrukne utførelser av foreliggende oppfinnelse;

Fig. 27 er et blokkskjema som viser en konfigurasjon av et talemodifikasjonsfilter ifølge  
30 en utførelse som gjør bruk av en rekke parametere blant de foretrukne utførelser av foreliggende oppfinnelse;

Fig. 28 er et blokkskjema som, som eksempel viser, en konfigurasjon for et taleanalyse/syntesystem;

35

Fig. 29 er et blokkskjema som viser en bruksmåte for et talemodifikasjonsfilter;

Fig. 30, fig. 31 og fig. 32 er blokkskjemaer som viser konfigurasjoner av de talemodifikasjonsfiltere som er beskrevet i referanse 1, referanse 2 og referanse 3;

Fig. 33, fig. 34 og fig. 35 er grafiske gjengivelser av log-potens i forhold til  
5 frekvensspektrumkarakteristika for de talemodifikasjonsfiltere som er beskrevet i referanse 1, referanse 2 og referanse 3; og

Fig. 36 er et blokkskjema som viser en konfigurasjon av det talemodifikasjonsfilter som er beskrevet i referanse 4.

10

Utførelser av foreliggende oppfinnelse vil nå bli beskrevet med henvisning til tegningene, der bestanddeler som er identiske eller svarer til tidligere kjent teknikk som er vist på figurene 28 til 36 har samme henvisningstall og ikke vil bli ytterligere forklart. Det skal påpekes at bestanddelene i de respektive utførelser også er betegnet med  
15 samme henvisningstall og ikke vil bli forklart flere ganger.

Det henvises først til figurene 1 og 2 der det er vist to utførelser som mottar LSP (LSP = linjespektrumpar) som spektralinformasjon i dekodet parametergruppe blant foretrukne utførelser av et filter 203 ifølge foreliggende oppfinnelse. Den utførelse som er vist på  
20 fig. 1 omfatter LSP modifikasjonsseksjoner 216 og 217 og LSP/LPC transformseksjoner 218 og 219 i tillegg til filterne 204 og 205. Dessuten omfatter utførelsen som er vist på fig. 2 LSP modifikasjonsseksjonen 216 og LSP/LPC transformseksjonen 218 i tillegg til filteret 204.

25 Disse utførelser kan benyttes i syntetiseringsenheten 200 som har en konfigurasjon som vist på fig. 30 eller fig. 3. Når det gjelder bruk av dekoderen 201 som er i stand til å gi LSP som utgang, som en del av en parametergruppe, kan filteret 203 direkte motta matningen fra dekoderen 201 som vist på fig. 29, mens det i eksemplet med bruk av dekoderen 201 som ikke er i stand til å gi LSP som utgangsinformasjon som en del av  
30 en parametergruppe, må matningen fra dekoderen 201 transformeres gjennom en transformseksjon 215 inn i LSP verdiområdet og deretter tilføres filteret 203, som vist på Fig. 3. Det skal påpekes at transformseksjonen 215 kan være integrert i dekoderen 201 eller i syntetisatoren 202.

35 LSP modifikasjonsseksjonene 216 og 217 mottar LSP  $\omega_i$  i form av en flerdimensjonal vektor fra dekoderen 201 eller transformseksjonen 215 og modifierer  $\omega_i$  i overensstemmelse med en på forhånd bestemt fremgangsmåte for å frembringe

henholdsvis modifisert LSP  $\omega_{h1_i}$  og  $\omega_{h2_i}$ . LSP/LPC transformseksjonene 218 og 219 transformerer  $\omega_{h1_i}$  og  $\omega_{h2_i}$  fra LSP verdiområdet til LPC verdiområdet for å frembringe modifiserte  $\alpha$ -parametere  $\alpha_{1_i}$  og  $\alpha_{2_i}$ . Filtrene 204 og 205 utfører i rekkefølge filtrering av syntetiserte talesignaler ved bruk av  $\alpha_{1_i}$  og  $\alpha_{2_i}$  som deres  
 5 respektive filterkoeffisienter. Som et resultat gir filteret 205 modifiserte syntetiserte talesignaler som sin utgang. Nå kan det antas at overføringsfunksjonene for filtrene 204 og 205 er  $1/A_1(z)$  og  $A_2(z)$ , og da kan overføringsfunksjonen for filteret 203 på fig. 1 gis som

$$10 \quad H(z) = A_2(z) / A_1(z) \quad \dots(3)$$

og overføringsfunksjonen for filteret 203 på fig. 2 kan gis som

$$H(z) = 1 / A_1(z) \quad \dots(4)$$

15

I den LSP-baserte utførelse (LSP = linjespektrumpar) av foreliggende oppfinnelse blir på denne måte LSP  $\omega_i$  som er mottatt som en av parametrene modifisert og de modifiserte LSP  $\omega_{h1_i}$  (og LSP  $\omega_{h2_i}$ ) blir transformert fra LSP verdiområdet inn i LPC verdiområdet for dermed å frembringe filterkoeffisienter  $\alpha_{1_i}$  (og  $\alpha_{2_i}$ ) som er  
 20 modifiserte parametere. En første fordel ved den LSP-baserte utførelse som fremkommer på denne måte ligger i at det er lett å påvise og sikre filteret 203 stabilt, siden stabiliteten kan kontrolleres i LSP verdiområde. Mer bestemt er det alminnelig kjent at det filter som bruker LSP  $\omega_i$  er stabilt når LSP  $\omega_i$  tilfredsstiller den følgende sekvensielle betingelse:

25

$$0 < \omega_1 < \omega_2 < \dots < \omega_p < \pi \quad \dots(5)$$

Så lenge den LSP som tilfredsstiller ligning (5) blir benyttet som filterkoeffisienten, kan derfor prosessen med frembringelse av  $\alpha_{1_i}$  og  $\alpha_{2_i}$  utføres uavhengig for respektive  $i$ ,  
 30 uten innføring av ustabilitet i filteret. Som et resultat oppnås en stor frihet når det gjelder filterutformningen. For eksempel er det mulig å utforme et filter som kan forbedre høyfrekvenskomponentene i talen ved å sette forbedringsgraden for de høyordens dimensjoner på en forholdsvis lav verdi. Omvendt, når det gjelder forhold der  $\alpha$ -parameteren eller autokorrelasjonskonstanten benyttes til å frembringe filterkoeff-  
 35 fisienten, kan bare den prosess som har vist at den ikke innfører ustabilitet i filteret benyttes til frembringelse av  $\alpha_{1_i}$  og  $\alpha_{2_i}$ , som i referansene 1 til 3, siden det i  $\alpha$ -parameterens verdiområde eller i autokorrelasjonens verdiområde er vanskelig å påvise

og sikre stabilitet for filteret som benytter filterkoeffisientene basert på slike parametere. Modifikasjonsprosessen som utføres for respektive i eller med justering av forbedringsgraden langs frekvensaksen kan som følge av dette ikke utføres uten å tillate innføring av ustabilitet i filteret når det benyttes filterkoeffisienter som er basert på  $\alpha$ -parameteren eller autokorrelasjonen.

En andre fordel med den LSP-baserte utførelse ligger i større anvendbarhet ved systemer som sender eller lagrer LSP som den spektrale informasjon. Særlig de fleste talekode/dekode-systemer som er blitt utviklet i de senere år bruk av LSP som den spektrale informasjon. Den LSP-baserte utførelse av foreliggende oppfinnelse kan lett tilpasses slike typer av talekode/dekode-systemer. Det betyr at på grunn av det faktum at det ikke er noe behov for re-analyse av spekteret og transformasjon av parametrene, kan en god tilkoblingsmulighet oppnås for slike systemer i motsetning til teknikkens stand der filterkoeffisientene blir bestemt på grunnlag av inngående mel-skalert kepstrum som beskrevet i referanse 4.

Det fremgår av den ovenstående beskrivelse at overføringsfunksjonen  $H(z)$  for filteret 203 i den LSP-baserte utførelse av foreliggende oppfinnelse vil avhenge av den måte hvorpå LSP modifikasjonsoperasjonen og LSP/LPC transformeringsoperasjonen utføres for å komme frem til filterkoeffisientene  $\alpha_{1i}$  og  $\alpha_{2i}$ . En foretrukket fremgangsmåte for LSP modifikasjonsoperasjonen er først en proporsjonal delingsmodifikasjon og deretter en tilsluttet dimensjon-til-dimensjon avstandsutvidelse.

Den proporsjonale delingsmodifikasjon som er nevnt først er en fremgangsmåte der  $\omega_i$  blir proporsjonalt delt ved bruk av modifiserte koeffisienter  $\nu, \eta$  som tilfredsstillende  $0 \leq \nu \leq \eta < 1$  som proporsjonale delingsforhold. Når denne fremgangsmåte utføres i konfigurasjonen på Fig. 1, vil hver av modifikasjonsseksjonene 216 og 217 ha en funksjonell konfigurasjon innbefattende en proporsjonalt delende operasjonsseksjon 220 og en gradientstillende seksjon 221 for eksempel som vist på fig. 4. Den proporsjonalt delende operasjonsseksjon 220 frembringer  $\omega h_{1i}$  eller  $\omega h_{2i}$  i overensstemmelse med det følgende uttrykk for proporsjonal deling:

$$\begin{aligned} \omega h_{1i} &= \omega_i \times (1 - \nu) + \omega f_i \times \nu & \text{ eller } & & \dots(6) \\ \omega h_{2i} &= \omega_i \times (1 - \eta) + \omega f_i \times \eta \end{aligned}$$

35

der  $i = 1, 2, \dots, p$ .

Den gradientstillende seksjon 221 stiller  $\omega_{f_i}$  i den proporsjonalt delende operasjonsseksjon 220 på basis av den lineære forutsigelsesorden  $p$ . Det skal påpekes at  $\omega_{f_i}$  som benyttes i LSP modifikasjonsseksjonen 216 kan ha en verdi som er forskjellig fra  $\omega_{f_i}$  i seksjon 217. Modifikasjonen av  $\omega_{f_i}$  ved den proporsjonale deling kan også  
 5 anvendes i konfigurasjonen på fig. 2.

En første fordel med den proporsjonale deling er å sikre en bedre formantforbedrende virkning. Det vil si at når  $\omega_{h1_i}$  og  $\omega_{h2_i}$  som er frembragt ved den proporsjonale deling blir transformert fra LSP verdiområdet til LPC verdiområdet, blir formantene klangløse  
 10 med det resultat at en god formantforbedrende virkning kan oppnås. "Formantene blir klangløse" betyr her at "toppverdiene på formantene blir små", eller med andre ord, "spektrale karakteristika flates ut mens spekteret etterlates med en viss topp-dal struktur".

En andre fordel med den proporsjonale deling er å sikre stor frihet når det gjelder formgivningen for karakteristikkene i overensstemmelse med de krav brukerne har, som for eksempel det å kunne variere graden av modifikasjon av de syntetiserte talesignalene for hvert frekvensbånd. Særlig ved å utforme  $\omega_{f_i}$  ved siden av  $\nu$  og  $\eta$  kan karakteristikkene for filteret 203 varieres slik at de godt tilfredsstillende de krav brukerne  
 20 setter. Denne store frihet når det gjelder utformning vil føre til det resultat at innen et område med tillatelige spektralgradienter kan det lett oppnås en større formantforbedrende virkning som overgår vanlige kjente teknikker.

Det er tatt sikte på at det finnes flere fremgangsmåter til oppsetning av  $\omega_{f_i}$ . En første fremgangsmåte er å stille LSP som representerer et flatt spektrum som  $\omega_{f_i}$ . Den gradientstillende seksjon 221 som er implementert i overensstemmelse med denne fremgangsmåte stiller  $\omega_{f_i}$  på en slik måte at  $\omega_{f_i}$  ved sammenstående dimensjon-til-dimensjon avstand ( $= \omega_{f_i} - \omega_{f_{i-1}}$ ) resulterer i en viss verdi representert som  $\pi / (p + 1)$ , i henhold til det følgende uttrykk  
 30

$$\omega_{f_i} = \pi \times i / (p + 1) \quad \dots(7)$$

Fig. 5 viser idémessig frembringelsen av  $\omega_{h1_i}$  som et eksempel, der operasjonen med modifisering-ved-proporsjonal-deling vil finne sted når  $\omega_{f_i}$  stilles ifølge uttrykket (7).  
 35 Det skal påpekes at det her er antatt at  $p = 10$ . Denne fremgangsmåte har fordelen ved at den er funksjonelt enkel når det gjelder seksjonen 221 for gradientinnstilling.

En andre fremgangsmåte er å stille LSP for et fast gradientspektrum som  $\omega f_i$ . Den gradientinnstillende seksjon 221 som er utformet i overensstemmelse med denne fremgangsmåte stiller  $\omega f_i$  på en slik måte at den tilstøtende dimensjon-til-dimensjon avstand for  $\omega f_i$  lineært øker eller avtar i henhold til det følgende uttrykk som

5 fremkommer ved å tilføye leddet  $\delta(i)$  som i avhenger av til høyre side av uttrykket (7)

$$\omega f_i = \pi \times i / (p + 1) + \delta(i) \quad \dots(7a)$$

I dette tilfelle vil fagfolk på området lett se av den ovenstående beskrivelse og det som

10 er vist på fig. 5 hvorledes den proporsjonsdelende modifikasjonsvirkning finner sted. Denne fremgangsmåte har for det første den fordel at den muliggjør styring av klarhet ved innstilling av proporsjonalkoeffisienten for  $\omega_i$  siden en stort sett fast gradient kan innføres i karakteristikkene for filteret 203. For det annet har den fordelen ved å kunne redusere behandlingsprosedyrene siden overføringsfunksjonen  $H(z)$  for dette filter 203

15 kan inneholde karakteristikkene for en fast høyfrekvent forbedringsprosess som kan utføres omtrent samtidig med den ordinære formantforbedrende prosess. For det tredje har fremgangsmåten fordelen ved å kunne anvendes for å undertrykke klarhetsvariasjonen ved å endre  $\delta(i)$  til  $\delta(\omega_i)$  og ved å modifisere dens funksjonsblokk med den stiplede linje på fig. 4.

20 En tredje fremgangsmåte er å stille som  $\omega f_i$  en LSP som fremkommer ved modifisering av LSP-verdien for et gjennomsnittlig støyspektrum, for eksempel ved hjelp av proporsjonsdeleprosessen. Den gradientinnstillende seksjon 221 som er utformet ifølge denne fremgangsmåte stiller  $\omega f_i$ , som vist på fig. 6, ved modifisering av LSP  $\omega_i'$  som

25 representerer det gjennomsnittlige støyspektrum på basis av det proporsjonale delingsforhold  $\nu'$  eller  $\eta'$ , i henhold til det følgende uttrykk

$$\begin{aligned} \omega f_i &= \omega_i' \times (1 - \nu') + \omega_i' \times \nu' \quad \text{eller} \\ \omega f_i &= \omega_i' \times (1 - \eta') + \omega_i' \times \eta' \quad \dots(7b) \end{aligned}$$

30 der  $i = 1, 2, \dots, p$ .

Fordelen ved denne fremgangsmåte ligger i forbedret oppfatningsmulighet på grunn av evnen til en viss forbedring av talespekteret i stedet for støyspekteret. Det skal nevnes at  $\omega_i'$  kan fremkomme ved gjennomsnittsberegning med en gjennomsnittsooperende

35 seksjon 223, av  $\omega_i$  i en periode som er blitt antatt å være en støyperiode i en vurderingsseksjon 222 som er vist på fig. 6. Det er også fordelaktig at modifikasjonsprosessen som  $\omega_i'$  gjennomgår blir stilt slik at det ikke innføres en altfor

stor spektralvariasjon i de modifiserte syntetiserte talesignaler. Hvis for eksempel  $\omega_i$  ble gjort for uskarp, vil det bli mulig å hindre ekstrem spektralvariasjon i å oppstå i de modifiserte syntetiserte talesignaler.

- 5 En fjerde fremgangsmåte er å stille som  $\omega_i$  en LSP som fremkommer ved modifisering for eksempel ved hjelp av den proporsjonale deleprosess, av en gjennomsnittlig verdi for  $\omega_i$  under en periode opp til nå etter start av prosessen eller i løpet av en tidligere forhåndsbestemt periode. Som vist på fig. 7 vil den gradientinnstillende seksjon 221 som er utformet ifølge denne fremgangsmåte finne en gjennomsnittsverdi  $\omega_i'$  for den
- 10 tidligere LSP  $\omega_i$  ved hjelp av den gjennomsnittsberegneende seksjon 223 og stiller  $\omega_i'$  på basis av denne  $\omega_i'$  og det proporsjonale delingsforhold  $\nu'$  eller  $\eta'$  og i henhold til uttrykket (7b). Fordelen ved denne fremgangsmåte ligger i den forbedrede oppfatningsmulighet som kan føres tilbake til evnen til å forbedre variasjoner i talespekteret. Det er også fordelaktig for utførelse av denne fremgangsmåte at det tas i
- 15 betraktning for eksempel å modifisere  $\omega_i'$  på en slik måte at det ikke innføres spektrale variasjoner som er for store i de modifiserte syntetiserte talesignaler.

- På fig. 8 er det vist log-potens i forhold til frekvensspektrumkarakteristikkene for filteret 203 som er vist på fig. 1, og som fremkommer når  $\omega_i$  blir modifisert ifølge uttrykkene
- 20 (6) og (7). På grafen representerer henholdsvis A, B, C og D syntetisatorens 202 karakteristikk  $= 1 / A(z)$ , filterets 204 karakteristikk  $= 1 / A_1(z)$ , filterets 205 invers-karakteristikk  $= 1 / A_2(z)$ , og filterets 203 overføringsfunksjon  $H(z) = A_2(z) / A_1(z)$  med  $\nu = 0,5$  og  $\eta = 0,8$ . Som vist på denne graf, er karakteristikken D på grafen flatet ut samtidig med at den etterlater seg spektrumets topp-dal struktur i en viss
- 25 utstrekning sammenlignet med karakteristikken D på fig. 33. På fig. 8 kan det på denne måte sees en større formantforbedring sammenlignet med fig. 33. Dessuten oppviser karakteristikken D på denne graf mindre forvrengninger med hensyn på spektrumets topp-dal struktur enn karakteristikkene D på fig. 34. Videre vil karakteristikken D på denne graf ikke lenger oppvise de to fenomener som er blitt observert i
- 30 karakteristikkene B og C på fig. 35, dvs. forskyvning av formanter med lavest frekvens og integrasjon av to formanter på midten. Som et alternativ til den proporsjonale delingsprosess kan andre prosesser som har en virkning der formanter i LSP verdiområdet blir uskarpe, anvendes for å oppnå tilsvarende fordeler.

- 35 Oppfinneren i denne sak har hørselsmessig sammenlignet den modifiserte syntetiserte tale som fåes fra filteret 203 i denne utførelse med modifisering av  $\omega_i$  i henhold til den fremgangsmåte som er representert av uttrykkene (6) og (7), med den modifiserte

syntetiserte tale som avledes fra filteret 203 i tidligere kjente utførelser som er beskrevet. Som et resultat har det vist seg at talemodifikasjonsfilteret i denne utførelse byr på en fordel sammenlignet med tidligere kjente filtere når det gjelder undertrykkelse av degradering av klarhet og at dette ikke fører til noen egenartet forvrengt tale eller

5 fluktuerende tone.

Avstandsutvidelsen for sammenstående dimensjon-til-dimensjon som er en andre foretrukket utførelse av den LSP-modifiserende operasjon kan utføres av en utvidelsesseksjon 224 og en ensartet kompresjonsseksjon 225 som vist på fig. 9.

10 Utvidelsesseksjonen 224 frembringer  $s_i$  ved forflytning av  $\omega_i$ , der både  $s_i$  og  $\omega_i$  tilhører LSP verdiorråde, slik at den sammenstående dimensjon-til-dimensjon avstand  $s_i - s_{i-1}$  kan gjøres større enn den tilstøtende dimensjon-til-dimensjon avstand  $\omega_i - \omega_{i-1}$  (når det gjelder  $\omega_i - \omega_{i-1}$ , se fig. 5). Seksjonen 225 for ensartet kompresjon finner  $\omega h_{1i}$  fra  $s_i$ . Det skal særlig påpekes at  $s_i$ , så vel som  $\omega_i$ , er en flerdimensjonal vektor. Når denne

15 fremgangsmåte utføres i konfigurasjonen på fig. 2, finner seksjonen 225 for ensartet kompresjon  $\omega h_{1i}$  etter det følgende uttrykk

$$\omega h_{1i} = s_i / s_{p+1} \times \pi \quad \dots(8)$$

20 og utvidelsesseksjonen 224 finner  $s_i$  i henhold til det følgende uttrykk

$$s_i = s_{i-1} + \text{maks}(\omega_i - \omega_{i-1}, th) \quad \dots(9)$$

der  $i = 1, 2, \dots, p + 1$

25  $\omega_0 = 0, \omega_{p+1} = \pi, s_0 = 0$

$th$ : terskelverdi.

Slik det fremgår av de ovenfor beskrevne uttrykk (8) og (9) er avstandsutvidelsen ved sammenstående dimensjon-til-dimensjon en prosess for å sikre en minste avstand  $th$

30 mellom den  $(i - 1)$ te dimensjon og den  $i$ -te dimensjon fra resultatet av sammenligning mellom  $\omega_i - \omega_{i-1}$  med  $th$ , som særlig angitt med det andre ledd på høyre side av uttrykket (9). Denne prosess lar LSP som er knyttet til  $(i + 1)$ te eller øvre dimensjoner forskyve seg sammen oppad over en avstand svarende til  $th - (\omega_i - \omega_{i-1})$ . Dessuten er faktoren  $\pi / s_{p+1}$  som inneholdes i den høyre side av uttrykket (8) en faktor for ensartet

35 kompresjon av de sammenstående dimensjon-til-dimensjon avstander som svar på forholdene i  $\omega_i$  området 0 til  $\pi$  og i  $s_i$  området 0 til  $s_{p+1}$  for linjespektrumparet LSP. Det skal påpekes at foreliggende oppfinnelse ikke skal betraktes som begrenset av dette

definerende uttrykk og at andre definerende uttrykk kan anvendes så lenge de representerer prosesser til utvidelse av små sammenstående dimensjon-til-dimensjon avstander. Dessuten kan  $\omega_1$  ved utvidelsen av sammenstående dimensjon-til-dimensjon avstander anvendes i konfigurasjonen på fig. 1. Dette vil muliggjøre ytterligere økning  
 5 av friheten til utformning av karakteristikkene for filteret 203.

Fig. 10 viser log-potens i forhold til frekvensspektrumkarakteristikkene som vil fremkomme når denne fremgangsmåte anvendes på filteret 203 på fig. 2. I denne graf representerer A, B og C henholdsvis syntetisatorens 202 karakteristikker =  $1 / A(z)$ ,  
 10 filterets 204 ( $th = 0,3$ ) karakteristikker =  $1 / A_1(z; th = 0,3)$  og filterets 204 ( $th = 0,4$ ) karakteristikker =  $1 / A_1(z; th = 0,4)$ . Som det fremgår av denne graf, gjør denne fremgangsmåte det mulig å anvende karakteristikker som er sammenlignbare med fig. 33 og fig. 34 bare i filteret 204 (med andre ord, uten å bruke filteret 205 eller en annen bestanddel som svarer til dette). Dette betyr at et godt talemodifikasjonsfilter kan  
 15 utformes med en lavere ordens filter enn i kjente filtere og at stort samme formantforbedrende virkning som ved vanlige filtere kan virkeliggjøres med et mindre antall bestanddeler. Videre har oppfinneren i foreliggende sak hørselsmessig sammenlignet den modifiserte syntetiserte tale som fremkommer i den utførelse med det som oppnås med tradisjonelle teknikker. Som resultat har det vist seg at bruk av  
 20 talemodifikasjonsfilteret i denne utførelse vil sikre en tonekvalitet som ikke på noen måte står tilbake for de eksisterende filtere.

De to typer fremgangsmåter til modifikasjon, dvs. den proporsjonale delingsmodifikasjon og utvidelse av sammenstående dimensjon-til-dimensjon utelukker ikke  
 25 hverandre og kan således benyttes sammen. Det er også for eksempel tenkelig at en av LSP modifikasjonsseksjonene 216 og 217 utfører den proporsjonale deling, mens den andre har styring med den tilstøtende dimensjon-til-dimensjon utvidelse. Som et alternativ kan det, som vist på fig. 11, benyttes en utformning som innbefatter vekslingsanordninger 228 og 229 for valgvis bruk av den proporsjonale  
 30 delingsmodifiserende seksjon 226 som tjener til å modifisere  $\omega_1$  ved den proporsjonale deling og seksjonen 227 for utvidelse av dimensjon-til-dimensjon avstanden som tjener til å utvide avstandene ved sammenstøtende dimensjon-til-dimensjon for LSP. Modifikasjonsseksjonen 226 for proporsjonal deling kan ha en hvilken som helst av de ovenfor beskrevne utformninger som er vist på fig. 4, 6 og 7. Som et alternativ kan det,  
 35 som vist på fig. 12, anvendes en utformning der modifikasjonsseksjonen 226 for proporsjonal deling er koblet i kaskade med den tilstøtende utvidelsesseksjon 227 for sammenstøtende dimensjon-til-dimensjon avstand. På grunn av at slike utformninger

har en enkel LSP modifikasjonsseksjon som tjener både som modifikasjonsseksjonen 226 for proporsjonal deling og utvidelsesseksjonen 227 for tilstøtende dimensjon-til-dimensjon avstand, vil man få ytterligere øket frihet når det gjelder utformningen av filteret 203. Det kan også tenkes at sekvensen med modifikasjonsseksjonen 226 for  
 5 proporsjonal deling og utvidelsesseksjonen 227 for sammenstøtende dimensjon-til-dimensjon avstand reverseres som vist på Fig. 12. Det er naturlig at andre prosesser kan kombineres med begge eller enten den proporsjonale delingsmodifikasjon eller avstandsutvidelsen ved sammenstøtende dimensjon-til-dimensjon.

10 Videre kan en  $\omega_j$  adaptiv prosess utføres av LSP modifikasjonsseksjonene 216 og 217. En fremgangsmåte som kan tenkes for å gjøre den proporsjonale delingsbaserte  $\omega_j$  modifikasjonsprosess  $\omega_j$  adaptiv, er for eksempel en fremgangsmåte der et  $\omega_j$  rom blir delt opp i en rekke delrom (i det følgende betegnet som kategorier) som ikke overlapper hverandre og der  $\nu$  og  $\eta$  er klargjort (eller vekslet) for hver kategori. I dette tilfelle kan  
 15 LSP modifikasjonsseksjonen anordnes for hver kategori, for eksempel en LSP modifikasjonsseksjon 216-1 (eller 217-1) som svarer til en første kategori, en LSP modifikasjonsseksjon 216-2 (eller 217-2) svarende til en andre kategori, ... og en LSP modifikasjonsseksjon 216-N (eller 217-N) som svarer til en N-te kategori (se fig. 13). Som et alternativ kan en enkel LSP modifikasjonsseksjon 216 (eller 217) bygges opp  
 20 sammen med en modifisert koeffisientvekslende seksjon 230 som tjener til å veksle  $\nu$  og  $\eta$  som svar på kategoriene eller  $i$  (se fig. 14). Den  $\omega_j$  adaptive prosess har fordelen ved å kunne skape en fleksibel prosess som for eksempel gjør det mulig å svekke formantforbedringen bare for en bestemt kategori, som for eksempel en kategori som skaper forvrengninger når formantforbedringen økes. Dette ville sikre en ensartet eller  
 25 forvrengningsfri forbedring i karakteristikene for filteret 203. Det skal påpekes at siden  $\omega_j$  er en flerdimensjonal vektor, er kategorien det her vises til i alminnelighet et flerdimensjonalt vektorrom.

Det er å foretrekke at  $\omega_j$  modifiseringsprosessen i LSP modifikasjonsseksjonene 216 og  
 30 217 implementeres ved bruk av en omsettingstabell 231 som vist på fig. 15. Mer bestemt settes omsettingstabellen 231 opp for å korrelere  $\omega_j$  med  $\omega_{h1_j}$  eller  $\omega_{h2_j}$  slik at det blir mulig for LSP modifikasjonsseksjonen 216 eller 217 å frembringe  $\omega_{h1_j}$  eller  $\omega_{h2_j}$  som sin utgang når det gjelder  $\omega_j$ . Fordelen ved å benytte omsettingstabellen 231 ligger i en reduksjon av behandlingstiden. Denne fordel vil bli mer eller mindre merkbar  
 35 hvis et forholdsvis komplisert uttrykk benyttes som et prinsipielt uttrykk for  $\omega_j$  modifikasjonsprosessen.

$\omega_i$  modifikasjonsprosessen i LSP modifikasjonsseksjonene 216 og 217 kan utformes med et nevralt nett 232 som har på forhånd innlærte  $\omega_i$  modifikasjonskarakteristikker fastlagt for eksempel med uttrykket (6) som vist på fig. 16. En første fordel ved bruk av nevraltnett 232 ligger i en reduksjon av behandlingstiden. Denne fordel vil bli mer

5 merkbar hvis et forholdsvis komplisert uttrykk blir benyttet som det prinsipielle uttrykk for  $\omega_i$  modifikasjonsprosessen. En annen fordel ved å bruke nevraltnett 232 ligger i at en minnekapasitet kan reduseres på grunn av det faktum at det ikke er noe behov for å lagre omsettingstabellen 231 sammenlignet med det tilfelle da omsettingstabell 231 benyttes.

10

En tredje fordel ved bruk av nevraltnett 232 ligger i reduksjonen av forvrengning. I for eksempel  $\omega_i$  adaptive utførelser som er vist på figurene 13 og 14 opptrer forvrengninger ofte ved grensene ved kategorier i det modifiserte eller halvmodifiserte syntetiserte talesignal på grunn av brå endringer av  $\nu$  og  $\eta$  som opptrer på grunn av en svak

15 variasjon av  $\omega_i$  forbi kategoriens grenser. Forvrengningene har tilbøyelighet til å bli merkbare, særlig når delingen av  $\omega_i$  rommet er forholdsvis grov. I utførelsen med omsettingstabell som vist på fig. 15 opptrer forvrengninger ofte ved en grense for tabelladressen på samme måte som i utførelsene på figurene 13 og 14. I motsetning til dette oppstår det i nevraltnettutførelsene som er vist på fig. 16 ingen forvrengninger, siden

20 det ikke finnes noen kategori som fører til brå endring i  $\nu$  og  $\eta$ .

Den LSP-baserte utførelse av foreliggende oppfinnelse skal ikke være begrenset til den utformning som foretar LPC filtrering og invers-LPC filtrering, og som vil tillate andre parametere enn LPC å bli benyttet som sine filterkoeffisienter. For eksempel, som vist

25 på figurene 17 og 18, kunne foreliggende oppfinnelse utføres ved bruk av et LSP filter 233 (og et invers-LSP filter 234) som bruker filterkoeffisienten  $\omega h_{1i}$  (og  $\omega h_{2i}$ ) slik de er. Fordelen ved denne utformning ligger i at det ikke er noe behov for LSP/LPC transformseksjonene 218 og 219.

30

Det vises nå til fig. 19 der det er inngitt en utførelse som innfører PARCOR som spektralinformasjon. Denne utførelse omfatter PARCOR modifikasjonsseksjoner 235 og 236 og PARCOR/LPC transformseksjoner 237 og 238 i tillegg til LPC-filte

35 retet 204 og det inverse-LPC filter 205. PARCOR modifikasjonsseksjonen 235 innfører PARCOR  $\phi_i$  som spektralinformasjonen fra dekode

ren 201 eller transformseksjonen 215 og modifiserer denne  $\phi_i$  for å frembringe modifisert PARCOR  $\phi h_{1i}$ . På samme måte frembringer PARCOR modifikasjonsseksjonen 236 modifisert PARCOR  $\phi h_{2i}$ . PARCOR/LPC transformseksjonen 237 transformerer  $\phi h_{1i}$  fra et PARCOR

verdiområde til et LPC verdiområde for å skape en filterkoeffisient  $\alpha_{1i}$  for LPC-filteret 204. PARCOR/LPC transformseksjonen 238 transformerer også  $\phi_{h2i}$  fra PARCOR verdiområdet til LPC verdiområdet for å frembringe en filterkoeffisient  $\alpha_{2i}$  for det inverse-LPC filter 205.

5

PARCOR modifikasjonsseksjonene 235 og 236 frembringer henholdsvis  $\phi_{h1i}$  og  $\phi_{h2i}$  ved bruk av modifiserte koeffisienter  $\nu$  og  $\eta$  som tilfredsstillers for eksempel  $0 \leq \eta \leq \nu < 1$  og i henhold til de følgende uttrykk

$$\begin{aligned} \phi_{h1i} &= \phi_i \times \nu^{(i \times i)} \\ \phi_{h2i} &= \phi_i \times \eta^{(i \times i)} \end{aligned} \quad \dots(10)$$

der  $i = 1, 2, \dots, p$ .

15 Utførelse av denne modifikasjon gjør formanter klangløse i PARCOR verdiområdet.

Som en følge av dette vil denne utførelse sikre den samme karakteristiske forbedringsvirkning som den ovenstående LPC-baserte utførelse (f.eks. formantforbedrende virkning og forbedring av evnen til å justere graden av 20 forbedringen) så vel som fri styring/innstilling av karakteristikkene for filteret 203 etter de behov brukerne har. Det er naturlig at foreliggende oppfinnelse ikke skal betraktes som begrenset av uttrykket (10) og at andre prosesser kan anvendes som gjør formantene klangløse i PARCOR verdiområdet. Videre, når det gjelder filteret som bruker som sin filterkoeffisient PARCOR eller den parameter som frembringes på basis 25 av PARCOR, er det forholdsvis enkelt å påvise og sikre stabilitet på PARCOR verdiområdet, siden stabilitetstilstanden gis av den følgende enkle ligning:

$$- 1 < \phi_i < 1 \quad \dots(11)$$

30 Med andre ord, så lenge ligning (11) er tilfredstilt, er filteret som bruker PARCOR-basert filterkoeffisient stabilt. Ifølge denne utførelse blir derfor friheten når det gjelder filterutforming forbedret. For eksempel kan man bruke som PARCOR modifikasjonsprosess prosessen med å modifisere PARCOR  $\phi_i$  uavhengig for de respektive  $i$ . I tillegg vil anvendelse i systemer for overføring eller lagring av PARCOR som spektral- 35 informasjon sikre en god tilkoblingsmulighet på grunn av det faktum at det ikke er nødvendig med spektrum re-analyse og parametertransform. Fig. 20 viser grafisk log-potens i forhold til frekvensspektrumkarakteristikker for filteret 203 på fig. 19. På

grafen betegner A, B, C og D henholdsvis syntetisatorens 202 karakteristikk  $= 1 / A(z)$ , filterets 204 karakteristikk  $= 1 / A1(z)$ , filterets 205 invers-karakteristikk  $= 1 / A2(z)$ , og filterets 203 karakteristikk  $= A2(z) / A1(z)$ , med  $\nu = 0,98$  og  $\eta = 0,9$ . Som det fremgår av sammenligningen mellom figurene 20 og 33, gjør denne utførelse det

5 mulig for spektrumets topp-dal struktur å opptre mer eller mindre sterkt enn i den utførelse som er vist i referanse 1. Ved hørselsmessige sammenligninger av den modifiserte syntetiserte tale har oppfinneren i foreliggende sak fastslått at bruk av filteret 203 i denne utførelse helt bestemt ikke vil føre til noen særegen forvrengt tale eller noen fluktuerende tone og vil sikre en god formantforbedrende virkning.

10

Det skulle være klart for fagfolk på området ut fra denne beskrivelse at detaljene ved den PARCOR-baserte utførelse kan utvikles fra samme synspunkt som den LSP-baserte utførelse. For fagfolk på området skulle det også være lett å tenke seg på grunnlag av den foreliggende beskrivelse å utelukke invers-LPC filtrering og bestanddeler som er

15 knyttet til dette som vist på fig. 21 og å anvende en utformning som innbefatter et PARCOR-filter 239 og et invers-PARCOR filter 240 med modifisert PARCOR  $\phi h_{1i}$  og  $\phi h_{2i}$  benyttet som dets filterkoeffisienter som vist på fig. 22.

20

En utførelse som innfører LAR som spektralinformasjon er gjengitt på fig. 23. Denne utførelse omfatter i tillegg til LPC-filteret 204 og det inverse-LPC filter 205 LAR-modifikasjonsseksjonene 241 og 242 og LAR/LPC transformseksjonene 243 og 244. LAR-modifikasjonsseksjonen 241 innfører LAR  $\Psi_i$  som spektralinformasjon fra dekoderen 201 eller transformseksjonen 215 og modifiserer denne  $\Psi_i$  for å frembringe modifisert LAR  $\Psi h_{1i}$ . På samme måte frembringer også LAR-modifikasjonsseksjonen

25 242 modifisert LAR  $\Psi h_{2i}$ . LAR/LPC transformseksjonen 243 transformerer  $\Psi h_{1i}$  fra LAR verdiområdet til LPC verdiområdet for å frembringe en filterkoeffisient  $\alpha_{1i}$  for LPC-filteret 204. LAR/LPC transformseksjonen 244 transformerer  $\Psi h_{2i}$  fra LAR verdiområdet til LPC verdiområdet for å frembringe en filterkoeffisient  $\alpha_{2i}$  for det inverse-LPC filter 205.

30

LAR-modifikasjonsseksjonene 241 og 242 frembringer  $\Psi h_{1i}$  resp.  $\Psi h_{2i}$  ved bruk av modifiserte koeffisienter  $\nu$  og  $\eta$  som tilfredsstillers for eksempel  $0 \leq \eta \leq \nu < 1$ , og i henhold til de følgende uttrykk

35

$$\begin{aligned} \Psi h_{1i} &= \Psi_i \times \nu^i \\ \Psi h_{2i} &= \Psi_i \times \eta^i \end{aligned} \quad \dots(12)$$

der  $i = 1, 2, \dots, p$ .

Utførelse av en slik modifikasjon lar formantene bli klangløse på PARCOR verdiområdet.

5

Som en følge av dette vil denne utførelse sikre den samme forbedringsvirkning av karakteristikkene som den man får med den ovenfor beskrevne LPC-baserte utførelse og den PARCOR-baserte utførelse (f.eks. formantforbedrende virkning og forbedring av muligheten for å justere graden av forbedring) så vel som fri styring/innstilling av karakteristikkene for filteret 203 alt etter de krav brukerne har. Det er naturlig at foreliggende oppfinnelse ikke skal anses å være begrenset av uttrykket (12) og at andre prosesser kan anvendes som lar formantene bli klangløse i LAR verdiområdet. Siden det er fastslått og sikret at filteret er stabilt når det benyttes filterkoeffisienter som er frembragt på basis av LAR, er LAR-modifikasjonsprosessen i denne utførelse ikke begrenset til trekket som gjelder filterets stabilitet. Av den grunn er friheten når det gjelder formgivning av filteret i denne utførelse større enn for tidligere kjente anordninger. I tillegg vil bruk av systemene som overfører eller lagrer PARCOR som spektralinformasjon sikre god tilkoblingsmulighet på grunn av det faktum at det ikke er noe behov for spektrum re-analyse og parametertransform.

20

Fig. 24 viser grafisk log-potens i forhold til frekvensspektrumkarakteristikk for filteret 203 på fig. 23. I grafen betegner A, B, C og D henholdsvis syntetisatorens 202 karakteristikk  $= 1 / A(z)$ , filterets 204 karakteristikk  $= 1 / A_1(z)$ , filterets 205 invers-karakteristikk  $= 1 / A_2(z)$ , og filterets 203 karakteristikk  $= A_2(z) / A_1(z)$ , der  $\nu = 0,9$  og  $\eta = 0,7$ . Sammenligning mellom figurene 24 og 33 har vist at denne utførelse lar spekteret bli flatet ut samtidig med at det blir tilbake en spektrum topp-dal struktur i en viss utstrekning, noe som resulterer i en bedre formantforbedrende virkning sammenlignet med konfigurasjonen som er beskrevet i referanse 1. I sammenligning med fig. 34 viser fig. 24 dessuten mindre forvrengninger i topp-dal strukturen av spekteret. På fig. 24 fremkommer ikke lenger fenomenet med integrering av to formanter på midten, noe som vil fremgå av sammenligningen mellom karakteristikkene B og C på fig. 35. Ved hørselsmessige sammenligninger av modifisert syntetisert tale har oppfinneren i foreliggende sak fastslått at bruk av filteret 203 i denne utførelse helt tydelig ikke skaper noen særegen forvrengt tale eller noen svingende tone, og vil sikre god formantforbedrende virkning.

35

Det skulle ut fra denne beskrivelse være klart for fagfolk på området at detaljer ved denne LAR-baserte utførelse kan bygges opp ut fra samme synspunkt som den LSP-baserte utførelse og den PARCOR-baserte utførelse. Det skulle også være lett å se på grunnlag av denne beskrivelse for fagfolk på området å utelukke invers-LPC filtrering og tilhørende bestanddeler som vist på fig. 26 og å anvende en konfigurasjon som innbefatter et PARCOR-filter 239 og invers-PARCOR filter 240 med modifiserte LAR  $\Psi_{h1_i}$  og  $\Psi_{h2_i}$  benyttet som filterkoeffisienter. For å transformere den modifiserte LAR  $\Psi_{h1_i}$  og  $\Psi_{h2_i}$  fra LAR verdiområde til PARCOR verdiområde synes det dessuten på fig. 26 LAR/PARCOR-transformerende seksjoner 246 og 247. Siden LAR/PARCOR-transformeringsprosessen stort sett er forholdsvis enkel og lett å utføre sammenlignet med LAR/LPC-transformering, kan LAR/PARCOR-transformeringsseksjonene 246 og 247 utføres med færre prosesstrinn eller med mindre kretser enn LAR/LPC-transformeringsseksjonene 243 og 244. Ifølge utførelsen på fig. 27 er derfor filterkoeffisientene  $\alpha_{1_i}$  og  $\alpha_{2_i}$  utledet med kortere periode enn, og hele prosessen med filteret 203 er redusert fra, utførelsesformene på fig. 23 og fig. 25.

På grunnlag av denne beskrivelse skulle det for fagfolk på området være lett selektivt å kombinere den ovenfor beskrevne LSP-baserte utførelse, PARCOR-baserte utførelse og LAR-baserte utførelse. Det skulle dessuten på grunnlag av denne beskrivelse for fagfolk på området være enkelt å kombinere hver utførelse av foreliggende oppfinnelse med vanlig LPC-basert utstyr. Disse forskjellige kombinasjoner bidrar til virkeliggjørelse av et filter 203 som har høy frihet når det gjelder utformning av karakteristikkene som ikke kunne virkeliggjøres på annen måte. For eksempel, som vist på fig. 27, kan filterkoeffisienten  $\alpha_{1_i}$  for filteret 204 defineres med samme fremgangsmåte som referanse 1 mens filterkoeffisienten  $\alpha_{2_i}$  for filteret 205 kan defineres med samme fremgangsmåte som den PARCOR-baserte utførelse. Denne utformning vil føre til et filter 203 som har en lavere spektralgradient enn karakteristikkene D for fig. 33 og mindre forvrengninger i området ved formantene enn karakteristikkene D på fig. 34.

Foran eller bak filteret 203 eller parallelt med filteret 203 kan det være anbragt et annet filter for å foreta behandling som forbedrer tonehøyde, behandling som forbedrer høyfrekvens, behandling som forbedrer formanter, etc.

P a t e n t k r a v

1.

Filter, k a r a k t e r i s e r t v e d at det omfatter:

- 5 filteringsanordninger for filtrering av syntetiserte talesignaler med en overføringsfunksjon definert med filterkoeffisienter for å frembringe modifiserte syntetiserte talesignaler; og
- anordning til frembringelse av filterkoeffisienter på basis av den spektralinformasjon som er representert i form av en flerdimensjonal vektor og tilhører et på forhånd bestemt
- 10 verdiområde og gjelder innmatede talesignaler, på en slik måte at formantkarakteristikkene for det modifiserte syntetiserte talesignal blir forbedret ifølge den spektrale informasjon og i sammenligning med de for de syntetiserte talesignaler, hvilken spektrale informasjon kan være en hvilken som helst av LSP-informasjon, PARCOR-informasjon og LAR-informasjon.

15

2.

Filter som angitt i krav 1, k a r a k t e r i s e r t v e d at filterkoeffisientene tilhører et LPC verdiområde.

20

3.

Filter som angitt i krav 2, k a r a k t e r i s e r t v e d at anordningen til frembringelse av filterkoeffisient innbefatter:

- en modifikasjonsanordning for modifisering av spektralinformasjonen i det på forhånd bestemte verdiområde for å frembringe modifisert spektralinformasjon; og
- 25 anordning for transformering av den modifiserte spektralinformasjon fra det på forhånd bestemte verdiområde inn i et LPC verdiområde for å frembringe filterkoeffisientene.

4.

Filter som angitt i krav 3, k a r a k t e r i s e r t v e d at

- 30 modifikasjonsanordningen innbefatter utflatningsanordning til modifisering av spektralinformasjonen for å redusere toppe i formantene av de modifiserte syntetiserte talesignaler.

5.

- 35 Filter som angitt i krav 4, k a r a k t e r i s e r t v e d at spektralinformasjonen er LSP-informasjon, og at utflatningsanordningen innbefatter en proporsjonal delingsanordning for proporsjonal deling i henhold til en modifisert

koeffisient, hvilken spektralinformasjon og referanseinformasjon tilhører ett og samme verdiorråde som spektralinformasjonen hører til for å frembringe den modifiserte spektralinformasjon.

5 6.

Filter som angitt i krav 5, k a r a k t e r i s e r t v e d a t proporsjonaldelingsanordningen proporsjonalt deler spektralinformasjon og referanseinformasjonen for å innføre en fast spektralgradient i de modifiserte syntetiserte talesignaler.

10

7.

Filter som angitt i krav 5, k a r a k t e r i s e r t v e d a t proporsjonaldelingsanordningen proporsjonalt deler spektralinformasjonen og referanseinformasjonen for i de modifiserte syntetiserte talesignaler å innføre en spektrumgradient som viser et gjennomsnittlig støyspektrum.

15

8.

Filter som angitt i krav 5, k a r a k t e r i s e r t v e d a t proporsjonaldelingsanordningen proporsjonalt deler spektralinformasjonen og referanseinformasjonen for i det modifiserte syntetiserte talesignal å innføre en spektrumgradient som viser en historie som spektralinformasjonen har fulgt så langt.

20

9.

Filter som angitt i krav 4, k a r a k t e r i s e r t v e d a t spektralinformasjonen er enten PARCOR-informasjon eller LAR-informasjon, og at utflatningsanordningen innbefatter anordning for multiplisering, for hver av en rekke dimensjoner som utgjør spektralinformasjonen, av spektralinformasjonen med en modifisert koeffisient eller med potensen for den modifiserte koeffisient for å frembringe den nevnte modifiserte spektralinformasjon.

30

10.

Filter som angitt i krav 9, k a r a k t e r i s e r t v e d a t potensen er avhengig av den nevnte dimensjon.

35

11.

Filter som angitt i krav 3, k a r a k t e r i s e r t v e d a t spektralinformasjonen er LSP-informasjon, og at modifikasjonsanordningen innbefatter

en avstandsutvidende anordning for utvidelse av avstandene mellom sammenstående dimensjoner blant en flerhet av dimensjoner som representerer spektralinformasjonen for dermed å frembringe den modifiserte spektralinformasjon.

5 12.

Filter som angitt i krav 11, k a r a k t e r i s e r t v e d at den avstandsutvidende anordning innbefatter:

utvidelsesanordning for utvidelse av avstanden ut over referanseavstanden når avstandene mellom sammenstående dimensjoner er mindre enn en referanseavstand; og

10 komprimeringsanordning for lik komprimering av avstandene i forhold til alle de sammenstående dimensjoner etter utvidelsen av avstandene mellom sammenstående dimensjoner med utvidelsesanordningen, for dermed å sikre at utstrekningen av den spektrale informasjon i sin helhet faller sammen med utstrekningen før utvidelsen.

15 13.

Filter som angitt i krav 3, k a r a k t e r i s e r t v e d at spektralinformasjonen er LSP-informasjon og ved at modifikasjonsanordningen innbefatter:

en proporsjonal delingsanordning for proporsjonal deling, i overensstemmelse med en

20 modifisert koeffisient, av spektralinformasjonen og referanseinformasjonen som tilhører ett og samme verdiområde som spektralinformasjonen hører til;

avstandsutvidende anordning for utvidelse av avstandene mellom sammenstående dimensjoner blant en flerhet av dimensjoner som representerer spektralinformasjonen; og

25 vekslingsanordning for selektiv bruk av enten proporsjonaldelingsanordningen eller avstandsutvidelsesanordningen for å frembringe den modifiserte spektralinformasjon.

14.

Filter som angitt i krav 3, k a r a k t e r i s e r t v e d at spektralinformasjonen er LSP-informasjon og ved at modifikasjonsanordningen innbefatter:

en proporsjonal delingsanordning for proporsjonal deling av spektralinformasjonen og referanseinformasjonen som tilhører et og samme verdiområde som

spektralinformasjonen hører til i overensstemmelse med en modifisert koeffisient;

35 avstandsutvidende anordning for utvidelse av avstandene mellom sammenstående dimensjoner blant en flerhet av dimensjoner som representerer spektralinformasjonen; og

kaskadekobleanordning til bruk av både proporsjonaldelingsanordninger og avstandsutvidelsesanordninger i samvirkning for å frembringe den modifiserte spektralinformasjon.

5 15.

Filter som angitt i krav 3, k a r a k t e r i s e r t v e d a t modifikasjonsanordningen innbefatter en omsettingstabell for lagring av spektralinformasjonen i korrelasjon med den modifiserte spektralinformasjon, hvilken omsettingstabell frembringer den modifiserte spektralinformasjon som skal frembringes  
10 som reaksjon på tilførsel av spektralinformasjonen.

16.

Filter som angitt i krav 3, k a r a k t e r i s e r t v e d a t modifikasjonsanordningen innbefatter et nevralt nett som ved innlæring har fått en evne  
15 til å transformere spektralinformasjonen til den modifiserte spektralinformasjon, hvilket nevralnett frembringer modifisert spektralinformasjon som skal frembringes som resultat av tilførsel av spektralinformasjonen.

17.

20 Filter som angitt i krav 3, k a r a k t e r i s e r t v e d a t modifikasjonsanordningen innbefatter:  
en flerhet av kategorispesifikke modifikasjonsanordninger hver innrettet for hver av en flerhet av kategorier som ikke overlapper hverandre og som er fremkommet ved klassifisering av det på forhånd bestemte verdiområde;  
25 der hver av flerheten av kategorispesifikke anordninger innbefatter:  
anordning for modifisering av spektralinformasjonen i en tilsvarende kategori for å frembringe modifisert spektralinformasjon; og  
anordning for transformering av den modifiserte spektralinformasjon fra det på forhånd bestemte verdiområde til LPC verdiområde for å frembringe en filterkoeffisient.

30

18.

Filter som angitt i krav 3, k a r a k t e r i s e r t v e d a t modifikasjonsanordningen innbefatter:  
anordning for modifisering, ifølge en modifisert koeffisient, av spektralinformasjonen i  
35 det på forhånd bestemte verdiområde for å frembringe modifisert spektruminformasjon;

anordning for transformering av den modifiserte spektruminformasjon fra det på forhånd bestemte verdiområde til et LPC verdiområde for å frembringe filterkoeffisientene; og

anordning for justering av den modifiserte koeffisient alt etter hvilken kategori  
5 spektralinformasjonen hører til blant flerheten av kategorier, som er fremkommet ved deling av det på forhånd bestemte verdiområde og som ikke overlapper hverandre.

19.

Filter som angitt i krav 1, k a r a k t e r i s e r t v e d a t  
10 filterkoeffisientene tilhører et hvilket som helst av et LSP verdiområde og et PARCOR verdiområde.

20.

Filter som angitt i krav 19, k a r a k t e r i s e r t v e d a t  
15 anordningen til frembringelse av filterkoeffisient innbefatter: modifikasjonsanordning for modifisering av spektralinformasjonen i det på forhånd bestemte verdiområde for å frembringe modifisert spektralinformasjon; og anordning for tilførsel av den modifiserte spektralinformasjon som filterkoeffisientene til filtreringsanordningen.

20

21.

Filter som angitt i krav 1, k a r a k t e r i s e r t v e d a t  
filtreringsanordningen innbefatter et syntesefilter for implementering av nevneren i overføringsfunksjonen for å sikre at formantkarakteristikkene i de modifiserte  
25 syntetiserte talesignaler blir forbedret sammenlignet med de fra syntetiserte talesignaler.

22.

Filter som angitt i krav 21, k a r a k t e r i s e r t v e d a t  
filtreringsanordningen videre innbefatter et invers filter for undertrykkelse av en  
30 spektralgradient som er innført i de modifiserte syntetiserte talesignaler av syntesefilteret.

23.

Talesyntetiserende anordning, k a r a k t e r i s e r t v e d a t den  
35 omfatter:

anordning til frembringelse av syntetiserte talesignaler på basis av spektralinformasjon som er representert i form av en flerdimensjonal vektor og tilhører et på forhånd bestemt verdiområde og gjelder innmatede talesignaler;

anordning for filtrering av syntetiserte talesignaler ved hjelp av en overføringsfunksjon som er definert av filterkoeffisienter for å frembringe modifiserte syntetiserte  
5 talesignaler; og

anordning til frembringelse av filterkoeffisienter på basis av spektralinformasjonen på en slik måte at formantkarakteristikkene for de modifiserte syntetiserte talesignaler blir forbedret i overensstemmelse med spektralinformasjonen og i sammenligning med de

10 for de syntetiserte talesignaler;

hvilken spektralinformasjon er en hvilken som helst av LSP-informasjon, PARCOR-informasjon og LAR-informasjon.

24.

15 Talesyntetiserende anordning, k a r a k t e r i s e r t v e d a t d e n omfatter:

anordning til frembringelse av et syntetisert talesignal på basis av en første spektralinformasjon som er representert i form av en flerdimensjonal vektor og tilhører et på forhånd bestemt verdiområde og gjelder inngangstalesignalene;

20 anordning for transformering av den første spektralinformasjon til en andre spektralinformasjon som tilhører et verdiområde som er forskjellig fra det på forhånd bestemte verdiområde;

anordning for filtrering av syntetiserte talesignaler med en overføringsfunksjon som er definert med filterkoeffisienter for å frembringe modifiserte syntetiserte talesignaler; og

25 anordning til frembringelse av filterkoeffisientene på basis av den andre spektralinformasjon for å sikre at formantkarakteristikkene for de modifiserte syntetiserte talesignaler blir forbedret i overensstemmelse med den andre spektralinformasjon og i sammenligning med de for de syntetiserte talesignaler;

30 hvilken spektralinformasjon er en hvilken som helst av LSP-informasjon, PARCOR-informasjon og LAR-informasjon.

25.

Talesyntetiserende anordning, k a r a k t e r i s e r t v e d a t d e n omfatter:

35 anordning for frembringelse av syntetiserte talesignaler på basis av en første spektralinformasjon som er representert i form av en flerdimensjonal vektor og tilhører et på forhånd bestemt verdiområde og gjelder innmatede talesignaler;

anordning for analyse av de syntetiserte talesignaler for å frembringe en andre spektralinformasjon;

anordning for filtrering av syntetiserte talesignaler med en overføringsfunksjon som er definert med filterkoeffisienter for å frembringe modifiserte syntetiserte talesignaler; og

- 5 anordning til frembringelse av filterkoeffisientene på basis av den andre spektralinformasjon for å sikre at formantkarakteristikkene for de modifiserte syntetiserte talesignaler blir forbedret i henhold til den andre spektralinformasjon og i sammenligning med de for de syntetiserte talesignaler;
- 10 hvilken spektralinformasjon er en hvilken som helst av LSP-informasjon, PARCOR-informasjon og LAR-informasjon.

26.

Talelagrende/overførende system, k a r a k t e r i s e r t v e d a t det omfatter:

- 15 anordning for analyse av innmatede talesignaler til frembringelse av spektralinformasjon som er representert i form av en flerdimensjonal vektor og tilhører et på forhånd bestemt verdiområde og gjelder de innmatede talesignalene;
- anordning for lagring eller overføring av spektralinformasjonen;
- anordning til frembringelse av syntetiserte talesignaler på basis av
- 20 spektralinformasjonen som er blitt lagret eller overført;
- anordning for filtrering av de syntetiserte talesignaler med en overføringsfunksjon som er definert av filterkoeffisienter for å frembringe modifiserte syntetiserte talesignaler; og
- anordning for frembringelse av filterkoeffisientene på basis av spektralinformasjonen for å sikre at formantkarakteristikkene for de modifiserte syntetiserte talesignaler blir
- 25 forbedret i henhold til spektralinformasjonen og i sammenligning med de for de syntetiserte talesignaler;
- hvilken spektralinformasjon er en hvilken som helst av LSP-informasjon, PARCOR-informasjon og LAR-informasjon.

30 27.

Talelagrende, overførende system, k a r a k t e r i s e r t v e d a t det omfatter:

- anordning for analysering av innmatede talesignaler for å frembringe første spektralinformasjon som er representert i form av en flerdimensjonal vektor og tilhører
- 35 et på forhånd bestemt verdiområde og gjelder de innmatede talesignalene;
- anordning for lagring eller overføring av den første spektralinformasjon;

anordning for frembringelse av et syntetisert talesignal på basis av den første spektralinformasjon som er blitt lagret eller overført;

anordning for transformering av den første spektralinformasjon til en andre spektralinformasjon som tilhører et verdiområde som er forskjellig fra det på forhånd

5 bestemte verdiområde;

anordning for filtrering av de syntetiserte talesignaler med en overføringsfunksjon som er definert av filterkoeffisienter for å frembringe modifiserte syntetiserte talesignaler; og

anordning til frembringelse av filterkoeffisientene på basis av den andre

spektralinformasjon for å sikre at formantkarakteristikkene for de modifiserte

10 syntetiserte talesignaler blir forbedret i henhold til den andre spektralinformasjon og i sammenligning med de for de syntetiserte talesignaler;

hvilken spektralinformasjon er en hvilken som helst av LSP-informasjon, PARCOR-informasjon og LAR-informasjon.

15 28.

Talelagrende/overførende system, k a r a k t e r i s e r t v e d a t det omfatter:

anordning til analysering av innmatede talesignaler for å frembringe første

spektralinformasjon som er representert i form av en flerdimensjonal vektor og tilhører

20 et på forhånd bestemt verdiområde og gjelder de innmatede talesignalene;

anordning for lagring eller overføring av den første spektralinformasjon;

anordning til frembringelse av syntetiserte talesignaler på basis av den første spektralinformasjon som er blitt lagret eller overført;

25 anordning for analyse av de syntetiserte talesignaler for å frembringe en andre spektralinformasjon;

anordning for filtrering av de syntetiserte talesignaler med en overføringsfunksjon som er definert av filterkoeffisienter for å frembringe modifiserte talesignaler;

og

anordning til frembringelse av filterkoeffisientene på basis av den andre

30 spektralinformasjon for å sikre at formantkarakteristikkene for de modifiserte syntetiserte talesignaler blir forbedret i overensstemmelse med den andre

spektralinformasjon og i sammenligning med de for de syntetiserte talesignaler;

hvilken spektralinformasjon er en hvilken som helst av LSP-informasjon, PARCOR-informasjon og LAR-informasjon.

35

29.

Fremgangsmåte til talemodifikasjon, k a r a k t e r i s e r t v e d  
at den omfatter:

- et første trinn med filtrering av syntetiserte talesignaler med en overføringsfunksjon  
5 definert av filterkoeffisienter for å frembringe modifiserte syntetiserte talesignaler; og  
et andre trinn med frembringelse av filterkoeffisientene på basis av  
spektralinformasjonen som er representert av en flerdimensjonal vektor og tilhører et på  
forhånd bestemt verdiområde og gjelder de syntetiserte talesignaler for å sikre at  
formantkarakteristikker for de modifiserte syntetiserte talesignaler blir forbedret i  
10 overensstemmelse med spektralinformasjonen og i sammenligning med de for de  
syntetiserte talesignaler; hvilket andre trinn ligger foran utførelsen av det første trinn;  
hvilken spektralinformasjon er en hvilken som helst av LSP-informasjon, PARCOR-  
informasjon og LAR-informasjon.

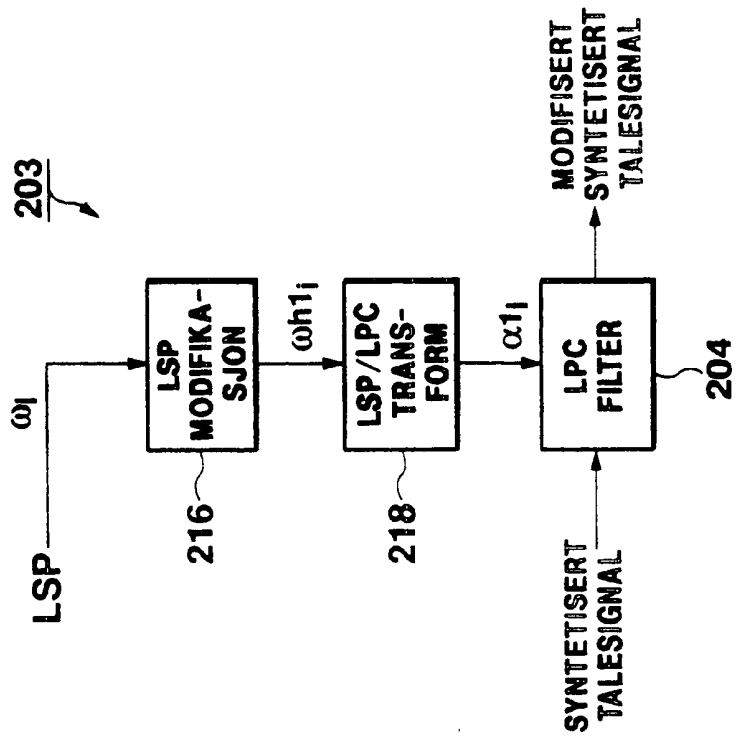


Fig. 2

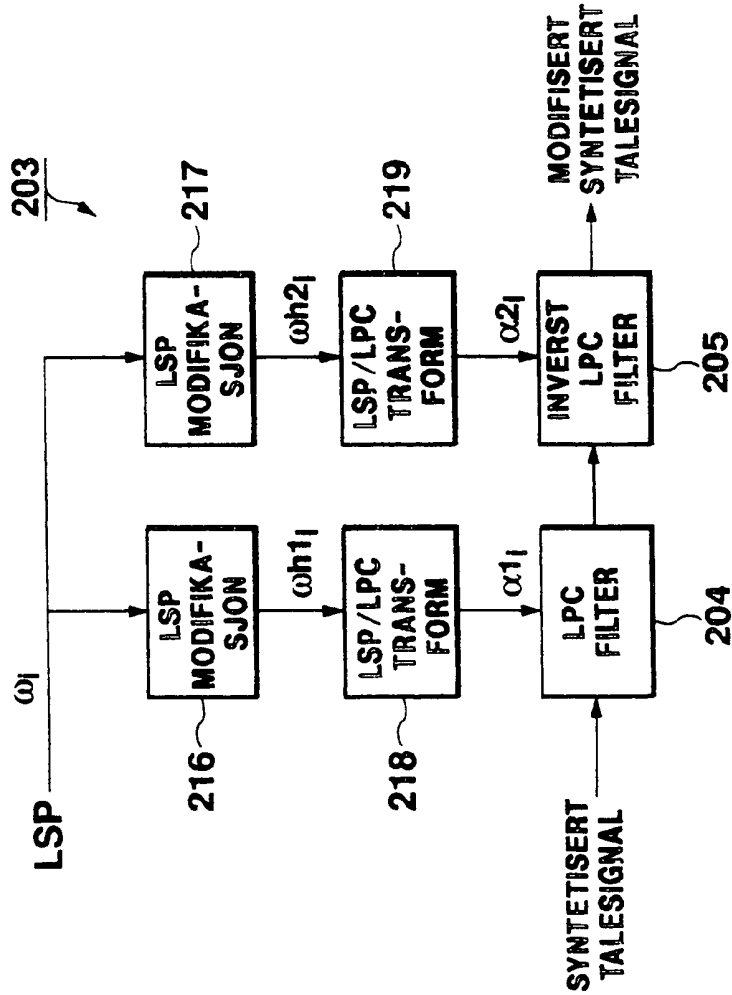


Fig. 1

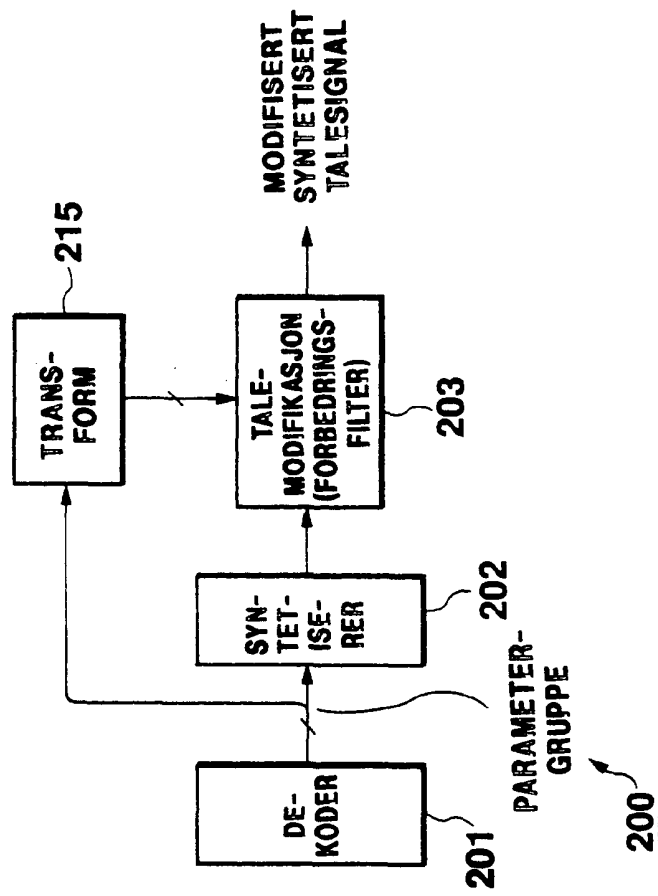


Fig. 3

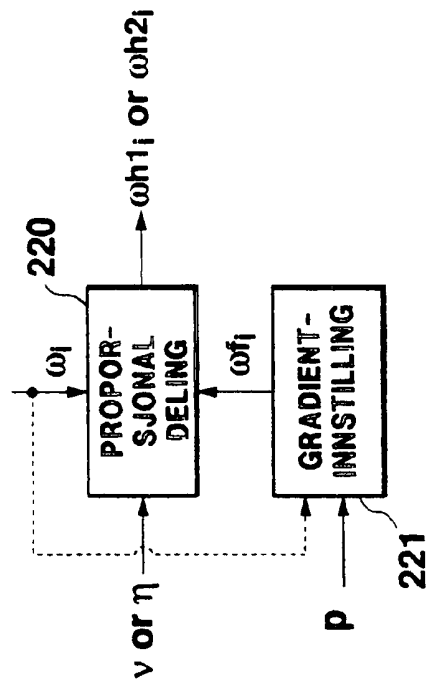


Fig. 4

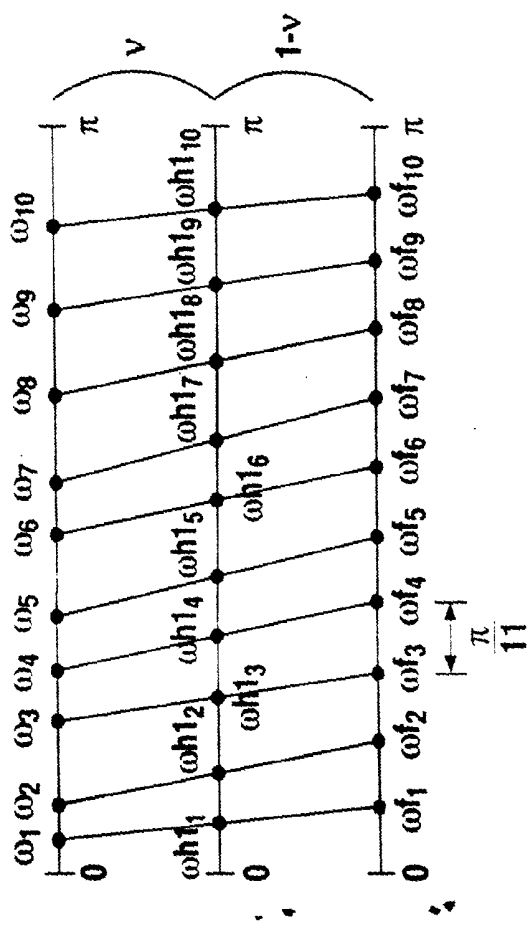


Fig. 5

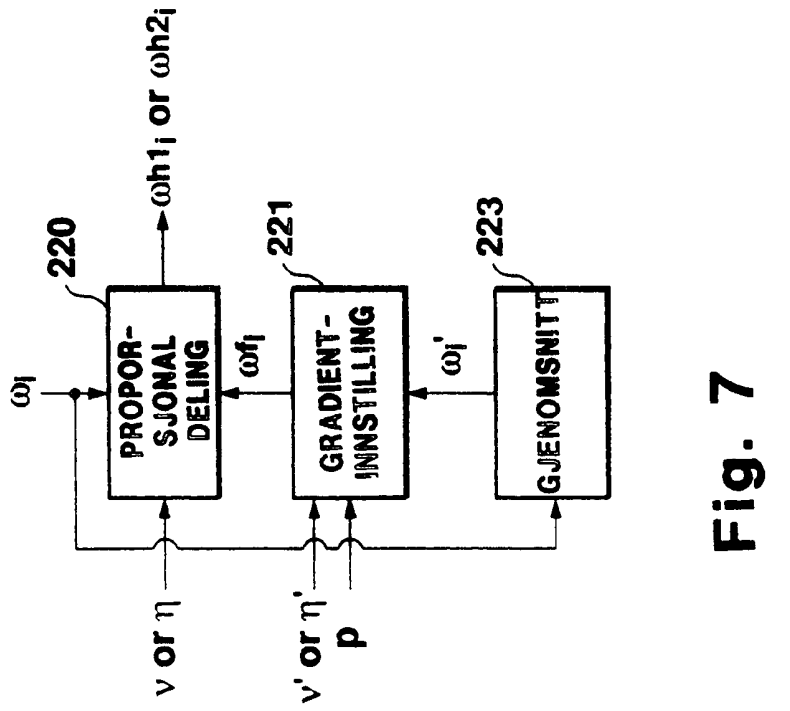


Fig. 6

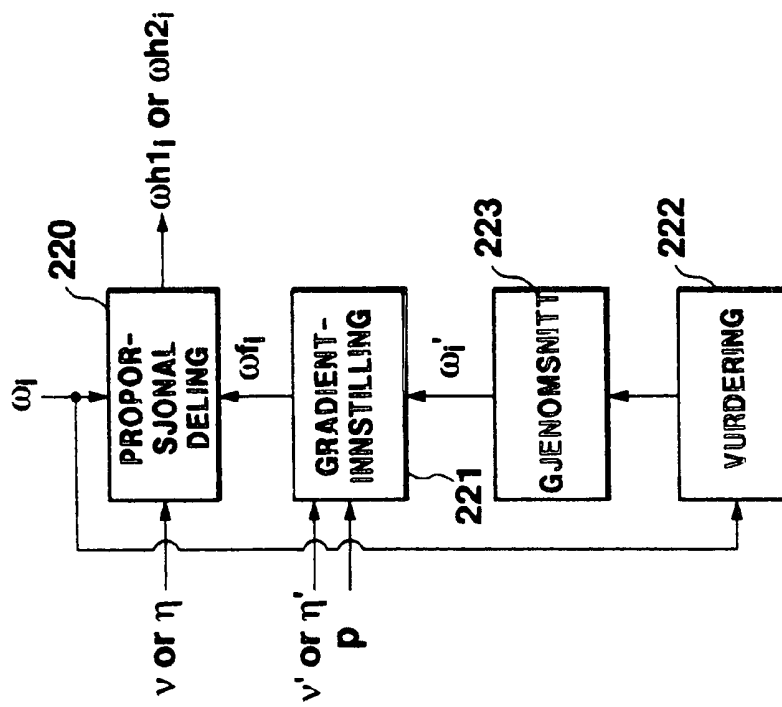
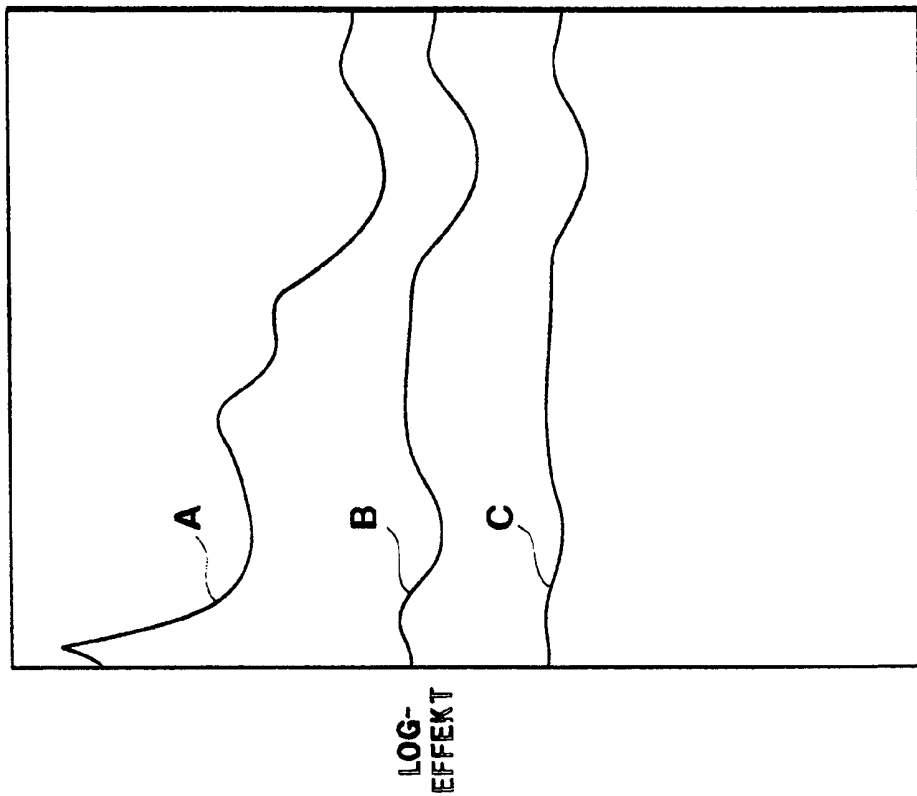
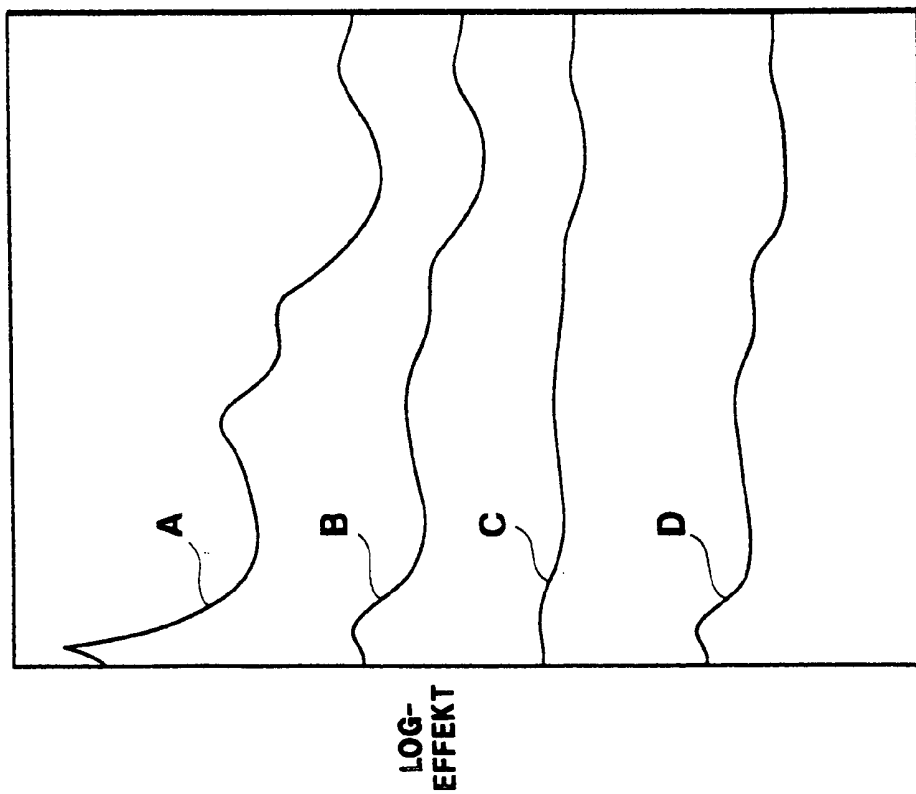


Fig. 7



FREKVENS

Fig. 10



FREKVENS

Fig. 8

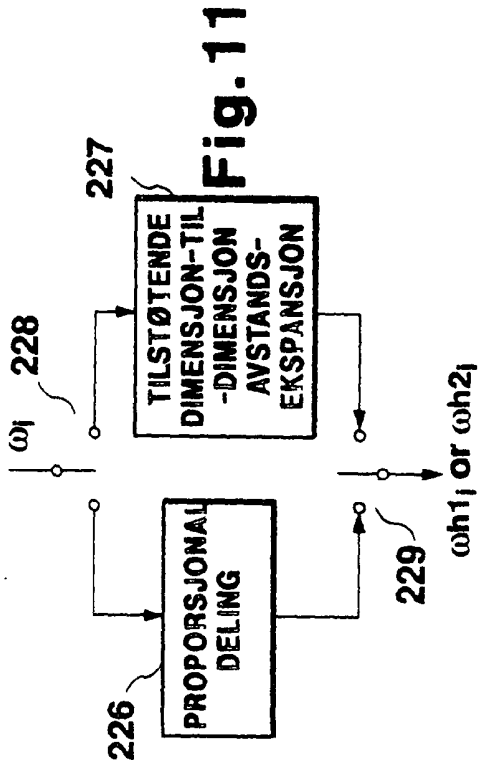


Fig. 11

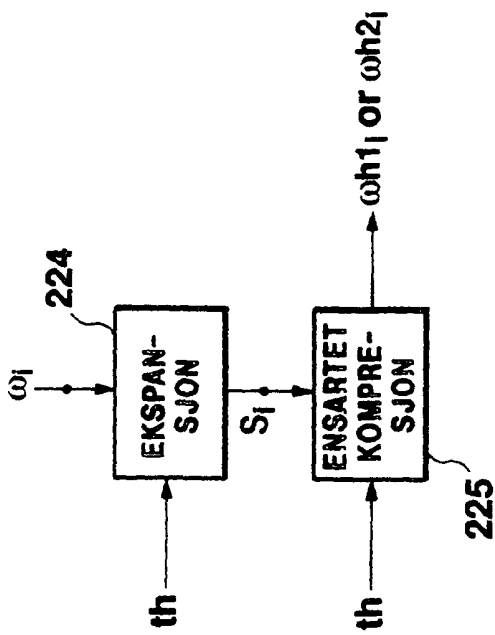


Fig. 9

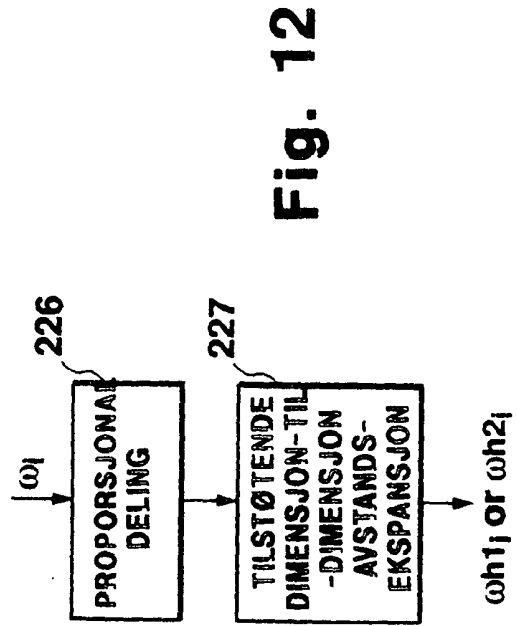


Fig. 12

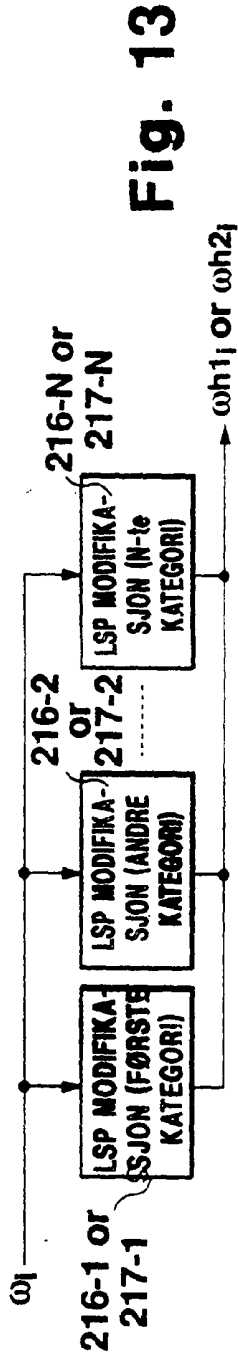


Fig. 13

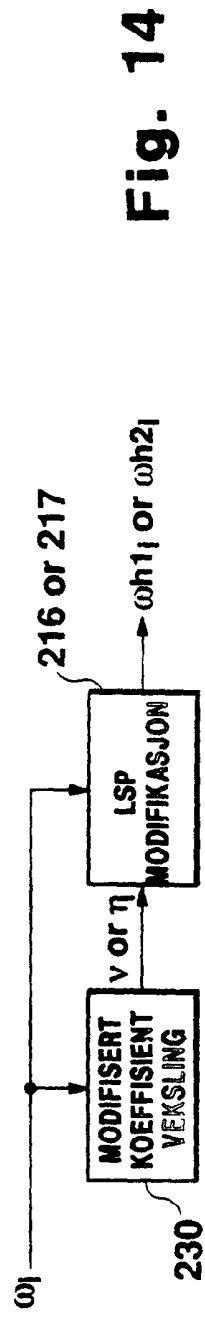


Fig. 14

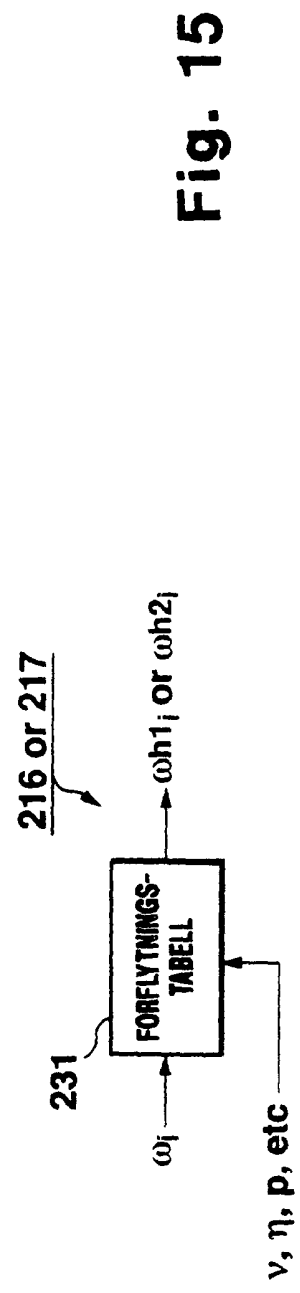


Fig. 15

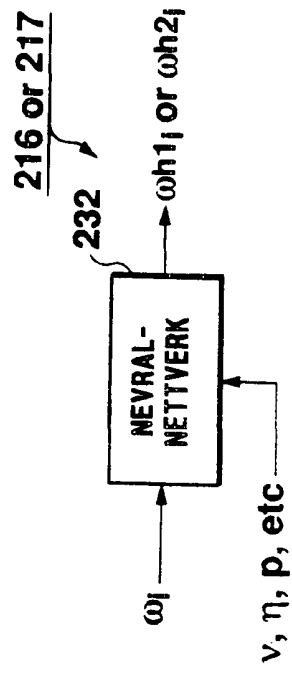


Fig. 16

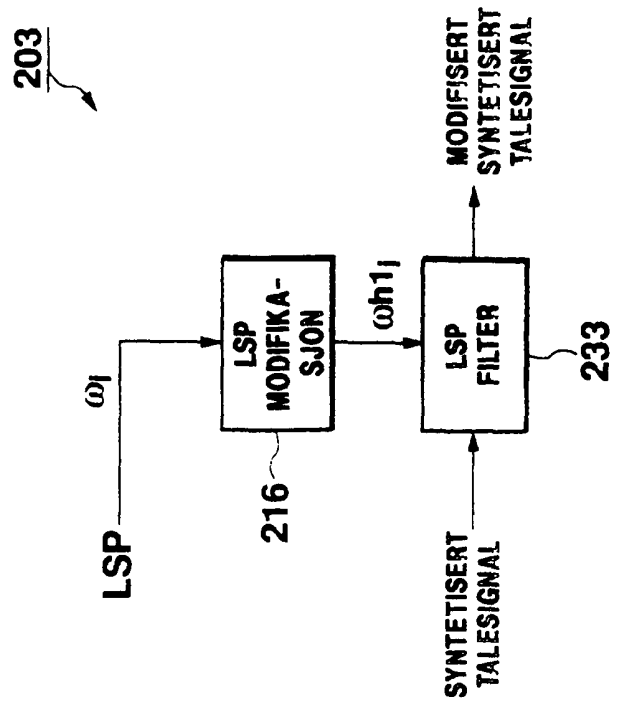


Fig. 18

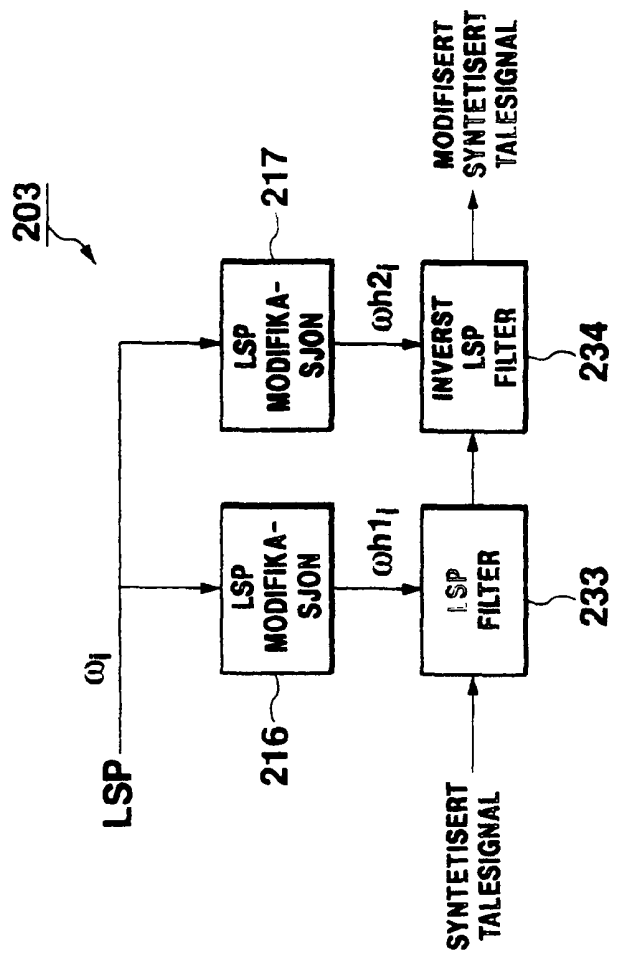


Fig. 17

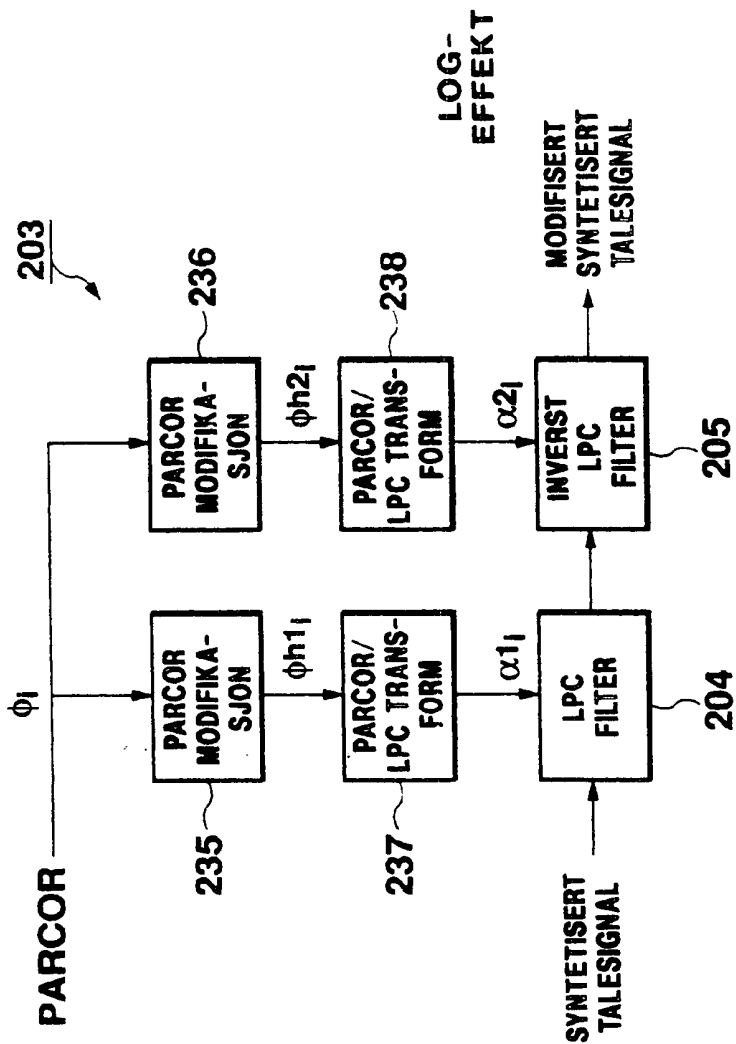
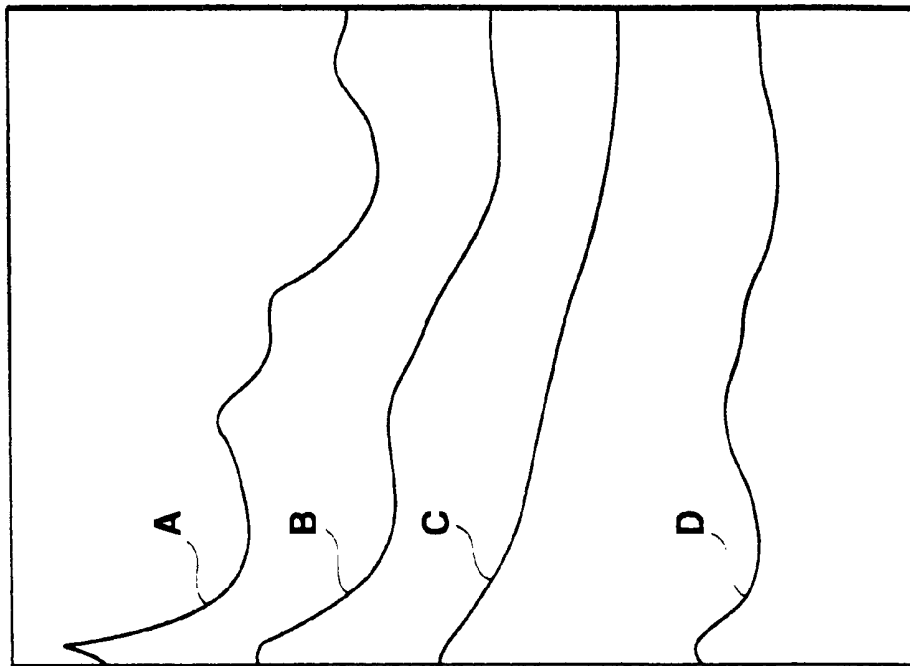


Fig. 19



FREKVENS

Fig. 20

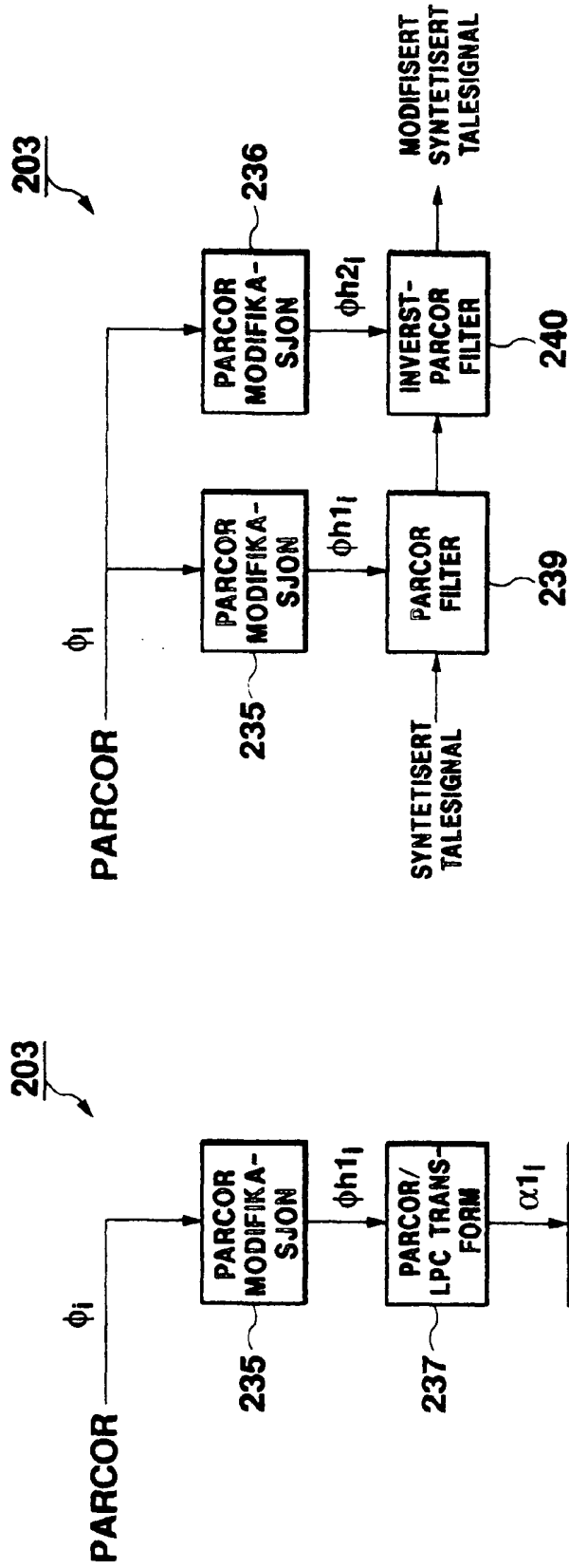


Fig. 22

Fig. 21

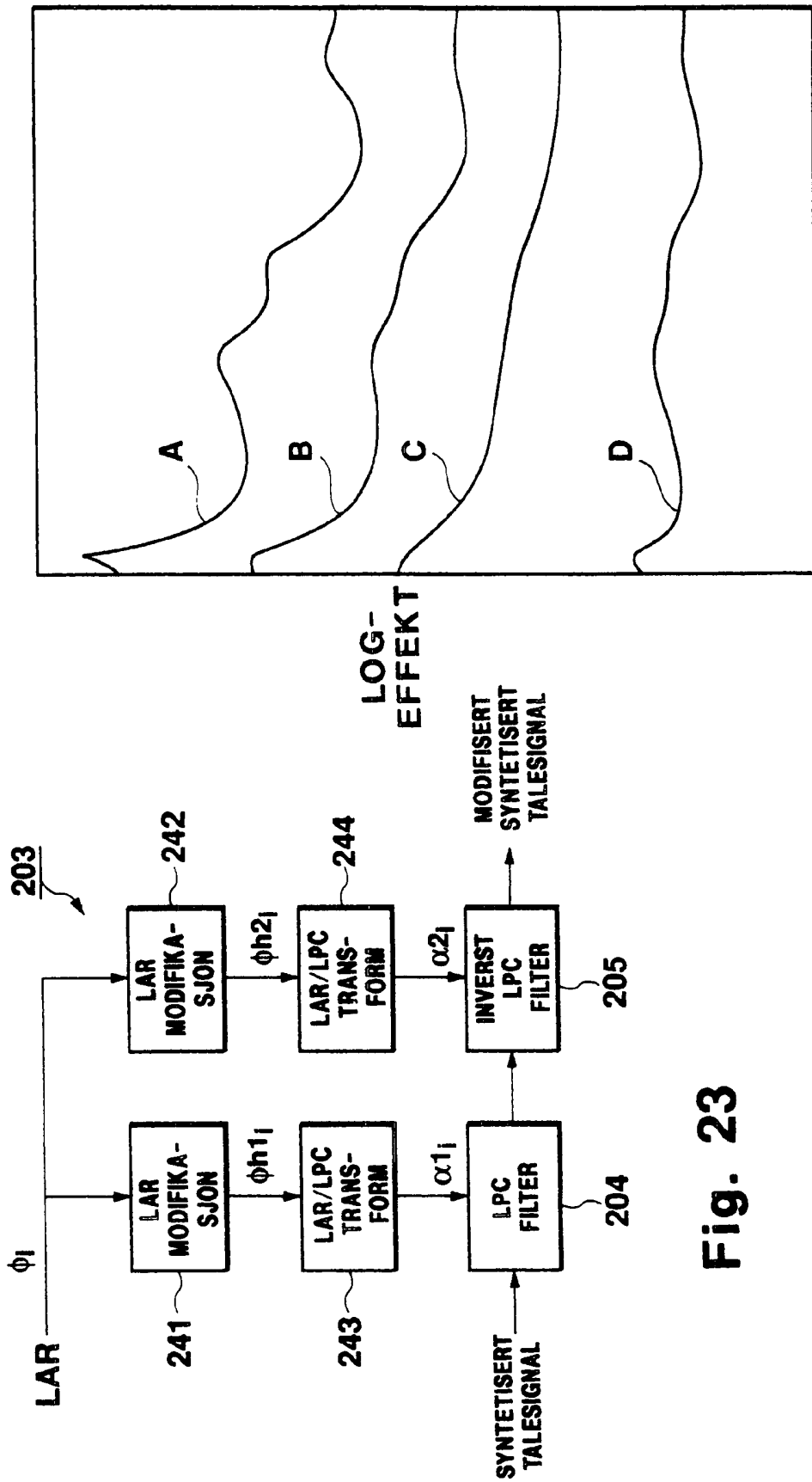


Fig. 23

Fig. 24

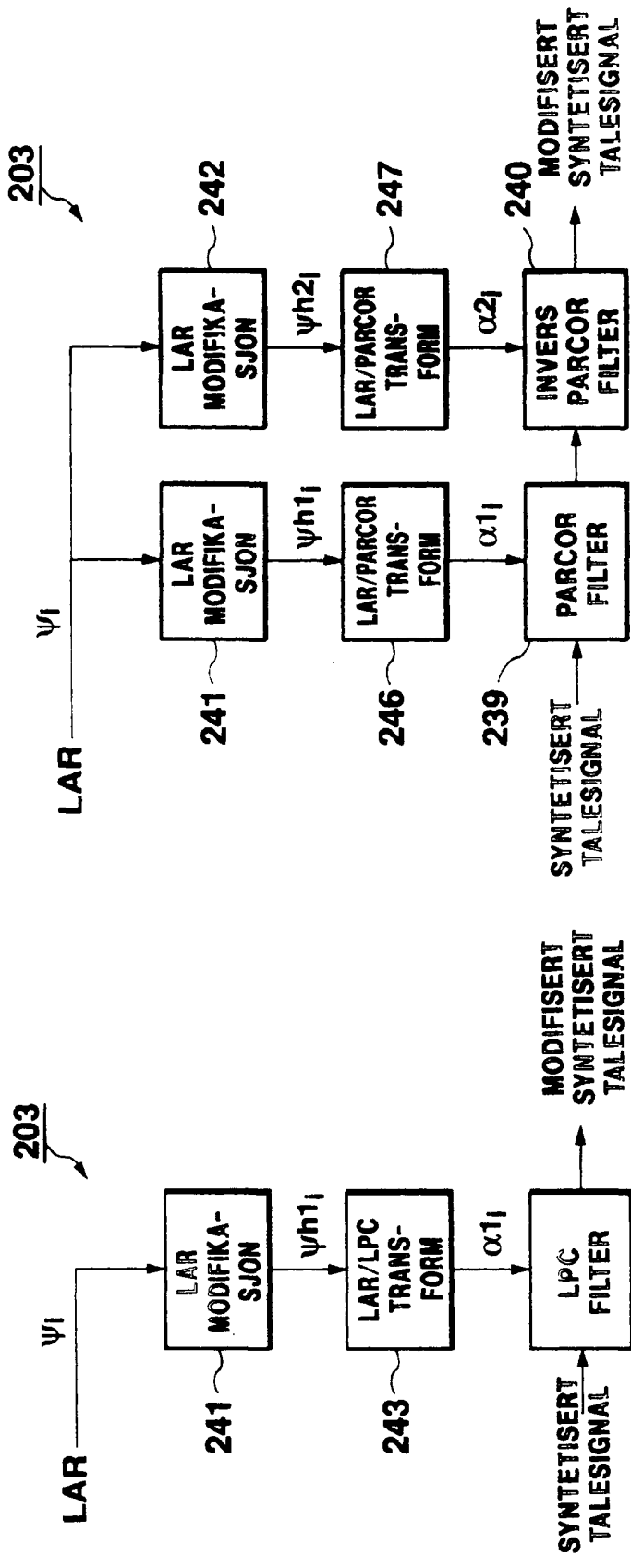


Fig. 26

Fig. 25

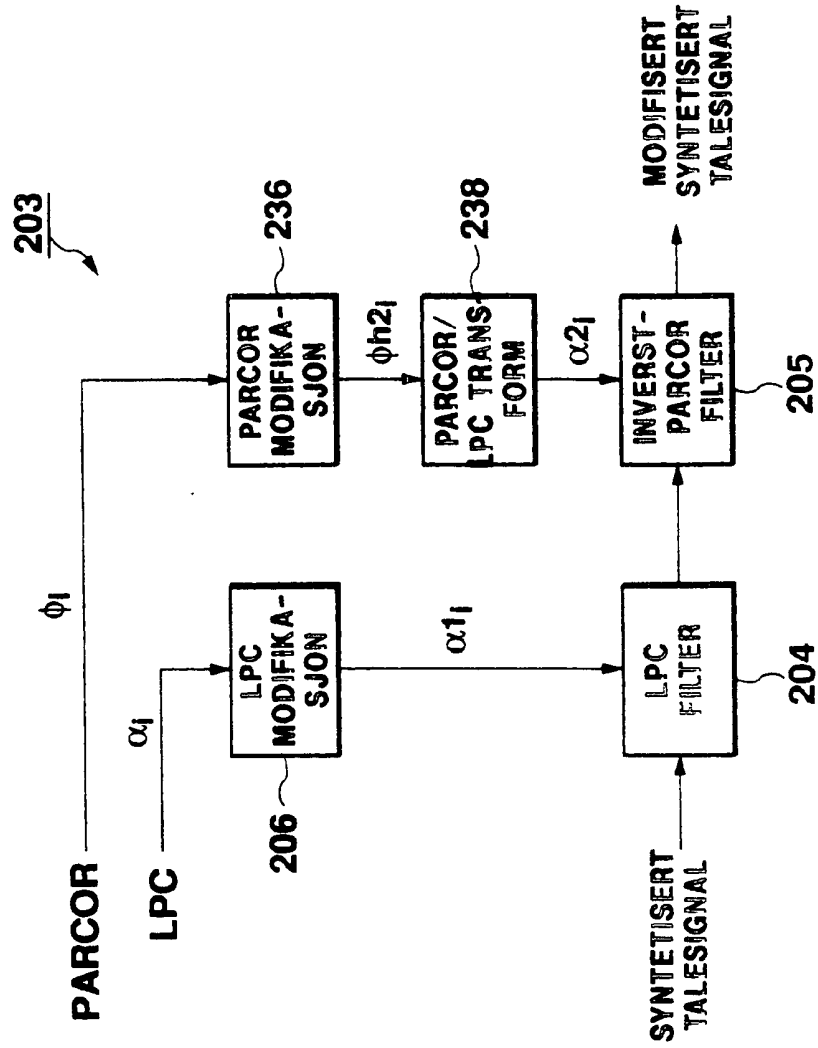


Fig. 27

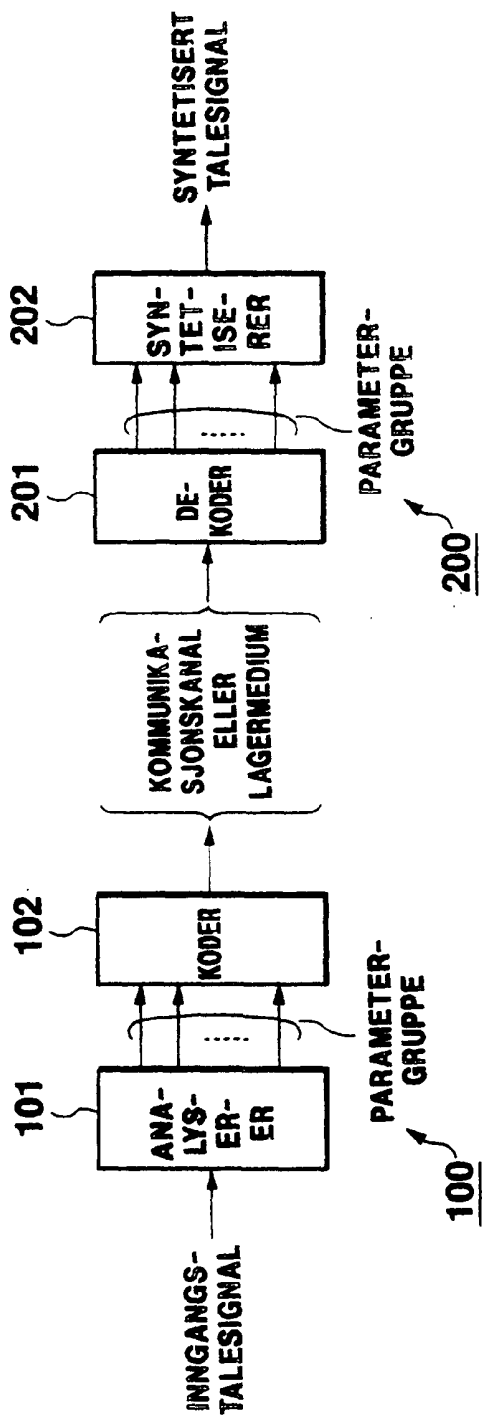


Fig. 28 TIDL. KJENT

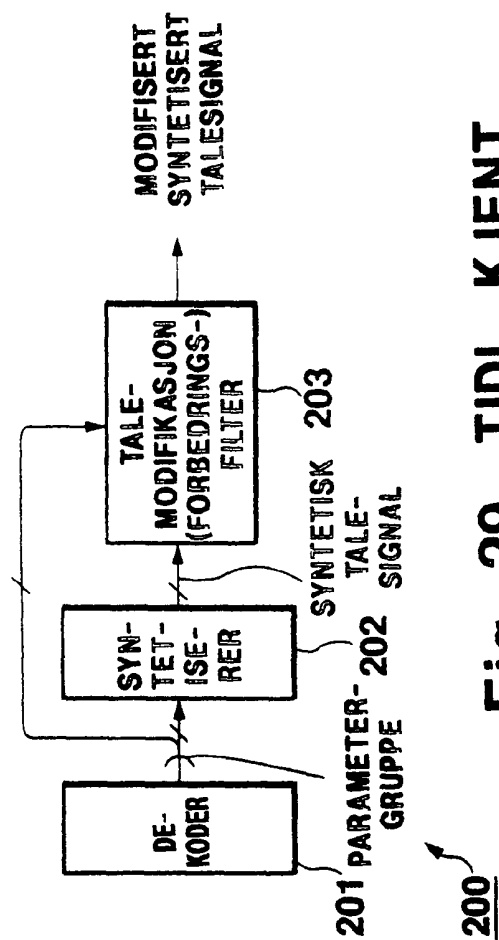


Fig. 29 TIDL. KJENT

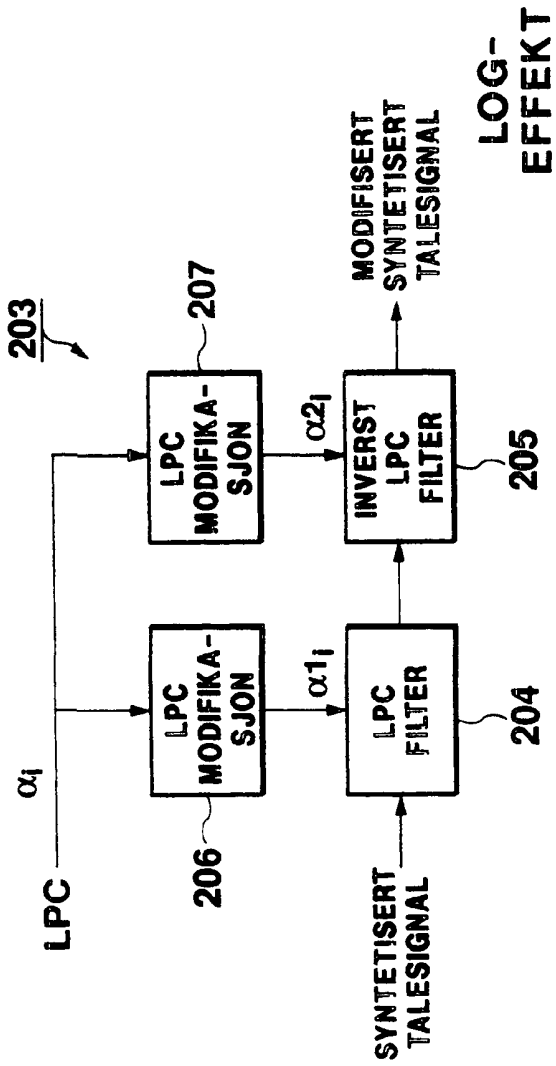
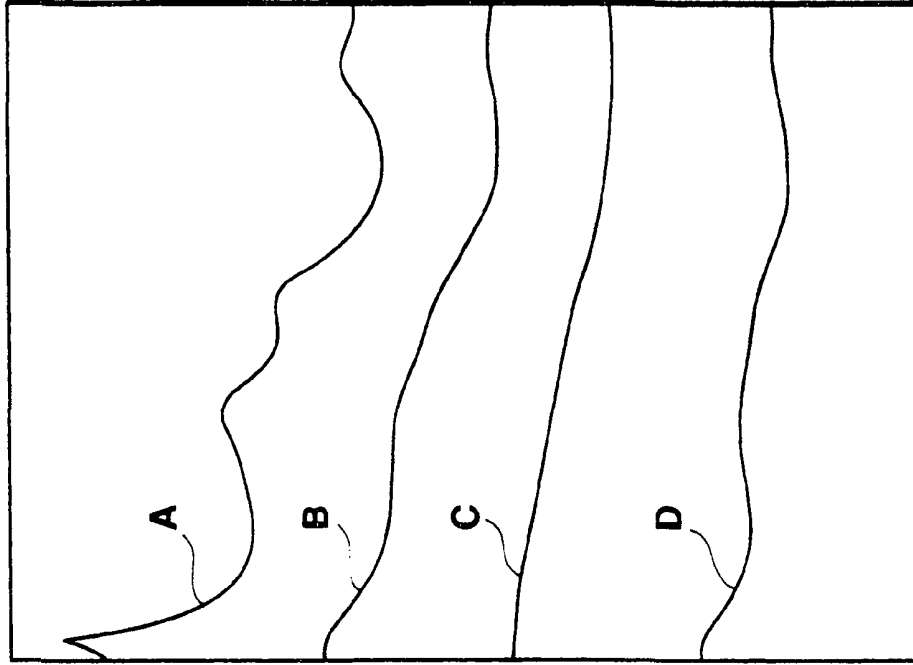
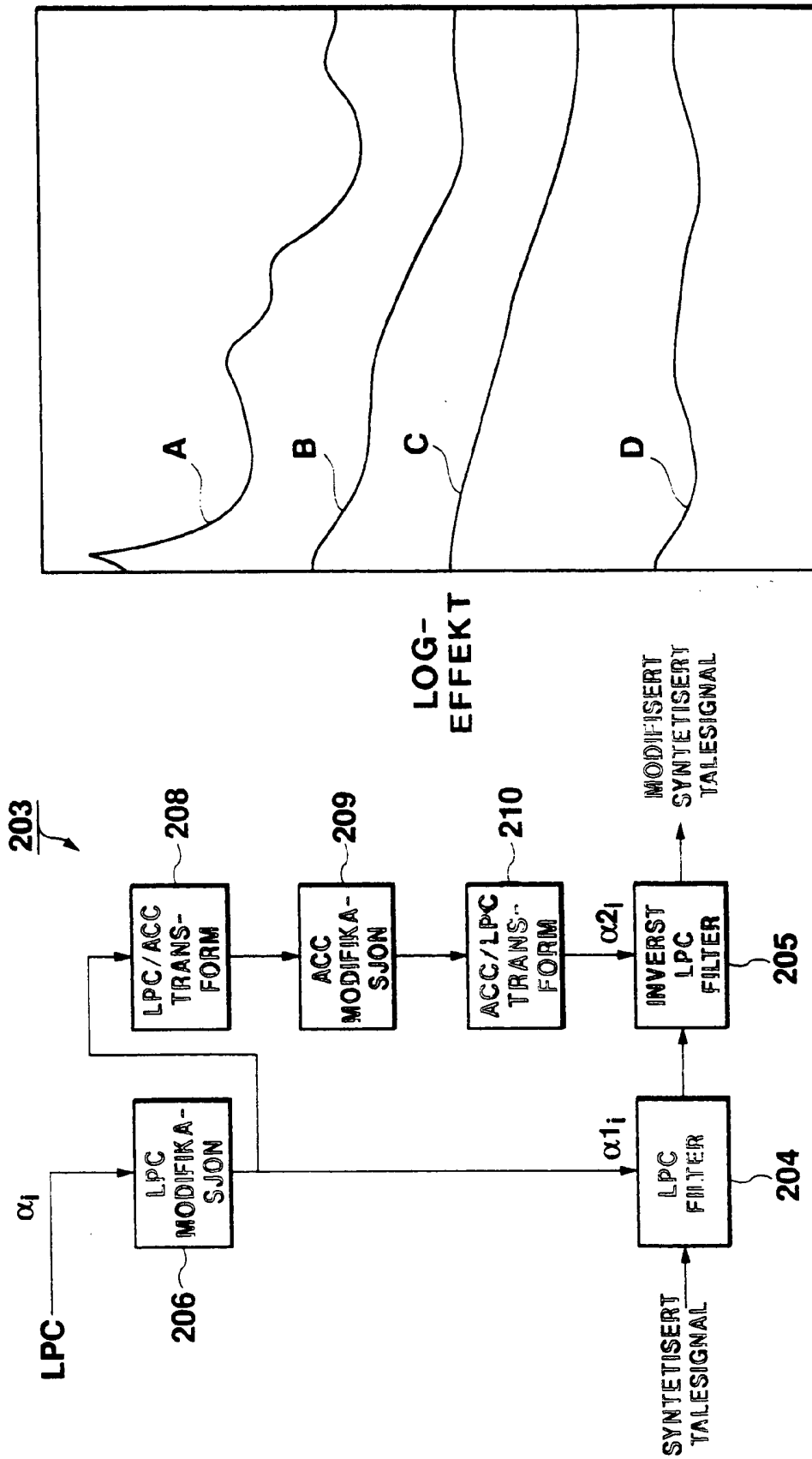


Fig. 30 TIDL. KJENT



FREKVENNS

Fig. 33 TIDL. KJENT



FREKVENNS

Fig. 31 TIDL. KJENT

Fig. 34 TIDL. KJENT

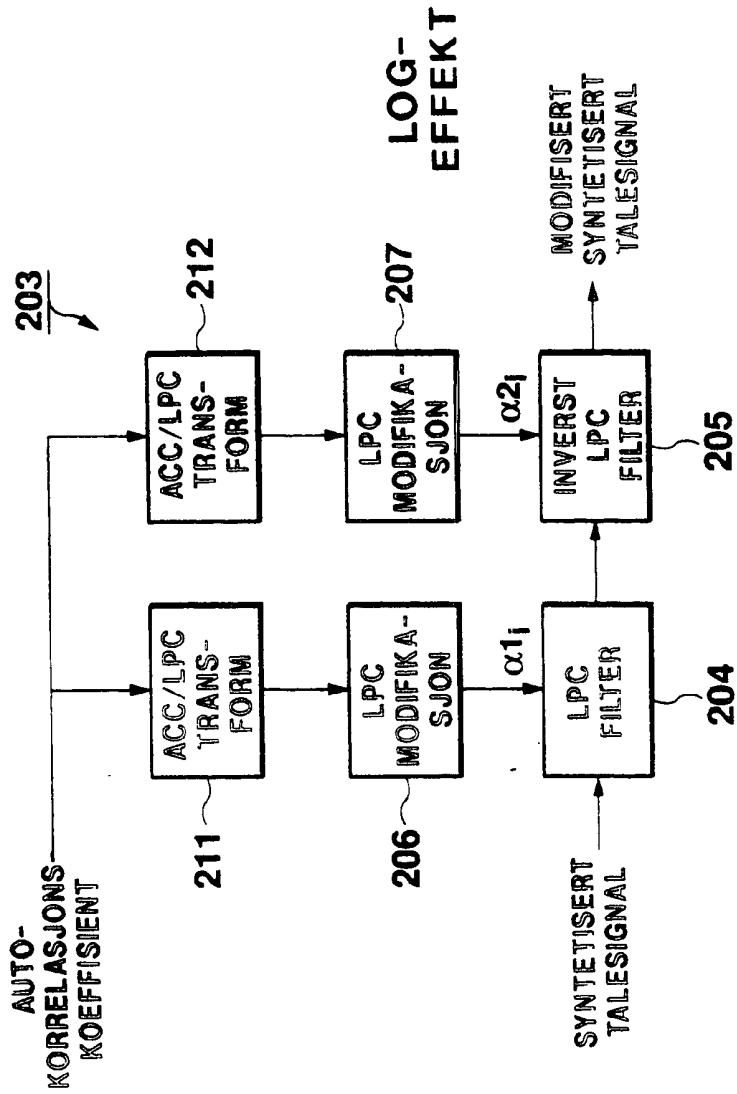
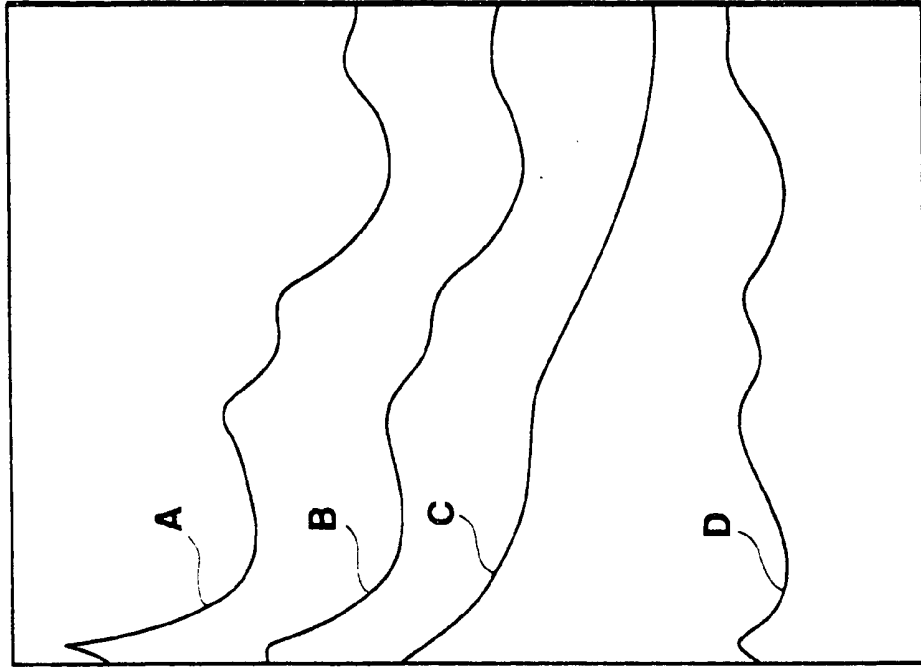


Fig. 32 TIDL. KJENT



FREKVENNS

Fig. 35 TIDL. KJENT

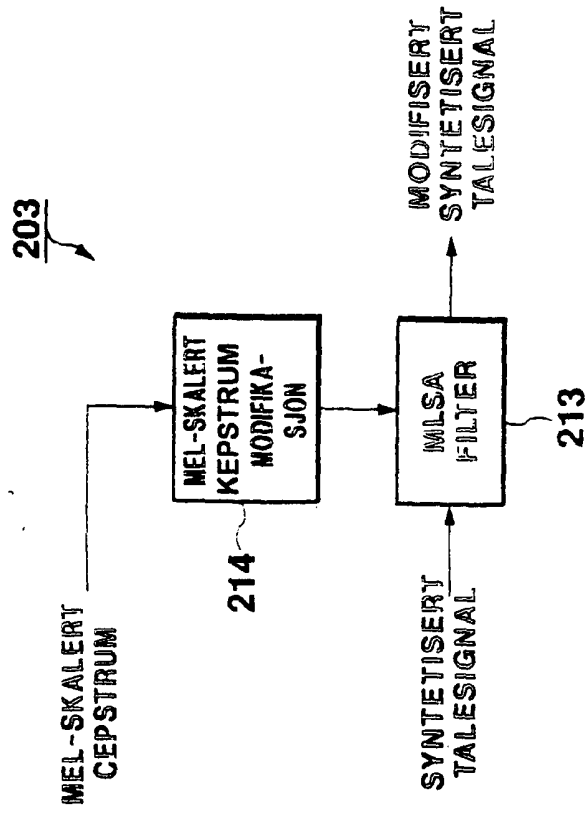


Fig. 36 TIDL. KJENT