

# KORT INDFØRING I AKUSTISKE GRUNDBEGREBER

NIELS REINHOLT PETERSEN

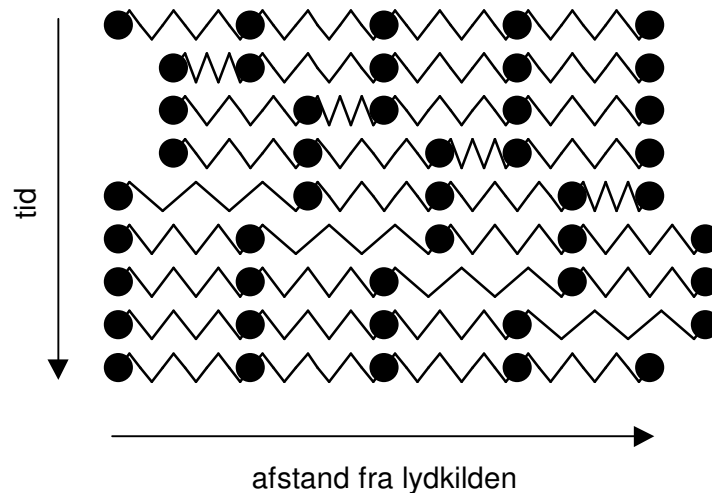
Akustikken er læren om lydes frembringelse og udbredelse. Dvs. den fysiske beskrivelse af lyd. Det fysiske fænomen, der af øret opfattes som lyd er forstyrrelser i luftens tryk i forhold til normaltrykket (= den til enhver tid gældende barometerstand).

## SINUSSVINGNINGEN

For at lyd skal frembringes kræves, at et eller andet bringer forstyrrelse i luftens trykforhold. Sådanne forstyrrelser kan naturligvis afstedkommes af et utal af ting: stemmelæbernes vibrationer, en klaverstreng der anslås, en bog der falder på gulvet osv. osv. Den lyd giver, vi først skal beskæftige os med som et eksempel her, er stemmegaflen. Dens bevægelsesmønster er nemlig ganske simpelt, og den lyd, den frembringer har en ukompliceret opbygning, som kan danne grundlag for beskrivelsen af mere komplicerede lydfænomener. Når stemmegaflen er anslået, vil dens ben bevæge sig frem og tilbage omkring en hvilestilling. Bevægelsens hastighed vil være afhængig af benenes masse og længde og af det materiale, stemmegaflen er lavet af. Udsvingets størrelse vil være afhængigt af, hvor kraftigt stemmegaflen blev anslået. Stemmegafleens ben 'sparker' til de nærmeste luftmolekyler. Disse presses sammen – deres indbyrdes afstand bliver mindre – og lufttrykket stiger. Nu vil de sammentrykkede luftmolekyler ved stemmegafleens ben forsøge at genoprette den normale indbyrdes afstand (nå tilbage til den normale barometerstand). De skubber derfor til de næste molekyler, der igen skubber til de næste osv. Det vil sige, at vi får et område med højt tryk, der bevæger sig væk fra stemmegaflen. Nu svinger stemmegafleens ben tilbage gennem hvilestillingen til den modsatte yderstilling. Der opstår nu en fortynding, et undertryk, i den luft, der ligger lige op ad stemmegafleens ben. Dette undertryk vil – ligesom overtrykket – bevæge sig bort fra stemmegaflen. Vi får altså en række bølger af overtryk og undertryk, der bevæger sig ud i rummet, bort fra stemmegaflen. Dette er illustreret med en 'mekanisk' model i figur 1. Læg mærke til, at der er tale om to forskellige former for bevægelse:

- der sker en transport af energi over en relativt stor afstand
- de enkelte luftmolekyler foretager ganske små bevægelser frem og tilbage omkring en hvilestilling.

Når ændringerne i lufttrykket når ørets trommehinde, sættes denne i bevægelse svarende til variationen mellem over- og undertryk. Bevægelsen overføres via øreknoglerne til det indre øre, hvor den omsættes til nerveimpulser, som så af hjernen registreres som lyd.



**figur 1.** Udbredelsen af overtryk og undertryk i luften. Kuglerne symboliserer luftmolekylerne og fjedrene imellem dem symboliserer de kræfter, der søger at holde en konstant afstand mellem molekylerne, dvs. opretholde det normale atmosfæretryk.

Den hastighed tryk- og fortyndingsbølgerne udbredes med – lydens hastighed – er ved 20° Celsius 342 m pr. sekund.

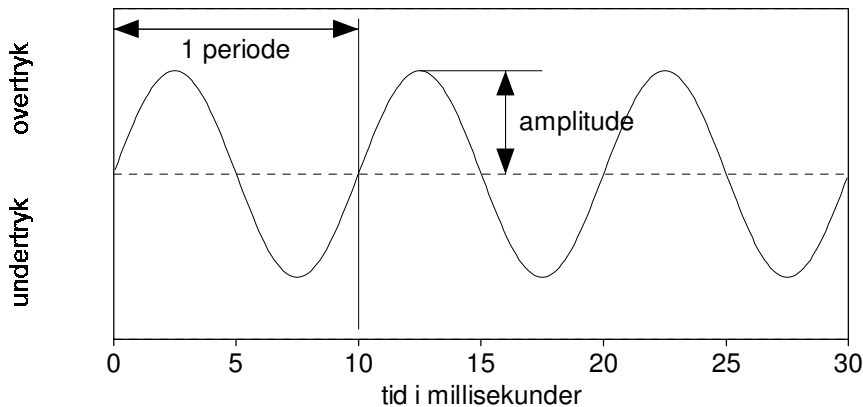
Nu vil vi naturligvis være interesseret i at måle og beskrive lydbølgens egenskaber. Vi vil gerne vide, hvor store trykændringerne er, og hvor hyppigt de skifter mellem overtryk og undertryk. Et almindeligt barometer vil ikke kunne bruges, det er for langsomt. Men man har et hurtigtvirkende barometer i form af mikrofonen, som omsætter trykvariationerne til en elektrisk vekselspænding, som kan visualiseres på forskellig måde, for eksempel i form af et *oscillogram* (oversat: svingningsudskrift), der viser trykvariationen som funktion af tiden. I figur 2 er der vist et oscillogram af den lyd som stemmegaflen afgiver. Vi kan på kurven se, at trykket først stiger fra 0 (atmosfæretrykket) til et trykmaksimum (største overtryk), derefter falder det til et trykminimum (største undertryk) for så at vende tilbage til 0. Herefter gentages dette svingningsmønster så længe stemmegaflen er i gang. Den lyd, der er afbildet i figur 2, er en *sinustone*. Den hedder sådan, fordi den kan beskrives ved den matematiske sinusfunktion. Sinussvingningen er den simpleste lyd, man overhovedet kan forestille sig, og den er – som vi skal se – den basale byggeklods, som alle andre lyde, selv de mest komplicerede, er bygget op af.

Sinussvingningen kan entydigt beskrives ud fra to uafhængige parametre, nemlig *amplituden*, som angiver udsvingets størrelse, altså trykkets afvigelse fra nul, og *periodetiden*, som angiver, hvor lang tid en enkelt svingning (nul-maksimum-minimum-nul) varer. Af periodetiden afledes *frekvensen*, som er et mål for hvor mange svingninger, der er pr. sekund. Frekvensen angives i Hertz (forkortet Hz). Hvis periodetiden, som i figur 2,

er 10 millisekunder (forkortet ms) – altså  $10/1000$  sek. =  $1/100$  sek. – så er frekvensen 100 Hz; 100 svingninger ‘fylder’ 1 sekund. Det vil sige, at frekvensen er omvendt proportional med periodetiden:

$$f = \frac{1}{T} \quad (1)$$

hvor  $f$  står for frekvensen i Hz og  $T$  for periodetiden i sekunder. Amplituden er bestemmende for den opfattede lydstyrke og frekvensen for den opfattede tonehøjde. Læg mærke til, at periodetid og frekvens er to udtryk for det samme, hvorimod frekvens/periodetid og amplitude er indbyrdes uafhængige størrelser.



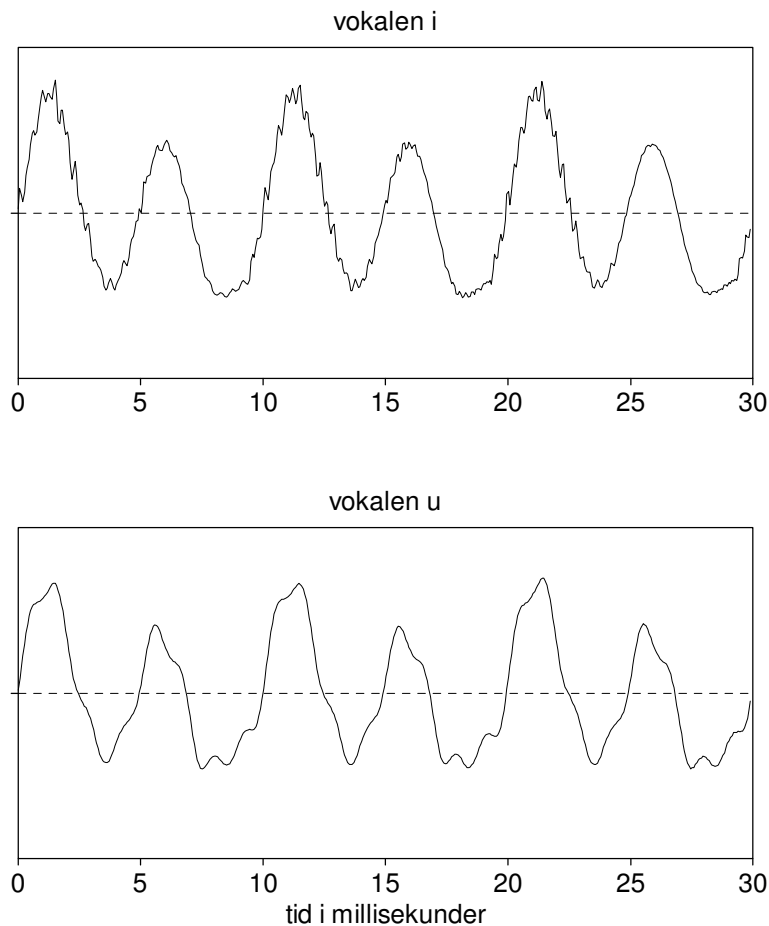
**Figur 2.** Tre perioder af en sinustone med frekvensen 100 Hz (svarende til periodetiden 10 millisekunder).

Det er ikke alle trykvariationer i luften, som opfattes som lyd. Der stilles visse krav til frekvens og amplitude. Man kan for eksempel ikke høre variationen mellem meteorologiske højtryk og lavtryk, den er for lavfrekvent. Man kan heller ikke høre de svingninger flagermus udsender, de er for højfrekvente. Det har vist sig, at frekvenser i området ca. 18 til 20.000 Hz opfattes som lyd. Med hensyn til amplituden rækker ørets dynamikområde (forskellen mellem den svageste og den kraftigste lyd man kan høre) fra 0.001 bar (1 bar = 760 mm kviksølvsøjle) som det kraftigste helt ned til  $10^{-8}$  ( $1/1000.000.000$ ) bar. Lyd over grænsen opfattes som smerte, altså som et andet sanseindtryk, og lyd under grænsen opfattes slet ikke.

### KOMPLEKSE SVINGNINGER

Vi har nu set, hvordan man kan beskrive en sinustone entydigt ud fra frekvens og amplitude. To sinustoner med samme frekvens og amplitude lyder ens. Imidlertid er disse to parametre ikke tilstrækkelige, hvis vi skal gøre rede for mere almindeligt forekommende lyde. For eksempel kan det godt lade sig gøre at spille nøjagtigt samme tone med nøjagtigt samme styrke på en trompet og på en klarinet. Men vi kan alligevel sagtens høre forskel på de to toner. På samme måde vil to vokaler (feks. *i* og *u*) lyde forskelligt, selvom de siges på samme tone og med samme styrke. For at gøre rede for dette må vi indføre endnu et begreb, nemlig *kvalitet*, og vi skal se på, hvordan kvaliteten af lyde manifesteres akustisk.

I figur 3 er der vist oscillogrammer af vokalerne *i* og *u*. Vi kan se, at disse lyde ligesom sinustonen har et svingningsmønster, der gentages regelmæssigt. De er ligesom sinustoner *periodiske* lyde. Men de adskiller sig fra sinustoner ved at være *komplekse*. I oscillogrammet for *i* kan vi se, at samme svingningsmønster gentages hvert tiende millisekund. Vi har altså en



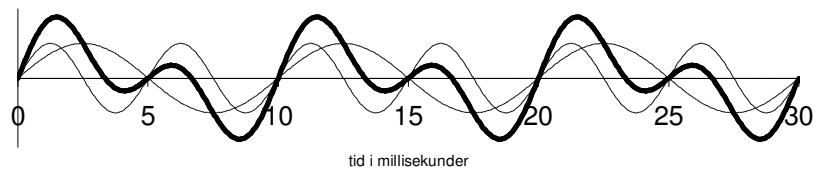
**Figur 3.** Oscillogrammer af vokalen *i* (øverst) og vokalen *u* (nederst). Oscillogrammerne viser tre perioder.

frekvens på 100 Hz. Men vi kan også inden for hver periode på 10 ms tælle to mindre svingninger, der ligner hinanden noget, selvom de ikke er helt ens. Og oven på dette mønster ligger der yderligere nogle meget hurtige, små svingninger. Det kunne altså se ud til, at svingningen består af ikke én men flere samtidige sinussvingninger, nemlig en på 100 Hz, en på 200 Hz og en ved en meget høj frekvens. I oscillogrammet af *u* har vi som ved *i* et mønster, der gentages hver 10 ms, og vi har også to mindre svingninger inden for hver 10 ms periode, men til forskel fra *i* indgår der ikke nogen højfrekvent sinustone i *u*'et. Det, der får netop disse to vokaler til at lyde forskelligt, kunne altså se ud til at være tilstedeværelsen versus fraværelsen af en højfrekvent sinustone. Eller for at sige det mere generelt: forskellige vokaler er sammensat af forskellige sinustoner. (I virkeligheden indgår der betydeligt flere sinustoner end de nævnte i de to vokaler – det vender vi tilbage til senere).

Alle komplekse lyde, og ikke bare vokaler, kan bygges op af sinustoner. Ved at lægge sinustoner sammen og manipulere med deres frekvens og

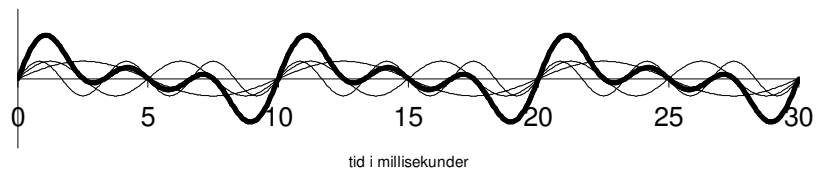
amplitude kan man fremstille en hvilken som helst kompleks svingningsform. Dette er der vist eksempler på i figur 4, 5 og 6, hvor de tyndt optrukne

1)  $f=100$  Hz  $a=1$ ; 2)  $f=200$  Hz  $a=1$



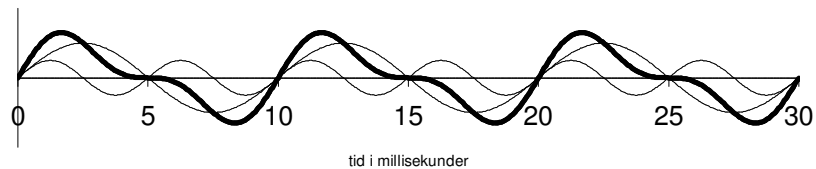
**Figur 4.** To sinustoner (tynde linier) med frekvenserne 100 og 200 Hz med samme amplitude og den komplekse svingning (fed linie), der er summen af de to. Betegnelserne  $f$  og  $a$  står for henholdsvis frekvens og amplitude.

1)  $f=100$  Hz  $a=1$ ; 2)  $f=200$  Hz  $a=1$ ; 3)  $f=300$  Hz  $a=1$



**Figur 5.** Tre sinustoner (tynde linier) med frekvenserne 100, 200 og 300 Hz med samme amplitude og den komplekse svingning (fed linie) der er summen af de tre. Betegnelserne  $f$  og  $a$  står for henholdsvis frekvens og amplitude.

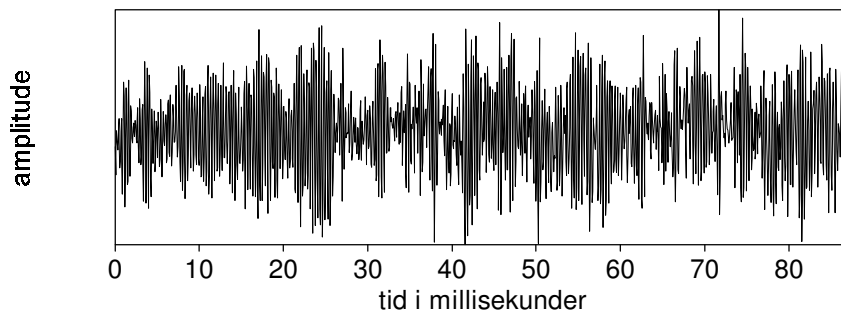
1)  $f=100$  Hz  $a=1$ ; 2)  $f=200$  Hz  $a=0.5$



**Figur 6.** To sinustoner (tynde linier) med frekvenserne 100 og 200 Hz, hvor 200 Hz tonens amplitude er det halve af 100 Hz tonens, og den komplekse svingning (fed linie), der er summen af de to. Betegnelserne  $f$  og  $a$  står for henholdsvis frekvens og amplitude.

linier viser de sinustoner, som den komplekse tone – angivet med tyk linie – er bygget op af. I figur 4 er den komplekse tone summen af en 100 Hz og en 200 Hz sinustone, som har samme amplitude. I figur 5 er der føjet endnu en sinustone til, nemlig en på 300 Hz med samme amplitude som de øvrige, og i figur 6 indgår de samme to sinustoner som i figur 4, 100 og 200 Hz, men her er 200 Hz tonens amplitude kun det halve af 100 Hz tonens.

De lyde, der er vist i figur 3-6 er alle karakteriseret ved at være periodiske og komplekse. Sådanne lyde, som altså er opbygget af flere sinustoner, og hvor samme svingningsmønster gentages regelmæssigt, kaldes *klange*. Nu er en masse af de lyde, vi kommer ud for, ikke periodiske. Et eksempel på en ikke-periodisk – *aperiodisk* – lyd er konsonanten *s*. Et oscillogram af denne lyd er vist i figur 7. Vi kan se, at trykvariationerne er helt tilfældi-



Figur 7. Oscillogram af støj i konsonanten s

ge; der kan ikke erkendes noget regelmæssigt gentaget svingningsmønster. En aperiodisk lyd kaldes også en *støj*. En støj kan ligesom en periodisk lyd opbygges af sinustoner. Hvis vi skal 'bygge' et s, skal vi bruge mange (i princippet uendelig mange) sinustoner i området ca. 2000 til ca. 15000 Hz.

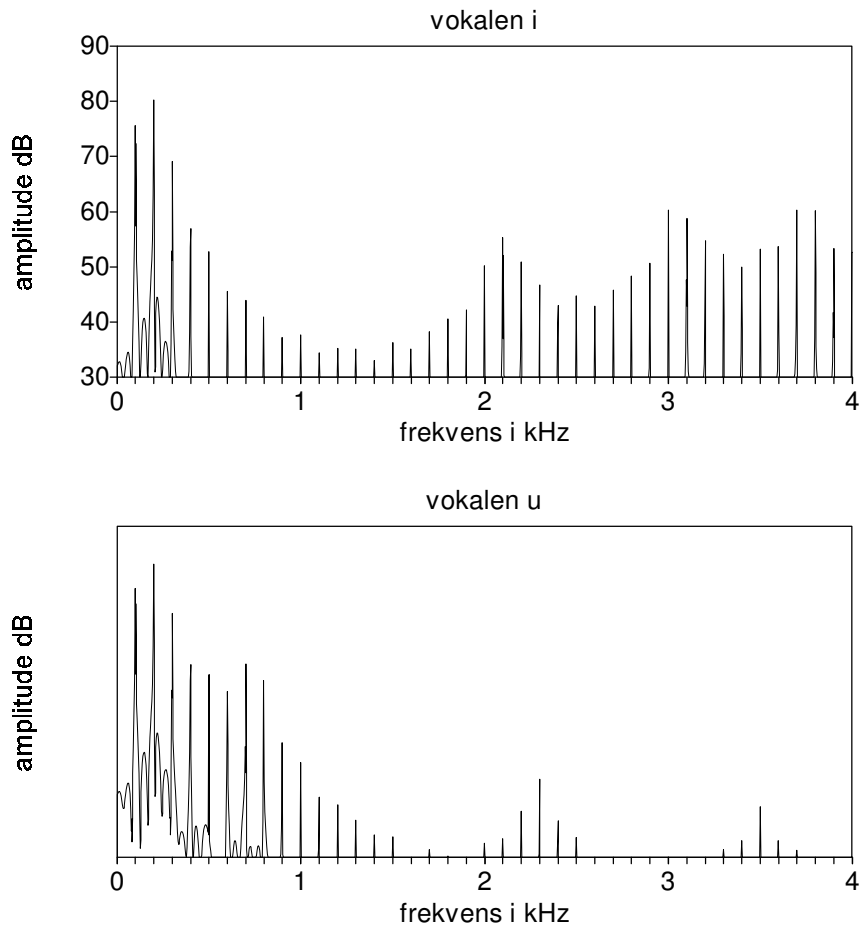
En særlig form for støjlyd er en såkaldt hvid støj (hvorfor den kaldes hvid, kommer vi tilbage til), som er sat sammen af i princippet uendeligt mange lige kraftige sinustoner i frekvensområdet 0 til uendeligt. En sådan 'teoretisk' lyd kan selvfølgelig ikke realiseres, hverken af naturen eller ved teknisk snilde, men noget, der ligner, er den støj, man hører, når man indstiller sin radio uden for stationerne.

### SPEKTRALANALYSE

Det, vi normalt er interesseret i, er ikke at stykke sinustoner sammen til forskellige lyde – det klarer naturen helt fint – men derimod at gå den modsatte vej, at finde ud af hvilke sinustoner, de lyde, vi kommer ud for, er opbygget af, og hvor kraftige disse sinustoner er i forhold til hinanden. Derved vil det være muligt at forklare, hvorfor samme toner spillet på forskellige musikinstrumenter og hvorfor forskellige vokaler lyder forskelligt. Denne analyse, hvor man så at sige skiller en lyd ad for at bestemme frekvens og amplitude af de sinustoner, den er bygget op af, kaldes *spektralanalyse*, og resultatet af analysen kaldes lydens *spektrum*. Når man taler om spektralanalyse, er det i analogi med den spektralanalyse, man kan foretage af lys ved hjælp af et prisme. Her får man også 'splittet lyset ad' i de farver, det er sammensat af. I figur 8 ses spektre af vokalerne *i* og *u*. Hver lodret linie angiver en sinustone. Placeringen på frekvensaksen (den vandrette akse) markerer sinustonens frekvens, og højden af linien angiver dens amplitude.

Fælles for de to vokaler (der er sagt med samme tonehøjde) er, at de sinustoner, de er bygget op af, ligger med samme afstand på frekvensaksen, den laveste ved 100 Hz, den næste ved 200 Hz, den tredje ved 300 Hz osv. Dette er karakteristisk for komplekse, periodiske lyde, hvortil vokalerne jo hører, og man siger at *deltone*ne (sinustonerne) er *harmonisk forbundne*. Den laveste deltone kaldes *grundtonen* eller *første deltone* eller *første harmoniske* eller  $F_0$ , den næste kaldes anden deltone eller anden harmoniske og så videre. Vi kan altså sige, at i komplekse periodiske lyde er deltonerne *hele multipla af grundtonen*. Hvis grundtonen havde været 215 Hz (det kunne være en kvindestemme), havde linierne i spektret ligget ved 215 Hz, 430 Hz, 645 Hz og så videre deropad. (Somme tider taler man om grundtone og overtoner, hvor første overtone svarer til anden harmoniske. Den nummere-

ring - som nok kan bringe forvirring - bruges vistnok især inden for musikken.)

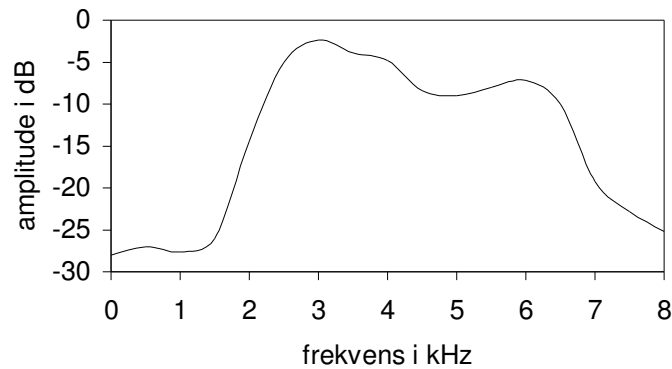


**Figur 8.** Spektre af vokalerne *i* (øverst) og *u* (nederst).

Det, der adskiller de to vokaler og er ansvarlig for deres forskellige kvalitet, er forholdet mellem deltonernes amplituder. Det ses af figur 8, at *u* har næsten al sin energi ved relativt lave frekvenser, mens deltonerne længere oppe i spektret er meget svage. Vokalen *i* har også en del energi ved lave frekvenser, men til forskel fra *u* har *i* også meget kraftige deltoner i området 2000-4000 Hz. Det, vi kan se i spektrene for de to vokaler, stemmer meget godt overens med de gæt, vi udledte pr. øjemål af oscillogrammerne i figur 4. Auditivt stemmer det også med, at *i* har et lysere og *u* et mørkere, mere uldent (undskyld) klangbillede. Vi kan altså sige, når det drejer sig om komplekse, periodiske lyde, at

- grundtonen bestemmer den opfattede tonehøjde
- deltonernes *generelle amplitude* bestemmer den opfattede lydstyrke
- deltonernes *indbyrdes amplitudeforhold* bestemmer den opfattede lydkvalitet.

Vi så tidligere, at aperiodiske lyde også kunne bygges op af sinus-toner, og vi kan derfor naturligvis også analysere dem i sinustoner. I figur 9 ses et spektrum af et *s*. Vi kan se, at der ikke er adskilte linier i spektret. Det har at gøre med, at deltonerne ligger uendeligt tæt. Så i virkeligheden skulle området under kurven have været sort af spektrallinier.



Figur 9. Spektrum af konsonanten *s*.

Når det drejer sig om støj, er det meningsløst at tale om harmonisk forbundne deltoner. Og vi kan jo heller ikke rigtig tale om tonehøjden af et *s*. Men kvaliteten af støjen er – ligesom ved periodiske lyde – bestemt af energifordelingen i spektret. For eksempel vil hovedparten af energien i *sj*-lyden (som i Schubert) ligge lavere i frekvens end i *s*, typisk i området 1500-8000 Hz. Et spektrum af en støj, hvor man giver indhyldningskurven for de uendeligt mange tætliggende sinustoner, kaldes et *kontinuert spektrum*, hvorimod spektret af en periodisk lyd med klart adskilte deltoner kaldes et *liniespektrum*. Ovenfor omtaltes den såkaldt hvide støj. Spektret for en hvid støj vil være en vandret linie, fordi de sinustoner støjen er bygget op af alle har samme amplitude. Når man kalder den type støj for hvid, er det i analogi med sollys, hvor (bogstavelig talt) alle regnbuens farver er repræsenteret med ligelig styrke.

## TALENS AKUSTIK

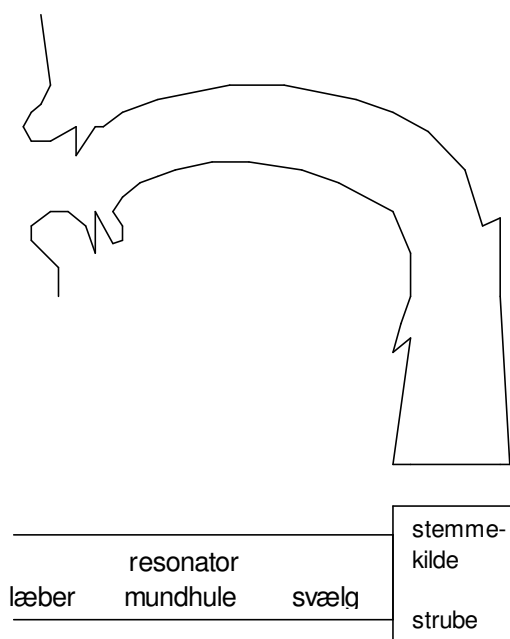
Nu skulle vi være rustet til at komme ind på talens (eller rettere taleproduktionens) akustik. Man kan, når det drejer sig om orale vokaler, meget forenklet betragte taleorganerne som et rør – et luftfyldt hulrum –, der er åbent i den ene ende (ved læberne) og lukket i den anden (ved struben) med en lydkilde, som kan sætte luften i røret i svingninger. Se figur 10.

### Resonans

Inden vi går videre, skal vi se på de akustiske forhold i luftfyldte hulrum. Det er formentlig en almen erfaring, at man kan få en tone, dvs. en periodisk lyd, frem, når man puster hen over halsen på en flaske, og at tonehøjden af denne lyd er afhængig af flaskens størrelse (eller af hvor meget man har drukket af det flydende indhold), altså af det luftfyldte hulrums størrelse. Hvor kommer den tone fra? Det, vi påvirker med, er en luftstrøm, som, idet den rammer flaskehalsens kant, danner *turbulens*, dvs. hvirvler. Denne turbulens giver anledning til en støj; nu har vi lyd, men ingen tone. Men vi ved, at støjen er bygget op af sinustoner, og det kunne derfor se ud som om, flasken så at sige har udvalgt en eller nogle få af disse sinustoner og forstærket dem, mens den har undertrykt andre. Og det er præcis det, der er sket. Alle luftfyldte hulrum har den egenskab, at luften i dem lettere sættes i svingninger af nogle frekvenser end af andre. Man kalder dette fænomen *resonans*, og man siger, at hulrummet har en (eller flere) *resonansfrekvenser* eller *egenfrekvenser*. Resonansfrekvensen er afhængig af hulrummets størrelse og form. Groft kan man sige, at store hulrum eller hulrum med en lang og



snæver hals, har lave resonansfrekvenser, og at små hulrum eller hulrum med en vid kort hals har høje resonansfrekvenser.



**Figur 10.** Snit af taleorganerne (øverst) og taleorganerne betragtet som en simpel rørmodel.

Nu springer vi et øjeblik til en analogi, som kan illustrere resonanfænomenet: Alle, der har været ude for børn (eller voksne for den sags skyld), der holder af at gynges, ved, at man kan bibringe gyngen et ganske betragteligt udsving blot ved at give den ganske små skub med passende mellemrum, det vil sige med en passende frekvens. Og den passende frekvens er netop gyngens egenfrekvens, dens resonansfrekvens. Skubber man med en højere frekvens, har gyngen ikke nået sit maksimale udsving, når skubbet kommer, og det vil sige, at skubbet modarbejder gyngens bevægelse. Skubber man med en lavere frekvens, er gyngen allerede på vej væk igen, når skubbet kommer, og man får ikke tilført den energi. I begge tilfælde stopper gyngen. Hvis man vil tvinge gyngen til at bevæge sig med en frekvens, der afviger fra resonansfrekvensen, så må man tage fat i den og løbe frem og tilbage, og det er jo noget mere energikrævende end at give den et lille elegant skub med den rigtige frekvens.

### Kildespektrum, overføringsfunktion

Så vender vi tilbage til flasken: Netop den af de sinustoner, støjen er sammensat af, som svarer til hulrummets resonansfrekvens, sætter luften i hulrummet i svingninger – og disse svingninger er så kraftige, at de kan sætte luften uden for flasken i svingninger – og vi hører en tone. Lidt mere formelt kan vi beskrive, hvad der sker ved hjælp af 1) *kildespektret*, dvs. spektret for den lyd, vi påvirker resonatoren (dvs. flasken) med, 2) hulrummets *overføringsfunktion*, som er den kurve, der beskriver hvilke frekvenser der forstærkes og hvilke der undertrykkes, og endelig 3) spektret for den *resulterende lyd*, dvs. den lyd, der kommer ud af hulrummet.

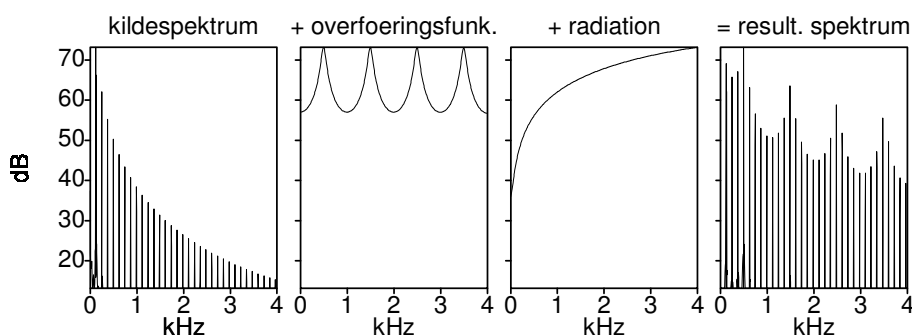
I beskrivelsen af resonatorens overføringsfunktion indgår ud over resonansfrekvensen også *båndbredden*. Båndbredden siger noget om bredden af det frekvensområde, hvor resonatoren forstærker de sinustoner, kildespektret indeholder. Båndbredden bestemmes af hårdheden af det luftfyldte hulrums vægge. En flaske (af glas) har hårde vægge; den giver en kraftig forstærkning i en smalt frekvensområde, den har en lille båndbredde. En beholder med blødere vægge, vil forstærke frekvenserne i et bredere område, men til gengæld ikke forstærke dem så meget som flasken; en sådan resonator har en stor båndbredde.

### Vokaler

Nu kan vi tage rørmodellen af taleorganerne op igen: Her har vi også en lydkilde (i det virkelige liv struben) og et luftfyldt, rørformet hulrum (svælg og mundhule). En rørformet resonator har ikke kun én egenfrekvens, den har flere. Og hvis rørets tværsnitsareal er ens i hele dets længde (således som det er ved shwa-vokalen), så vil der med en rørlængde på ca. 17 cm, som svarer til længden fra struben til læberne hos en mand, være resonanser ved ca. 500 Hz, 1500 Hz, 2500 Hz og så videre med 1000 Hz afstand i princippet uendeligt langt ud ad frekvensskalaen. Overføringsfunktionen for en sådan talekanal (det kaldes det faktisk i fonetikerjargon) er vist i figur 11. Overføringsfunktionen fortæller som nævnt i hvilke frekvensområder, der sker en forstærkning af deltonerne, når man 'sætter lyd på'.

Lydkilden, struben, giver ikke, som i flaskeeksemplet, en støj men derimod en periodisk, kompleks tone. Spektret af denne tone, kildespektret, er vist til venstre i figur 11. Vi kan se, at grundtonen er den kraftigste af deltonerne, og at disses amplitude falder, jo højere frekvensen bliver: i dette tilfælde falder kildespektret med 12 dB pr. oktav (en oktav svarer til en fordobling af frekvensen).

Ud over kilden og overføringsfunktionen skal vi også tage en tredje komponent i betragtning, nemlig den såkaldte *radiationseffekt* (udstrålings-effekt). Lydsvingningerne i den luftsøjle, der står i talekanalen, skal jo overføres til luften udenfor og sætte den i svingninger. Dette påvirker spektret af



**Figur 11.** Kildespektrum, overføringsfunktion for et 17 cm langt rør med samme tværsnit i hele længden, radiationsfunktion og resulterende spektrum.  $F_0$  er 125 Hz, og faldet på stemmekilden er 12 dB pr. oktav.

det signal, man måler uden for munden. I modsætning til bidraget fra talekanalen (overføringsfunktionen) er bidraget fra radiationen konstant og kan approksimeres af en diskantthævning på 6 dB pr. oktav, således som det er vist i figur 11.

Til højre i figur 11 gives spektret for den resulterende lyd, altså den lyd, vi kan måle uden for munden. Dette spektrum er summen af kilde-spektrret, overføringsfunktionen og radiationen, eller sagt på en anden måde: den lyd kilden giver, bliver modificeret af resonanserne i talekanalen og radiationen. Toppene i det resulterende spektrum kaldes *formanter*. De nummereres nedefra, og vi kan se, at førsteformanten (forkortet F1) i vores vokal er 500 Hz, andenformanten (F2) er 1500 Hz, tredjeformanten (F3) er 2500 Hz og fjerdeformanten (F4).

Den beskrivelse af vokalproduktionen, der er givet her, kaldes *kilde-filter-modellen*, fordi den betragter talekanalens overføringsfunktion og radiationen som et filter, der modificerer den lyd, som stemmekilden (struben) producerer. Formelt kan modellen beskrives i følgende ligning

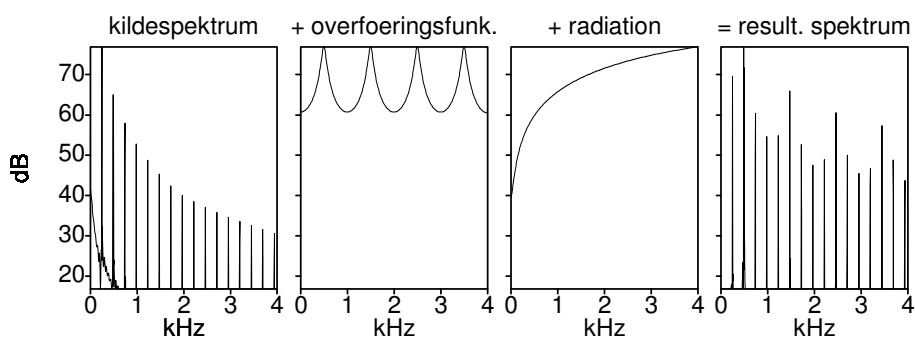
$$P(f) = U(f) + T(f) + R(f)$$

hvor  $P(f)$  er lydtrykket som funktion af frekvensen i det signal, man måler uden for munden,  $U(f)$  er lydtrykket i glottissignalet som funktion af frekvensen,  $T(f)$  er overføringsfunktionen, og  $R(f)$  er radiationseffekten som funktion af frekvensen.

Kilde-filter-modellen sætter fokus på den kendsgerning, at den lyd, vi iagttager uden for munden, er sat sammen af bidrag fra indbyrdes uafhængige komponenter, stemmekilde, ansatsrør og radiation. Radiationen er mindre interessant, for den bidrager altid med en diskantthævning på 6 dB/oktav, men det vigtige er, at de to andre komponenter, stemmekilde og ansatsrør *kan styres uafhængigt af hinanden*. Vi ved, at vi kan sige samme vokal på forskellige tonehøjder, med forskellige styrker og med forskellige stemmekvaliteter. Det vil sige, at vi kan holde overføringsfunktionen konstant, men ændre kilden. Vi kan også med samme tonehøjde, stemmekvalitet og samme stemmestyrke sige forskellige vokaler. Her holder vi kilden konstant og ændrer på overføringsfunktionen, dvs. på rørets udseende. Dette, at de to systemer arbejder uafhængigt af hinanden er egentligt banalt, men der er ikke desto mindre mange, der roder rundt i det.

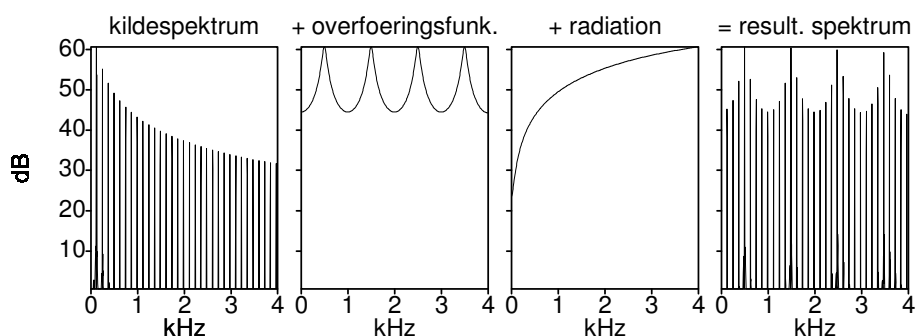
I figur 12-14 vises, hvordan det resulterende spektrum ser ud, når vi ændrer på tonehøjden og på kilde-spektrrets fald. Sammenligner vi det resulterende spektrum i figur 11 med dem i figur 12-14, kan vi se, at formanterne alle steder ligger ved 500, 1500, 2500 og 3500 Hz. Det er ikke forbavsende, for det er kun kilden og ikke overføringsfunktionen, vi har rørt ved.

I figur 12 er grundtonen ændret til 250 Hz i stedet for de 125 Hz i figur 11, og vi kan se, at afstanden mellem de harmoniske er blevet større (de ligger jo som vi ved ved hele multipla af grundtonen).

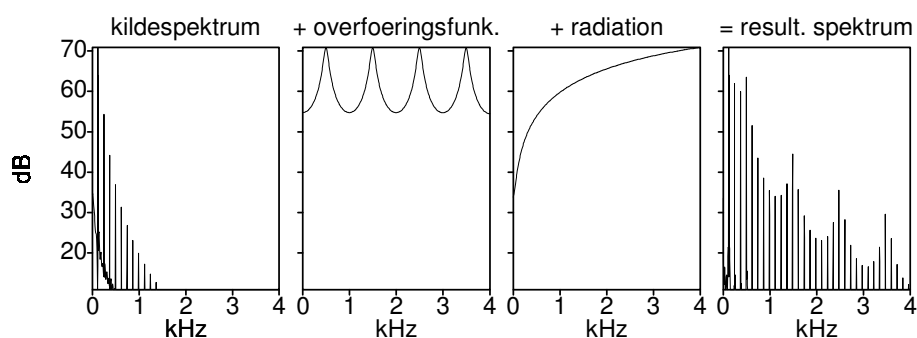


**Figur 12.** Kildespektrum, overføringsfunktion, radiationsfunktion og resulterende spektrum.  $F_0$  er 250 Hz, og faldet på stemmekilden er 12 dB pr. oktav.

I figur 13 har vi samme grundtone som i figur 11, nemlig 125 Hz, men faldet i deltonernes amplitude i kildespektret er her kun 6 dB pr. oktav, det vil sige, at faldet ikke er så kraftig som i figur 11. Det svarer til, at stemmen lyder mere komprimeret.



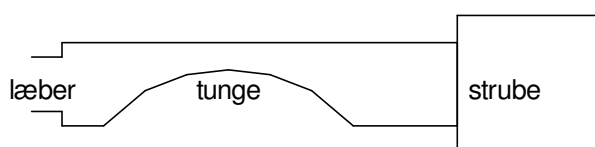
**Figur 13.** Kildespektrum, overføringsfunktion, radiationsfunktion og resulterende spektrum.  $F_0$  er 125 Hz, og faldet på stemmekilden er 6 dB pr. oktav.



**Figur 14.** Kildespektrum, overføringsfunktion, radiationsfunktion og resulterende spektrum.  $F_0$  er 125 Hz, og faldet på stemmekilden er 18 dB pr. oktav.

I figur 14 har vi et meget kraftigt fald i deltonernes amplitude i kildespektret, svarende til en underkomprimeret stemme. Disse ændringer i faldet på kildespektret kan også ses i de resulterende spektre: formanternes indbyrdes styrkeforhold er ændret, mens der ikke er sket noget med formantfrekvenserne. Det vi vil høre, er stadig en shwa-vokal, men med forskellig stemmekvalitet.

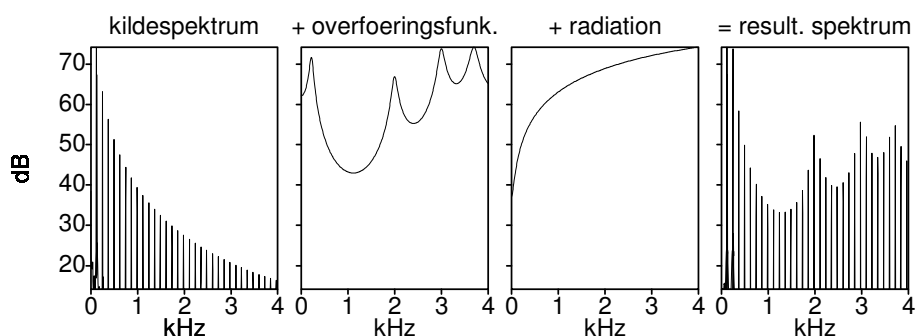
Hidtil har vi kun set på et rør, der har samme tværsnitsareal i hele sin længde, og således kun kan danne én vokal, shwa. Og det stemmer jo ikke med virkeligheden. For at kunne danne flere vokalkvaliteter, må vi modificere modellen, således som det er vist i figur 15: Vi sætter en variabel mun-



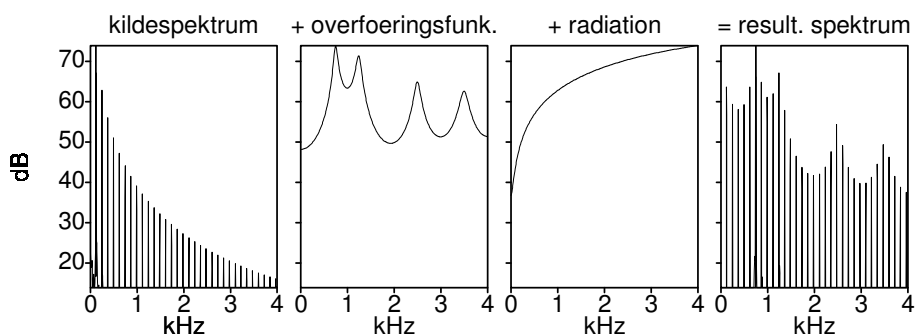
**Figur 15.** Modificeret rørmodel af taleorganerne med variabel åbning (læberne) og indsnævring (tungen).

ding på røret (læberne), og forsyner røret med en indsnævring (dannet af tungen). Munden kan gøres længere og snævrere, svarende til rundede læber, eller den kan gøres kortere og videre, svarende til urundede læber.

Indsnævringen, som tungen laver mod mundloftet og svælgets bagvæg, kan gøres større eller mindre, og den kan bevæges frem og tilbage. Ved at variere mundingens størrelse og konstriktionens størrelse og placering kan vi tilvejebringe ændringer i overføringsfunktionen, det vil sige flytte på formantfrekvenserne og således variere vokalkvaliteten. Figur 16 og 17 viser kildespektrum, overføringsfunktion, radiation og resulterende spektrum for vokalerne *i* og *α*. Vi kan se, at kildespektret (og naturligvis radiationen) er det samme for de to vokaler, nemlig med en  $F_0$  på 125 Hz og et fald på 12 dB/oktav. De har altså samme tonehøjde og stemmekvalitet. Det, der adskiller dem er overføringsfunktionen: Vokalen *i* har sine resonansfrekvenser (og dermed formanterne i det resulterende spektrum) ved 225, 2000, 3000 og 3700 Hz, og i vokalen *α* ligger de ved 750, 1250, 2500 og 3500 Hz.



**Figur 16.** Kildespektrum, overføringsfunktion, radiationsfunktion og resulterende spektrum for vokalen *i*.  $F_0$  er 125 Hz, og faldet på stemmekilden er 18 dB pr. oktav.



**Figur 17.** Kildespektrum, overføringsfunktion, radiationsfunktion og resulterende spektrum for vokalen *α*.  $F_0$  er 125 Hz, og faldet på stemmekilden er 18 dB pr. oktav.

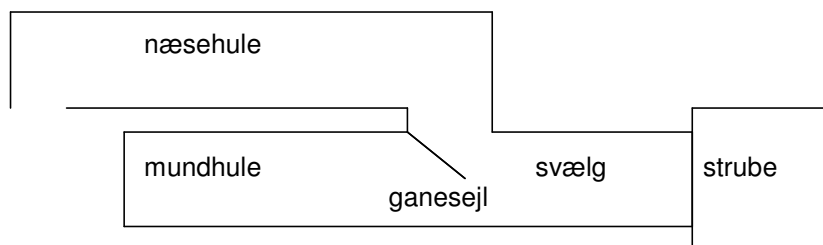
Som tidligere nævnt har røret i princippet et uendeligt antal resonanser, og der er dermed i det resulterende spektrum et uendeligt antal formanter. Imidlertid er det kun de tre laveste formanter, vi behøver interessere os for. Dette har to årsager. For det første er det de tre laveste formanter, der er mest følsomme for ændringer i rørets udseende, og derfor er de vigtigste bærere af den sproglige information. (I telefonen bliver al lydenergi over 3500 Hz – ca. svarende til overgrænsen af tredjeformantens variationsområde – skåret væk, og vi kan alligevel sagtens forstå, hvad der bliver sagt). For det andet bliver stemmekildens deltoner som vi har set svagere med stigende frekvens, og kommer man tilstrækkeligt langt op ad frekvensaksen, er der faktisk ikke noget energi tilbage for røret at forstærke.

## Konsonanter

Rørmodellen, i den form, vi indtil nu har beskæftiget os med, kan i princippet danne alle orale vokaler, men heller ikke andet. I det følgende skal det i grove træk omtales, hvordan andre lydtyper beskrevet ud fra modifikationer af modellen produceres.

### *Nasalkonsonanter*

*Nasalkonsonanter* har det til fælles med vokalerne, at de er periodiske lyde. De har altså samme stemmekildefunktion. Men røret er anderledes, mere kompliceret. Ved produktionen af nasalkonsonanter er det oprindelige rør lukket, men der er blevet tilføjet et nyt rør, nemlig passagen gennem næsen, se figur 18. Vi kunne også vende det om og sige, at vi har et rør, der er lukket i den ene ende (ved struben) og åbent i den anden (ved næseborene). Det er princippet i den oprindelige model, men derudover er der koblet et lukket



**Figur 18.** Model af taleorganerne ved produktion af nasalkonsonanter.

rør på, nemlig den lukkede mundhule. Vi ved, at et rør, der er lukket i den ene ende og åbent i den anden, har en overføringsfunktion med en række resonanstoppe, men vi har jo også det lukkede rør koblet på. Dette er også et luftfyldt hulrum, hvor luften kan sættes i svingninger af kildesignalet. Men de svingninger gør så at sige ingen nytte, røret er lukket, så svingningerne når ikke ud til omverdenen. Når vi derfor ser på spektret af den lyd der kommer ud af næsen, ser det ud som om der mangler deltoner i visse områder. Og det er faktisk også det, der er tilfældet. De deltoner i kildesignalet, som har sat luften i den lukkede mundhule i svingninger er blevet undertrykt. Man taler om, at mundhulen virker som en antiresonator, som bevirker, at der er nuller, antiformanter i spektret. Generelt kan det siges, at nasalkonsonanterne er relativt svage i forhold til vokalerne, og at størstedelen af deres energi ligger ved lave frekvenser, 200-300 Hz.

### *Frikativer*

De ustemte frikativer ligner vokalerne på den måde, at der er lukket til næsehulen. Men ellers adskiller de sig både hvad kilde og rør angår radikalt fra vokalerne. For det første er stemmekilden i den lukkede ende af røret 'slået fra', glottis er åben. For det andet er konstriktionen i røret så kraftig (passagen så snæver), at der dannes turbulens og dermed støj på dette sted. Lydkilden er altså flyttet frem i røret fra struben til artikulationsstedet. Den støj, der frembringes er tilnærmelsesvis en hvid støj. Overføringsfunktionen, og dermed kvaliteten af den støj, vi hører, bestemmes væsentligst af formen og

størrelsen af det hulrum, der ligger foran konstriktionen. Derfor lyder det alveolære *s* lysere end det palatale *sj*. Hulrummet bag konstriktionen (inklusive trachea og lungerne) spiller dog også en rolle for overføringsfunktionen, idet det fungerer som en antiresonator. Dette hulrum er stort, og dets antiresonans vil være lavfrekvent. Derfor har frikativer altid meget lidt energi i det lave frekvensområde.

De stemte frikativer (i dansk har vi kun *v*) ligner de ustemte, bortset fra at stemmekilden er 'slået til', stemmelæberne svinger. Det vil sige at vi har to lydkilder, en ved konstriktionen som producerer støj, og en ved glottis som producerer en periodisk lyd. Talekanalens form ved frikativerne betyder imidlertid, at alle de høje deltoner i glottisspektret dæmpes kraftigt, så kun de allerlaveste høres som en 'dyb brummen'.

Frikativen *h* (hvis man ellers vil klassificere den som en frikativ) er speciel. Her har vi ikke som i de andre frikativer flyttet kilden frem i røret. Den er i enden af røret ligesom ved vokalerne. Men den danner ikke en periodisk lyd. Glottis er netop så åben at der dannes turbulens, og dermed støj, på dette sted. Man kan beskrive *h* som en vokal (den gode gamle rørmodel), hvor den kilde, der sætter luften i røret i svingninger ikke er en periodisk kilde, men en støjkilde. Eller med andre ord: *h* er en hvisket vokal. Det er næsten rigtigt, men også kun næsten, for der er kobling til hulrummet under struben – trachea og lungerne – som virker som antiresonator. Derfor er førsteformanten i *h* normalt ret svag.

### **Lukkelyde**

Det, der er karakteristisk for de lydtyper vi hidtil har set på, er at de har 'deres egen lyd', som vi kan analysere og beskrive i relation til en ganske bestemt taleorganskonfiguration. Ved lukkelydene, derimod, kan vi kun beskrive deres akustiske forhold, hvis vi også ser på ændringerne af modellen som funktion af tiden. Lukkelydene er kendetegnet ved, at der er en fuldstændigt blokering et sted i røret i et lille stykke tid. For de ustemte lukkelydes vedkommende er der altså en periode, hvor der ikke er nogen lyd overhead. Og man behøver derfor ikke bekymre sig om deres akustiske beskrivelse. Det er – lidt frækt sagt – først når lukkelyden er forbi, når lukket opløses, at det bliver interessant. Selvom der ikke er nogen lyd under lukket, sker der noget alligevel. Der opbygges et lufttryk bag lukket, og idet lukket opløses, giver dette tryk anledning til en luftstrøm, som danner en kortvarig støj, en eksplosionsstøj, ved artikulationsstedet. Bortset fra forskellen i varighed kan lukkelydsekspllosioner beskrives ud fra nøjagtigt den samme model som frikativerne. Dette gælder de uaspirerede lukkelyde, *b*, *d* og *g*, og i dem sker der ikke mere. I *p* og *k* følges eksplosionen af en aspiration, som kan beskrives på samme måde som *h*. I *t* følges eksplosionen af en *s*-støj (en affrikation), som kan beskrives ud fra frikativmodellen ovenfor.

I de stemte lukkelyde (som vi ikke har nogen af i dansk) svinger stemmelæberne under lukket, og ligesom ved de stemte frikativer giver stemtheden sig til kende som en lavfrekvent brummen.