



Det här verket har digitaliserats vid Göteborgs universitetsbibliotek och är fritt att använda. Alla tryckta texter är OCR-tolkade till maskinläsbar text. Det betyder att du kan söka och kopiera texten från dokumentet. Vissa äldre dokument med dåligt tryck kan vara svåra att OCR-tolka korrekt vilket medför att den OCR-tolkade texten kan innehålla fel och därför bör man visuellt jämföra med verkets bilder för att avgöra vad som är riktigt.

This work has been digitized at Gothenburg University Library and is free to use. All printed texts have been OCR-processed and converted to machine readable text. This means that you can search and copy text from the document. Some early printed books are hard to OCR-process correctly and the text may contain errors, so one should always visually compare it with the images to determine what is correct.



Rapport

R22:1990

**Inverkan av högpassfiltrering på
taluppfattbarheten från högtalare
i efterklangsrrik miljö**

**Erik Rudolphi
Ulf Rosenberg**

V-HUSETS BIBLIOTEK, LTH



15000

400135427

Byggforskningsrådet

R22:1990

INVERKAN AV HÖGPASSFILTRERING PÅ TALUPPFATTBARHETEN I EFTERKLANGSRIK MILJÖ

Erik Rudolphi

Ulf Rosenberg

Denna rapport hänför sig till forskningsanslag 880855-1 från
Statens råd för byggnadsforskning till DNV Ingemansson AB,
Stockholm.

REFERAT

Vid talförstärkning med högtalare används ofta högpasfiltrering. Det görs dels för att kunna minska högtalarnas dimensioner, dels för att öka taluppfattbarheten, särskilt i efterklangrika lokaler. I syfte att finna regler för hur brytfrekvens och nivå bör sättas har taluppfattbarhet och naturlighet i återgivningen undersökts i en efterklangrik kyrka. Mätningar av taluppfattbarhet gjordes med RASTI och med lyssnartester med logatomer insatta i en bärfras. En subjektiv utvärdering av naturligheten vid olika inställningar gjordes också.

Bästa taluppfattbarhet erhöles vid 1000 Hz högpasfiltrering. Denna filtrering tycks även vara tillfredsställande ur naturlighetssynpunkt då den oförstärkta vokalklangen från talaren bidrar till upplevelsen. I lokaler med lång efterklangstid påverkas resultaten av lyssnartester med logatommetoden i hög grad av bärfrasens utformning.

I Bygghforskningsrådets rapportserie redovisar forskaren sitt anslagsprojekt. Publiceringen innebär inte att rådet tagit ställning till åsikter, slutsatser och resultat.

Denna skrift är tryckt på miljövänligt, oblekt papper.

R22:1990

ISBN 91-540-5174-6

Statens råd för byggnadsforskning, Stockholm

Svenskt Tryck Stockholm 1990

INNEHÅLLSFÖRTECKNING

	SAMMANFATTNING	5
1	FÖRORD	7
2	BAKGRUND	8
3	FAKTORER SOM PÅVERKAR TALUPPFATTBARHETEN	9
3.1	Energi och information i tal	9
3.2	Lågfrekvensmaskering	10
3.3	Direktljud, efterklang	10
3.4	Bakgrundsbuller	11
3.5	Efterklangstidens frekvensberoende	11
4	MÄTNING OCH BERÄKNING AV TALUPPFATTBARHET	12
4.1	Lyssnartester	12
4.2	$A_{l_{cons}}$	13
4.3	STI	13
4.4	RASTI	14
4.5	TDS- och ETC-analys	15
5	FÖRSÖKSUPPLÄGGNING	16
5.1	Laboratorieförsök	16
5.2	Fältförsök	17
5.2.1	Försöksuppställning	17
5.2.2	Filtreringar och nivåer	19
5.2.3	Lyssnartest	21
5.2.4	RASTI-mätningar	22
5.2.5	Ljudkvalitetsundersökning	22
5.2.6	Akustiska mätningar	23

6	RESULTAT	24
6.1	Lyssnartestet	24
6.2	RASTI-mätningarna	25
6.3	Ljudkvalitetsundersökningen	27
6.4	Akustiska mätningar	28
6.4.1	Efterklangstiden i Sofia kyrka	28
6.4.2	Systemets impulssvar	29
7	DISKUSSION	30
8	LITTERATUR	32
	BILAGA 1 Resultat av lyssnartest	34
	BILAGA 2 Resultat av mätningar med RASTI	36
	BILAGA 3 Resultat av ljudkvalitetsundersökning	37
	BILAGA 4 Matematisk beskrivning av STI	39

SAMMANFATTNING

Vid förstärkning av tal via högtalare är bandbredden en kritisk parameter. Begränsas bandbredden i basen kan högtalarnas dimensioner minska. Taluppfattbarheten i efterklangsrisk miljö kan dessutom ofta förbättras med en sådan bandbegränsning. Dock saknas fasta regler för val av brytfrekvens och lämplig nivå. Denna undersökning har gjorts för att söka öka kunskaperna på området.

Lyssnartester och mätningar med RASTI har gjorts i Sofia kyrka på Söder i Stockholm. Taluppfattbarheten har undersökts vid olika filterringar och med olika nivåkillnader mellan direktljud och efterklang, motsvarande olika avstånd till högtalaren.

Vid lyssnartesterna användes enstaviga ord, s.k. logatomer, insatta i bärfraserna "Det är ___ vi hör." eller "Nu är det ___ vi hör.". En jämförelse visade att i lokaler med lång efterklangstid är längden hos frasens första del av stor betydelse.

Sammanfattningsvis erhöles följande resultat:

- Högpasfilterring minskar den totalt utmatade ljudeffekten vid talförstärkning. Detta minskar risken för återkoppling och lämnar följaktligen utrymme för högre förstärkning vid höga frekvenser, vilket är positivt för taluppfattbarheten.
- Brytfrekvensen vid högpasfilterring bör läggas kring 1000 Hz för bästa taluppfattbarhet.

- I många efterklangsrika lokaler är troligtvis den oförstärkta vokalklangen från talaren tillräcklig, både för taluppfattbarheten och för att ge naturlig talklang även vid högpass-filtrerad högtalarförstärkning.

- Resultatet av lyssnarundersökningar med logatommetoden i lokaler med lång efterklangstid påverkas av den inledande bärfrasens längd.

1 FÖRORD

Denna undersöknings första del har utförts som ett examensarbete i elektroakustik vid Tekniska Högskolan i Stockholm, på uppdrag av DNV Ingemansson AB. Med stöd från Byggforskningsrådet har arbetet därefter kompletterats på några viktiga punkter.

Vi vill här framföra vårt tack till alla de som hjälpt oss i vårt arbete.

Speciellt vill vi tacka Sten Hagberg vid DNV Ingemansson AB, Johan Liljencrantz vid Institutionen för Talöverföring och Musikakustik, Tekniska Högskolan, Björn Hagerman, Ann-Catrin Lindblad och Ove Till vid Institutionen för Teknisk Audiologi, Karolinska Institutet, Fred Lundin vid Avdelningen för grundteknik, Televerket och Anders Olson vid Elfa Studio AB för värdefulla råd och synpunkter samt för hjälp med utrustning.

Ett särskilt tack riktas också till personalen vid Sofia kyrka för stor hjälpsamhet och stort tålamod.

2 BAKGRUND

Talförstärkning via högtalare är i dag något mycket vanligt, inte minst i offentliga lokaler. Alla torde också någon gång ha upplevt hur buller eller lång efterklang gjort det svårt att uppfatta vad som sägs. Det kan vara i en kyrka, på en järnvägsstation eller i en idrottshall. Behoven av väl dimensionerade anläggningar för talförstärkning är alltså stora.

I denna undersökning har intresset koncentrerats på att hitta högsta möjliga undre gränshänsyn för ett högtalarsystem utan att taluppfattbarhet eller naturlighet påverkas alltför mycket.

Stora högtalare, kanske bra för musikåtergivning, innebär inte alltid bra taluppfattbarhet. Särskilt i efterklangsrika lokaler medför en bredbandig återgivning ofta försämrade taluppfattbarhet. Begränsad bandbredd i basen minskar dessutom högtalarnas dimensioner.

Låga kostnader är ofta ett viktigt krav, vilket förutsätter att anläggningen inte överdimensioneras. I många fall är små och diskreta högtalare också önskvärda av estetiska skäl. Högtalaranläggningar i kulturbyggnader, som gamla kyrkor, innebär ofta en känslig avvägning mellan ljudkvalitet och estetik.

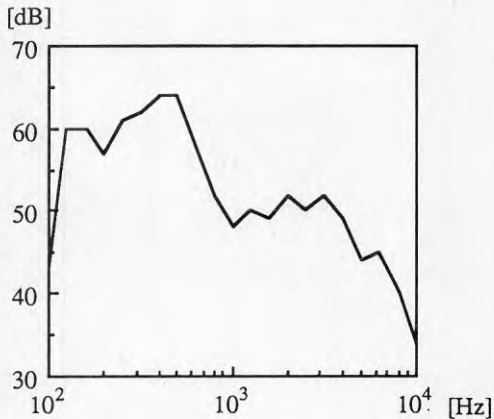
FAKTORER SOM PÅVERKAR TALUPPFATTBARHETEN

Då ett talat meddelande överförs mellan två personer påverkas uppfattbarheten av en mängd faktorer:

- Talspektrums utseende och dess variation över tiden.
- Absolut och relativ ljudnivå.
- Hur örat analyserar talsignalen.
- Lyssningsmiljöns egenskaper, t.ex. förekomst av buller eller lång efterklangstid.

3.1 Energi och information i tal

Figur 3.1 visar långtidsspektrum för tal hos en manlig talare. Här framgår att huvuddelen av energin finns i frekvenser under 500 Hz. Den för taluppfattbarheten viktigaste informationen finns emellertid vid frekvenser kring 2 kHz [1]. Det är alltså möjligt att filtrera bort låga frekvenser utan att minska taluppfattbarheten nämnvärt.



Figur 3.1 Talets energiinnehåll som funktion av frekvensen hos en manlig talare.

3.2 Lågfrekvensmaskering

Man vet att örats förmåga att urskilja ljud minskar då störnivån ökar. Försök har gjorts där två toner spelats upp [13]. Nivån på den ena tonen har därefter ökat tills den andra inte längre kunnat uppfattas. Resultaten av dessa försök har presenterats som hörtröskelkurvor, s.k. maskeringskurvor. Av dessa framgår att en låg ton lätt maskerar en högre. Det omvända förhållandet gäller inte alls i samma utsträckning.

3.3 Direktljud, efterklang

Undersökningar, [5,8], har visat att den ljudenergi som når örat inom 30 ms efter direktljudet är viktig för taluppfattbarheten, medan den energi som anländer senare än 95 ms har enbart menlig inverkan.¹ Detta har ett göra med örats s.k. integrationstid. Örat förmår inte separera ljud som är åtskilda mindre än 30 ms. Över 30 ms ökar separationsförmågan gradvis och vid 95 ms uppfattas ljuden som helt åtskilda.

¹ Utgående från detta har ett signal-störindex definierats som

$$S/N = 10 \log \frac{\int_{0 \text{ ms}}^{95 \text{ ms}} a(p,t) p^2(t) dt}{\int_{95 \text{ ms}}^{\infty} p^2(t) dt} \quad (3.1)$$

där $p(t)$ är ljudtrycket, och $a(p,t)$ är en viktsfunktion beroende av örats integrationsegenskaper.

En slutsats av detta är att kort efterklangstid gynnar taluppfattbarheten, liksom närhet till ljudkällan, så att andelen direktljud blir stor. Undersökningar [10] har visat att taluppfattbarheten förbättras om lyssnaren befinner sig inom ett avstånd på tre gånger efterklangsradien.² På längre avstånd har lyssnaren alltså ingen nytta av högtalaren.

3.4 Bakgrundsbuller

Ett viktigt krav för god taluppfattbarhet är lågt bakgrundsbuller. Bullernivåns inverkan på taluppfattbarheten kan inom vissa gränser kompenseras med en höjd talnivå. En tumregel är att talets direktljudsnivå bör ligga minst 10 dB över störnivån för god taluppfattbarhet. Höjs talnivån alltför mycket kommer örats inre distorsion att motverka förbättringen [2].

3.5 Efterklangstidens frekvensberoende

I många stora lokaler är efterklangstiden påtagligt längre vid låga frekvenser än vid höga. Detta medför att den lågfrekventa energin i talet ökar det maskerande lågfrekventa efterklangsfältet med sämre taluppfattbarhet som följd.

Sammantaget finns alltså en rad samverkande faktorer som ger anledning att förmoda att högpassfiltrering är en lämplig metod för att öka taluppfattbarheten vid talförstärkning i efterklangsrika lokaler.

² Efterklangsradien är det avstånd från ljudkällan där direktljud och efterklang är lika starka, och ges av

$$r_r = 0,057 \sqrt{QV/T} \quad (3.2)$$

där Q är källans riktfaktor, V är rummets volym och T är efterklangstiden. Se även kapitel 4.2 om AI_{cons} .

4 MÄTNING OCH BERÄKNING AV TALUPPFATTBARHET

Taluppfattbarhet är ett perceptivt begrepp och kan därmed inte direkt anges i fysikaliskt mätbara storheter. För att mäta taluppfattbarheten är man därför hänvisad till **subjektiva** mätmetoder, i form av lyssnartester.

Lyssnartester är arbetskrävande, och därför har även fysikaliska mätmetoder utvecklats. Metoder finns också för att redan på projekteringsstadiet för en anläggning beräkna taluppfattbarheten. Dessa metoder har korrelerats mot den subjektivt upplevda taluppfattbarheten genom lyssnartester. I detta kapitel presenteras några av de viktigaste metoderna för att ange taluppfattbarhet.

4.1 Lyssnartester

Vid lyssnartester sänds tal genom det system som skall undersökas, ett rum via en högtalaranläggning eller liknande. Ett antal lyssnare skriver ner vad de uppfattar av talinformationen. Andelen fel uppfattade ord ger en uppfattning om taluppfattbarheten.

Som talmaterial används vanligtvis listor med s.k. logatomer. Dessa är betydelselösa enstaviga ord av formen konsonant-vokal-konsonant. Resultatet mäts i procent artikulationsförluster, dvs. andelen felaktigt uppfattade vokaler, initial- och finalkonsonanter och ord i förhållande till det totala antalet. Logatomerna sätts ofta in i en bärfras, t.ex. "Det är _ _ _ vi hör". Även andra typer av talmaterial används, t.ex. vanliga ord i olika kombinationer.

Lyssnartester kräver ett stort statistiskt underlag för att uppnå god noggrannhet, vilket kräver ett flertal lyssnare. Detta speciellt som de individuella variationerna i förmågan att uppfatta tal är stora. Flera olika talare ökar också testets objektivitet.

4.2 Al_{cons}

Al_{cons} (Articulation Loss of Consonants) [10] är en enkel metod att beräkna taluppfattbarheten i rum där efterklngen är den dominerande faktorn. Metoden bygger på empiriska samband mellan rummets volym (V), efterklangstiden (T) och avståndet till ljudkällan (d).

$$Al_{\text{cons}} = \frac{200d^2T^2}{V} + a \text{ [%]} \quad \text{för } d < d_c \quad (4.1)$$

$$Al_{\text{cons}} = 9T + a \text{ [%]} \quad \text{för } d > d_c \quad (4.2)$$

där a är en korrektionsfaktor beroende av talarens och lyssnarens individuella förmåga och d_c är den kritiska radien, definierad som

$$d_c = 0,2\sqrt{QV/T} \quad (4.3)$$

vilket är lika med 3,16 gånger efterklngsradien. Q är ljudkällans riktfaktor.

Al_{cons} är definierat för efterklangstiden vid 1400 Hz och tar man således inte hänsyn till variationer ljudkällans i frekvensgång eller variationer hos efterklangstiden vid olika frekvenser.

4.3 STI

STI har utvecklats för att fysikaliskt **mäta** ett systems taluppfattbarhet och är den för närvarande tillförlitligaste metod som finns att tillgå. Den är emellertid krävande både i tid och tekniska resurser. Metoden fungerar på motsvarande sätt som kontrastöverföringsfunktionen inom optiken. En testsignal skickas genom systemet och analyseras på mottagarsidan. Resultatet ges i form av ett talöverföringsindex, STI (Speech Transmission Index).

Testsignalen består av brus uppdelat i sju oktavband med mittfrekvenser från 125 Hz till 8 kHz. Oktavbandsnivåerna är valda så att signalens spektrum efterliknar vanligt tal. Signalen amplitudmoduleras med sinusvåg i fjorton modulationsfrekvenser mellan 0,63 Hz och 12,5 Hz. Oktavbanden och modulationsfrekvenserna bildar tillsammans en matris med 7x14 mätpunkter.

På mottagarsidan kommer bakgrundsbuller och den efterklang som signalen alstrat att orsaka en minskning av **modulationsamplituden**, dvs. en utslätning av signalen. Genom att jämföra den mottagna **signalens modulationsdjup** med **testsignalens** kan ett signal-störförhållande beräknas. Ett medelvärde bildas för alla modulationsfrekvenser inom ett oktavband. Metoden korrigerar också för den maskering som ett lägre frekvensband förorsakar på ett högre.

Signal-störförhållandena i varje oktavband vägs slutligen samman och normaliseras till ett STI-värde mellan 0 och 1, där 1 betyder förlustfri talöverföring. Oktavbandsvärdena viktas efter deras betydelse för taluppfattbarheten.

För den intresserade ges en mer ingående matematisk beskrivning av STI i bilaga 4.

4.4 RASTI

En förenklad STI-metod, RASTI (Rapid Speech Transmission Index) använder endast nio mätpunkter av sammanlagt 98 i oktavbanden 500 Hz och 2 kHz. RASTI har utvecklats för att få en snabb och enkel mätmetod utan att pruta alltför mycket på noggrannheten. Metoden är standardiserad enligt ISO. Mätutrustningen består av två bärbara enheter. Resultaten visas direkt på mottagarens display. RASTI tar inte samma hänsyn till lågfrekvensmaskering, eftersom metoden endast arbetar med två oktavband.

4.5 TDS- och ETC-analys

Ett relativt nytt och kraftfullt redskap för att analysera akustiska problem är TDS-analys (Time Delay Spectrometry). Testsignalen består av en sinuston som sveps över det frekvensband man vill undersöka. På mottagarsidan sveps ett bandpassfilter med tidsfördröjning. På detta sätt kan frekvensgången i ett tidsfönster studeras. Tidsfönstertechniken medför att endast den signal som når mottagaren inom detta fönster registreras. All övrig signal, både tidigare och senare, undertrycks. Det går t ex att titta på direktljudet från en högtalare (utan att rummet påverkar resultatet) genom att välja en tidsfördröjning till tidsfönstret som motsvarar ljudets gångtid från ljudkälla till mikrofon. Genom att göra en serie svep med successiv tidsförskjutning av filtret kan också kurvlandskap med tid- energi- och frekvensaxlar ritas. Ur dessa data kan även fas- och Nyquistdiagram ritas.

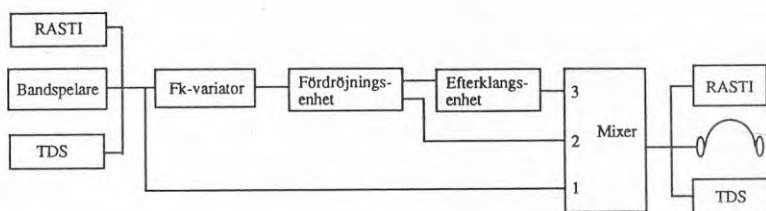
Med en liknande metod, ETC (Energy Time Curve), kan även impuls-svaret bestämmas. Ett sinussvep görs som ovan, men i stället för att svepa ett filter analyseras hela frekvenssvaret. Frekvenssvaret transformerar till tidsdomänen.

TDS och ETC kan mätas med en s.k. TEF-analysator (Time Energy Frequency). Med hjälp av denna kan även AI_{cons} och RASTI beräknas.

5 FÖRSÖKSUPPLÄGGNING

5.1 Laborieförsök

För att få en uppfattning om lämplig nivåinställning och filtrering för fältförsöken gjordes förförsök i laboratorieskala. Lokalens akustik simulerades med hjälp av en efterklangsenhet och tidsförskjutningen mellan olika ljudkällor åstadkoms med en digital fördröjningsenhet. Kopplingsschema visas i figur 5.1. Utvärderingen skedde dels subjektivt genom lyssning, dels genom mätningar med RASTI och TDS. Resultaten har inte vidare behandlats utan endast använts som underlag för försöken i fält.



Figur 5.1 Kopplingsschema för laborieförsöket. Mixerns ingång 1 motsvarar ljudet direkt från talaren och ingång 2 motsvarar ljudet från högtalare. Via ingång 3 simuleras efterklangsfältet.

5.2 Fältförsök

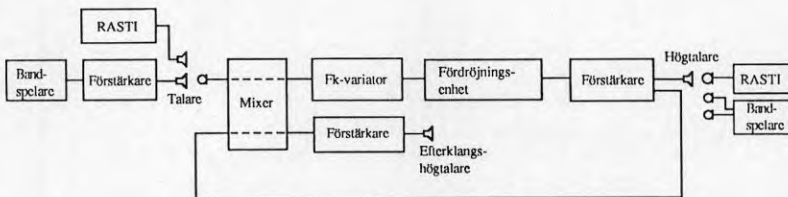
Fältförsöken har genomförts i två omgångar. Första undersökningen gjordes inom ramen för ett examensarbete under sommaren 1988. En kompletterande undersökning genomfördes i februari 1989 för att söka rätta ut några av de frågetecken som blev ett resultat av de första försöken.

För fältförsöken valdes Sofia kyrka på Södermalm i Stockholm. Denna kyrka har stor volym och lång efterklangstid och anses därför ur tal-synpunkt som akustiskt svår.

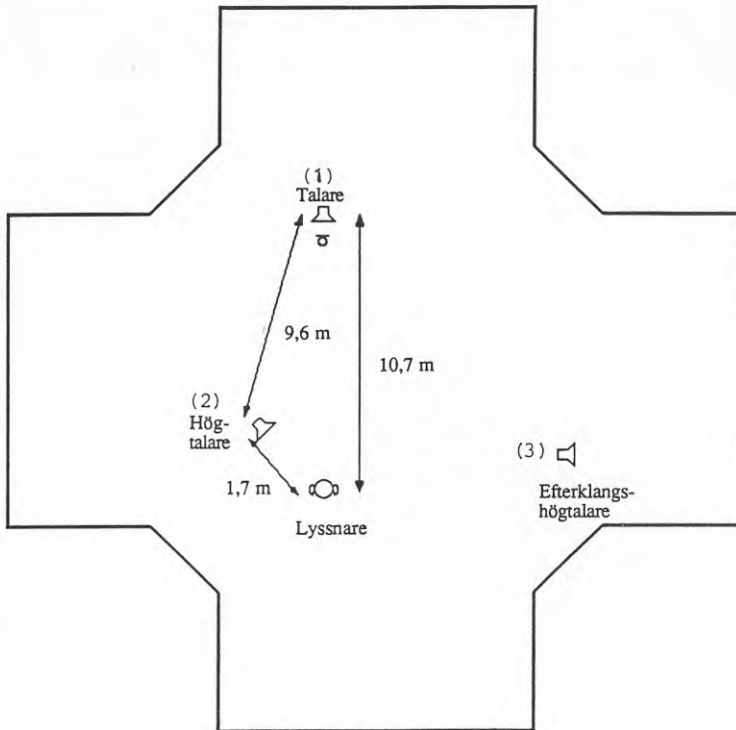
Mätningar gjordes med RASTI. Logatomlistor spelades in med konst-huvud för senare utvärdering i lyssnartest.

5.2.1 Försöksuppställning

Fältförsökets uppställning utformades för att efterlikna förhållandena under en gudstjänst. Figurerna 5.2 och 5.3 visar hur enheterna kopplades samman och högtalarnas placering i kyrkan.



Figur 5.2 Kopplingsschema för försöksuppställning i Sofia kyrka.



Figur 5.3

Principskiss över försöksuppställning i Sofia kyrka. En högtalare (1) placerad vid predikstolen representerade en talare. En mikrofon framför talaren fångade upp signalen som filterades och skickades via fördröjningsenhet och förstärkare till högtalaren för återgivning (2). Denna högtalare placerades så att lyssnarpositionen låg inom högtalarens direkrfält. Ytterligare en högtalare placerades på stort avstånd från talare och lyssnare för att alstra ett efterklangsfält (3). På detta sätt kunde efterklangsfältet hållas på en exakt nivå. Samtliga nivåer kunde regleras med hjälp av mixern.

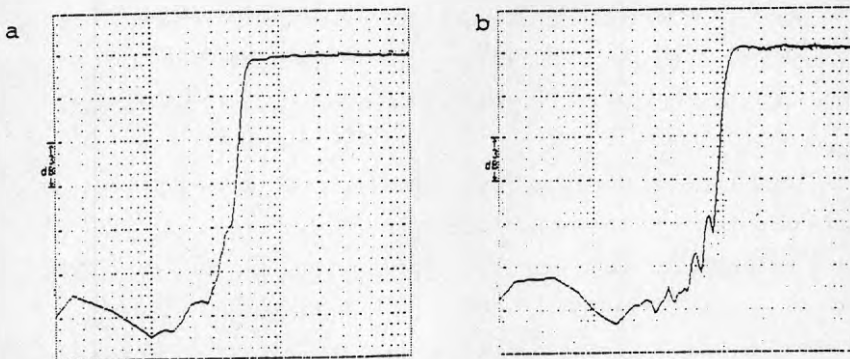
5.2.2 Filtreringar och nivåer

Vid försöket varierades dels filtreringen av talsignalen dels ljudnivån hos efterklangsfältet. Tre filtreringar provades:

- Rak kurva - ingen filtrering
- Högpas 500 Hz
- Högpas 1000 Hz

Frekvenskurvor för högpasfiltreringarna visas i figur 5.4.

För filtreringarna användes en fk-variator. Detta innebar att dämpningen endast var 18 dB utanför passbandet. Den bedömdes dock vara tillräckligt för att inte påverka resultatet nämnvärt.



Figur 5.4 Frekvenskurvor för högpasfiltreringarna.
Horisontella skalan logaritmisk 10-10000 Hz.
(a) högpas 500 Hz, (b) högpas 1000 Hz.

Efterklangsfältets, via högtalare 3, nivå varierades i förhållande till direktljudet från högtalare 2 vid sidan lyssnaren genom att ändra nivån från efterklangshögtalaren.

Vid den första undersökningen provades två inställningar:

- Direktljud/Efterklang = +3 dB
- Direktljud/Efterklang = -3 dB

I det senare kompletterande försöket provades att ytterligare öka efterklangsfältet för att eventuellt kunna påvisa en större effekt av filtreringarna:

- Direktljud/Efterklang = -9 dB

Förstärkningen justerades så att nivån från högtalare 2 i lyssnarpositionen låg 9 dB över motsvarande nivå från talaren, högtalare 1, mätt med stationär brussignal. Detta var den största möjliga nivåskillnad utan återkoppling.

Vid injustering av ljudnivåerna från högtalarna användes vitt brus bandbegränsat med ett 2 kHz oktavbandsfilter så att där frekvensgången var opåverkad oavsett vald filtrering. Nivåerna mättes med bullermätare.

Då energimaximum för talspektrum ligger kring 200 Hz, kommer en avsevärd del av ljudenergin att filtreras bort vid brytfrekvenserna 500 och 1000 Hz. Den **totala** ljudnivån varierade alltså i hög grad med filtreringen.

5.2.3 Lyssnartest

För lyssnartestet användes fem logatomlistor med vardera 50 ord, vilka spelades upp över systemet en gång för varje inställning. Resultatet spelades in i konsthuvdstereo. Logatomerna lästes tillsammans med bärfrasen "Det är _ _ _ vi hör."³

I den andra försöksomgången undersöktes även bärfrasens inverkan på taluppfattbarheten. I lokaler med mycket lång efterklangstid hinner en kort bärfras inte bygga upp nivån hos efterklangsfältet. Taluppfattbarheten mätt med lyssnartester bör därför påverkas av den inledande bärfrasens längd, något som det första försöket indikerat.

En ny inspelning av listorna gjordes med den längre bärfrasen "Nu är det _ _ _ vi hör.". Bärfrasens första del ökade därmed från ca 0,3 s för "Det är.." till ca 0,7 s för "Nu är det..". Utvärdering med den nya bärfrasen gjordes med rak frekvensgång och med 1000 Hz högpassfiltrering.

Vid utvärderingen av de båda undersökningarna deltog vardera tio försökspersoner, samtliga under 30 år. Fem av försökspersonerna medverkade i båda undersökningarna. Det inspelade materialet spelades upp i halvöppna hörtelefoner typ Yamaha HP-1. Listornas ordning kastades om så att varje försöksperson fick höra en ny lista för varje ny inställning. Endast två personer fick höra samma lista för en viss inställning, detta för att få bästa statistiska spridning.

³ Konsthuvud och listor lånades från Institutionen för Teknisk Audiologi vid Karolinska Institutet. En inspelning av listmaterialet gjordes på Ingemanssons.

5.2.4 RASTI-mätningar

Mätningar med RASTI gjordes för samtliga inställningar. "Talaren", högtalare 1, vid predikstolen ersattes då med RASTI's sändarenhet. Mätningarna gjordes under perioder på 32 s.

5.2.5 Ljudkvalitetsundersökning

Naturligheten i ljudåtergivningen är en viktig faktor som begränsar hur långt man kan gå upp i filtrering och förstärkning. För att undersöka detta gjordes även här ett antal inspelningar för senare bedömning av försökspersonerna. En sekvens ur logatombandet spelades upp två gånger, en gång utan talförstärkning, en gång med hela systemet inkopplat. Filtringarna av signalen var desamma som i logatomprovet. Nivåskillnaden mellan direktljud och efterklang var +3 dB hela tiden.

Det talförstärkta ljudets nivå varierades relativt nivån från "talaren". Följande nivåer provades:

- ± 0 dB
- + 6 dB
- + 9 dB
- + 12 dB

Nivåerna är givna relativt talarens ljudnivå i lyssnarpositionen.

Försökspersonerna ombads besvara följande frågor:

- 1) Hör du att högtalaren vid sidan är inkopplad?

Kryssa för Ja eller Nej.

Om du svarat ja på den första frågan, dvs. du har hört att högtalaren är inkopplad, besvarar du följande två frågor:

2) Varifrån tyckte du att ljudet kom?

1) Från "talaren" vid altaret.

X) Från en obestämd plats.

2) Från högtalaren vid sidan.

3) Tänk dig att högtalaren vid altaret är en verklig talare. Lyssna och jämför hur du tycker att högtalaren vid sidan färgar ljudet. Bry dig inte om, om återgivningen är starkare än talaren vid altaret. Betygsätt ljudåtergivningen med

5) Utmärkt

4) Bra

3) Acceptabel

2) Dålig

1) Oduglig

5.2.6 Akustiska mätningar

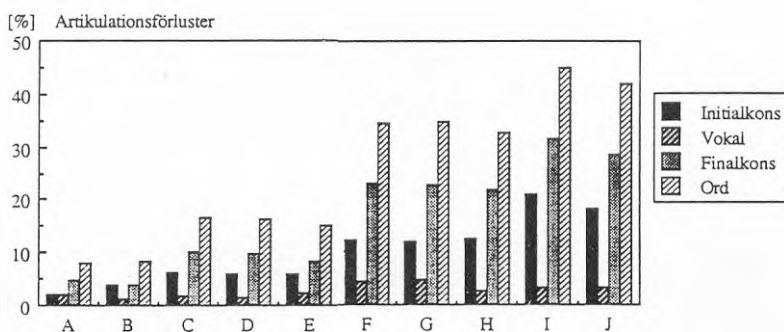
Ursprungligen avsågs att använda TEF-analys för mätning av kyrkans akustik och inmätning av högtalarsystemet. Den TEF-analysator som fanns att tillgå visade sig dock inte tillförlitlig. Kompletterande akustiska mätningar har därför gjorts på traditionellt sätt.

En studie av högtalarsystemets impulssvar gjordes genom att spela upp korta pulser genom "talaren". Via högtalarsystemet inkopplat spelades impulssvaret in i lyssnarpositionen för senare utvärdering med realtids-analysator. Inspelningen gjordes med 9 dB nivåskillnad mellan direktljud och efterklangsfält, dvs samma som i försöket.

6 RESULTAT

6.1 Lyssnartestet

Här mättes relativa förlusten av initialkonsonanter, vokaler, finalkonsonanter och ord. Med förlust av ord menas ord, där en eller flera delar missuppfattats. Lyssnartestetets resultat visas i figur 6.1.



Figur 6.1 Resultat av lyssnartestet i Sofia kyrka. Staplarna representerar procentuella förlusterna av initialkonsonanter, vokaler, finalkonsonanter och ord. Följande inställningar användes:

	Filtrering	Direktfält/Efterklingsfält
A	Rak frekvensgång	+3 dB
B	Högpäss 1000 Hz	+3 dB
C	Rak frekvensgång	-3 dB
D	Högpäss 500 Hz	-3 dB
E	Högpäss 1000 Hz	-3 dB
F	Rak frekvensgång	-9 dB
G	Högpäss 500 Hz	-9 dB
H	Högpäss 1000 Hz	-9 dB
I	Rak frekvensgång	-9 dB Förlängd bärfras
J	Högpäss 1000 Hz	-9 dB Förlängd bärfras

Lyssnartester av detta slag kräver omfattande material för att ge statistiskt signifikanta resultat. En så omfattande undersökning har inte varit ekonomiskt möjlig. Några tendenser tycks dock klara även om de inte är statistiskt signifikanta:

- Uppfattbarheten av konsonanter förbättras vid högpasfiltering. Tendensen är mer uttalad vid höga efterklangsnivåer.
- Uppfattbarheten av vokaler är i samtliga fall god och påverkas inte märkbart av filtreringarna.
- Resultaten av taluppfattbarhetsprov med logatommetoden i efterklangsrika lokaler påverkas i hög grad av bärfrasens utformning.

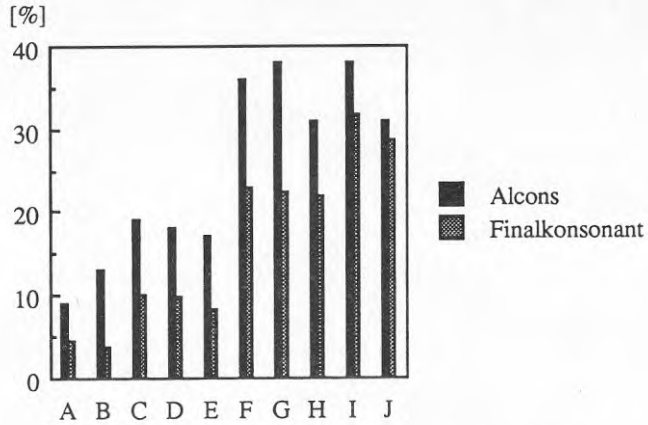
Det bör påpekas att inställningarna A - E och F - J utvärderats vid olika tillfällen och med delvis olika försökspersoner. Jämförelser mellan dessa grupper bör därför göras med försiktighet.

Siffervärden och skattade standardavvikelser redovisas i bilaga 1.

6.2 RASTI-mätningarna

Resultaten av RASTI-mätningarna visas i figur 6.2. För att underlätta jämförelser med lyssnartestet har RASTI-värdena räknats om till AI_{cons} enligt den empiriska formel som ges i bilaga 4. Siffervärden redovisas i bilaga 2.

Uppmätta RASTI-värden skiljer sig anmärkningsvärt mycket från lyssnartestets konsonantförluster. Denna skillnad minskar dock med den längre bärfrasen.



Figur 6.2 Resultat av RASTI-mätningar gjorda vid försök i Sofia kyrka omräknat till $A1_{\text{cons}}$. För jämförelse har även lyssnartestetets finalkonsonantförluster lagts in. Följande inställningar användes

	Filtrering	Direktfält/Efterklangsfält
A	Rak frekvensgång	+3 dB
B	Höpass 1000 Hz	+3 dB
C	Rak frekvensgång	-3 dB
D	Höpass 500 Hz	-3 dB
E	Höpass 1000 Hz	-3 dB
F	Rak frekvensgång	-9 dB
G	Höpass 500 Hz	-9 dB
H	Höpass 1000 Hz	-9 dB
I	Rak frekvensgång	-9 dB Förlängd bärfras
J	Höpass 1000 Hz	-9 dB Förlängd bärfras

6.3 Ljudkvalitetsundersökningen

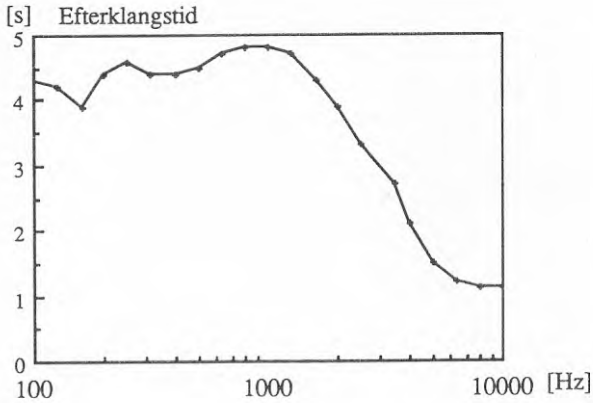
Ljudkvalitetsundersökningen uppvisar små variationer för de olika inställningarna. Lyssnaren tycks dock lättare lokalisera ljudkällan "rätt", dvs till talaren, då ljudet filtreras.

Betygen på ljudkvaliteten är över lag höga. Inte i något fall har betyget "oduglig" använts. Medelbetygen är något lägre vid 1000 Hz högpasstrerering, men spridningen på svaren är stor och tendensen kan inte sägas vara säker. Resultaten redovisas i bilaga 3.

6.4 Akustiska mätningar

6.4.1 Efterklangstiden i Sofia Kyrka

Som framgår av figur 6.3 ligger efterklangstiden i Sofia kyrka kring 4,7 s. För låga frekvenser är den något kortare och kurvan faller snabbt av för frekvenser över 1,25 kHz. Det senare till stor del beroende på den ökande luftabsorptionen vid höga frekvenser. Ur försökssynpunkt hade en lokal med ännu längre efterklangstider vid låga frekvenser troligen förstärkt tendenserna i försöket. Som synes har kurvan sin topp vid den högsta brytfrekvensen, 1 kHz.

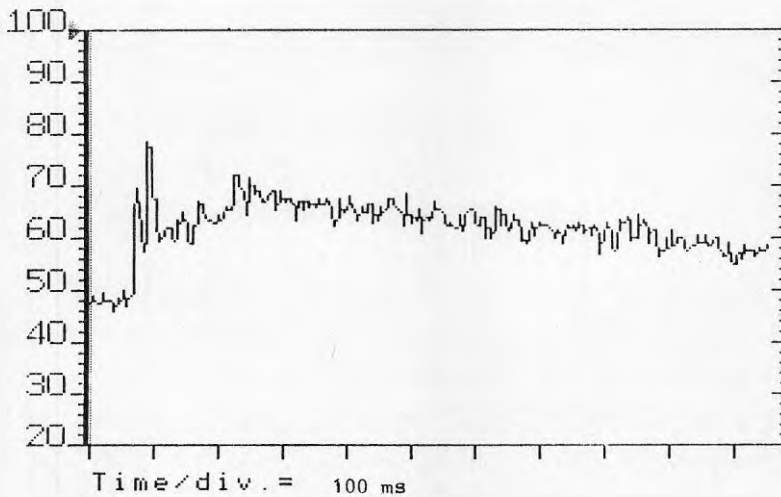


Figur 6.3 Efterklangstiden i Sofia kyrka.

6.4.2 Systemets impulssvar

Högtalarsystemets impulssvar visas i figur 6.4. Den första toppen svarar mot direktljudet från talaren. Nästa topp svarar mot direktljudet från högtalaren vid lyssnarens sida. Därefter byggs efterklangsfältet upp av efterklangshögtalaren vars direktljud anländer ca 100 ms efter direktljudet från talaren.

Den första "uppförsbacken" på efterklagsdelen av kurvan kan tyckas onaturlig, men detta bör likna förloppet i ett distribuerat högtalarsystem. Efterklangsfältet byggs då upp av ljudet från flera högtalare med olika tidsfördröjning.



Figur 6.4 Högtalarsystemets impulssvar i Sofia kyrka mätt i lyssnarpositionen.

7 DISKUSSION

Undersökningarna visar på en rad fördelar med högpasfiltrering. Ur materialet kan följande slutsatser dras.

- Brytpunkten bör läggas kring 1000 Hz för bästa taluppfattbarhet.
- Högpasfiltrering minskar den totala utmatade ljudeffekten vid talförstärkning, med minskad risk för återkopplingsproblem. Detta ger följaktligen utrymme att öka förstärkningen.
- Den minskade ljudeffekten vid högpasfiltrering gör att lyssnaren lättare lokaliserar ljudet "rätt" till talaren.
- I många efterklangrika lokaler är troligtvis den oförstärkta vokalklangen från talaren tillräcklig, både ur naturlighetssynpunkt och uppfattbarhetssynpunkt.

Brytpunktens optimala läge och filtrets branthet påverkas av lokalens efterklangskurva. Dessa parametrar måste därför alltid finjusteras i varje ny lokal "på känsla": Mer efterklang i basen kräver hårdare filtrering. Den undersökta lokalens gynnsamma efterklangskurva pekar dock på att brytpunkten i de flesta fall bör läggas närmare 1000 Hz än 500 Hz för bästa taluppfattbarhet.

Påståendet om vokalklangen är inte helt underbyggt i denna undersökning. Kunskaper saknas om hur mycket filtrering vår hörsel tål innan naturligheten påverkas och frågan är värd en fortsatt utredning. Är vår förmodan riktig kan det innebära en avsevärd minskning av högtalarnas dimensioner. Det blir också lättare att välja högtalare med gynnsamma riktningsegenskaper.

Intressant är också den stora skillnaden i taluppfattbarhet mellan de olika långa bärfraserna. Detta bekräftar de slutsatser som drogs efter den första försöksomgången där den stora skillnaden i konsonantbortfall för första och sista konsonanten var en första indikation. Den belyser väl logatommetodens brister i efterklangsrisk miljö, där en kort bärfras inte hinner bygga ett fullständigt maskerande efterklangsfält.

8 LITTERATUR

- [1] Davis, D. & C. (1987): "Sound System Engineering", Howard W. Sams & Co., Indianapolis.
- [2] Hagerman B. (1982): "Sentences for testing speech intelligibility in noise", *Scand. Audiol.* 11, s.79.
- [3] Hagerman B. & Lindblad A-C. (1981): "Modulation transfer function and speech transmission index - a method to predict speech intelligibility in Swedish?", Institutionen för teknisk audiologi, Karolinska Institutet.
- [4] "Instruction manual, Speech transmission meter type 3361 consisting of transmitter type 4225 and receiver type 4419", (1986), Brüel & Kjær.
- [5] Latham, H.G. (1979): "The signal-to-noise ratio for speech intelligibility - an auditorium acoustics design index", *Appl. Acoust.* 12, July, s. 253.
- [6] Liljencrantz, J. (1986): "Elektroakustik", Institutionen för talöverföring och musikakustik, KTH.
- [7] Lochner, J.P.A. & Burger J.F. (1959): "The intelligibility of reinforced speech", *Acustica* 9, s 31.
- [8] Lochner, J.P.A. & Burger J.F. (1961): "The intelligibility of speech under reverberant conditions", *Acustica* 11, s 195.
- [9] Lundin, F. (1986): "Rumsakustiska talöverföringsproblem", Institutionen för talöverföring och musikakustik, KTH.

- [10] Peutz, V.M.A. (1971): "Articulation Loss of Consonants as a Criterion for Speech Transmission in a Room", *J. Audio Eng. Soc.* 19:11, Dec. s. 915.
- [11] Steeneken H.J.M. & Houtgast, T. (1980): "A Physical method for measuring speech-transmission quality", *J. Acoust. Soc. Am.* 67, Nr 1, Jan, s 318.
- [12] Steeneken, H.J.M. & Houtgast, T. (1982): "Some Applications of the Speech Transmission Index (STI) in Auditoria", *Acoustica* 51, s 229.
- [13] Zwicker E. & Feldtkeller R. (1967): "Das Ohr als Nachrichtenempfänger", S. Hirzel Verlag, Stuttgart.

BILAGA 1 Resultat av lyssnartester

FÖRLUST AV INITIALKONSONANTER [%]

	Medel	S.dev
A	2,0	0,6
B	3,8	0,9
C	6,2	1,1
D	6,0	1,1
E	6,0	1,1
F	12,2	1,5
G	12,0	1,5
H	12,6	1,5
I	20,8	1,8
J	18,3	1,7

FÖRLUST AV VOKALER [%]

	Medel	S.dev.
A	2,0	0,6
B	1,2	0,5
C	1,8	0,6
D	1,6	0,6
E	2,4	0,7
F	4,4	0,9
G	4,8	1,0
H	2,6	0,7
I	3,2	0,8
J	3,4	0,8

FÖRLUST AV FINALKONSONANTER [%]

	Medel	S.dev.
A	4,6	0,9
B	3,8	0,9
C	10,0	1,3
D	9,9	1,3
F	8,2	1,2
E	23,0	1,9
G	22,5	1,9
H	21,8	1,8
I	31,6	2,1
J	28,6	2,0

FÖRLUST AV ORD [%]

	Medel	S.dev.
A	8,0	1,2
B	8,2	1,2
C	16,6	1,7
D	16,2	1,6
E	15,2	1,6
F	34,6	2,1
G	34,8	2,1
H	32,6	2,1
I	44,8	2,2
J	42,0	2,2

Kolumnerna representerar följande inställningar

	Filtrering	Direktfält/Efterklangsfält
A	Rak frekvensgång	+3 dB
B	Högpas 1000 Hz	+3 dB
C	Rak frekvensgång	-3 dB
D	Högpas 500 Hz	-3 dB
E	Högpas 1000 Hz	-3 dB
F	Rak frekvensgång	-9 dB
G	Högpas 500 Hz	-9 dB
H	Högpas 1000 Hz	-9 dB
I	Rak frekvensgång	-9 dB Förlängd bärfras
J	Högpas 1000 Hz	-9 dB Förlängd bärfras

BILAGA 2 Resultat av mätningar med RASTI

	RASTI	STI 500	STI 2000	AI _{cons}
A	0,54	0,49	0,58	9 %
B	0,47	0,40	0,53	13 %
C	0,40	0,40	0,41	19 %
D	0,41	0,38	0,42	18 %
E	0,42	0,40	0,43	17 %
F	0,28			36 %
G	0,27			38 %
H	0,31			31 %

Raderna representerar följande inställningar

	Filtrering	Direktfält/Efterklangsfält
A	Rak frekvensgång	+3 dB
B	Högpas 1000 Hz	+3 dB
C	Rak frekvensgång	-3 dB
D	Högpas 500 Hz	-3 dB
E	Högpas 1000 Hz	-3 dB
F	Rak frekvensgång	-9 dB
G	Högpas 500 Hz	-9 dB
H	Högpas 1000 Hz	-9 dB

BILAGA 3 Resultat av ljudkvalitetsundersökning

Fk	Först [dB]	Högt inkoppl		Varifrån			Ljudkvalitet					Medel
		Ja	Nej	1	X	2	5	4	3	2	1	
Rak	0	7	3	3	4		3	2	2			4,1
	6	10		2	5	3	4	4	1	1		4,0
	9	10				10	5	3	2			4,3
	12	10				10	5	2	2	1		4,1
500	0	8	2	6	1	1	3	3	2			4,1
	6	10		2	4	4	3	7				4,3
	9	8	2		4	4	5	2	1			4,2
	12	10				10	2	6	2			3,8
1000	0	4	6	1	2	1	1	2	1			4,0
	6	9	1	3	5	1	2	5		1		3,4
	9	10		1	4	5	2	6	2			4,0
	12	10		2	1	7	3	2	3	2		3,4

Siffrorna anger antal avgivna svar. Kolumnen Medel visar medelvärdet av ljudkvalitetsbetygen.

Undersökningen gjordes som parvis jämförelse mellan oförstärkt och förstärkt talare.

Försökspersonerna ombads besvara följande frågor:

- 1) Hör du att högtalaren vid sidan är inkopplad?

Kryssa för Ja eller Nej.

Om du svarat ja på den första frågan, dvs. du har hört att högtalaren är inkopplad, besvarar du följande två frågor:

- 2) Varifrån tyckte du att ljudet kom?
 - 1) Från "talaren" vid altaret.
 - X) Från en obestämd plats.
 - 2) Från högtalaren vid sidan.

- 3) Tänk dig att högtalaren vid altaret är en verklig talare. Lyssna och jämför hur du tycker att högtalaren vid sidan färgar ljudet. Bry dig inte om, om återgivningen är starkare än talaren vid altaret.
Betygsätt ljudåtergivningen med
 - 5) Utmärkt
 - 4) Bra
 - 3) Acceptabel
 - 2) Dålig
 - 1) Oduglig

BILAGA 4 Matematisk beskrivning av STI

Testsignalen består av brus indelat i sju oktavband med mittfrekvenser från 125 Hz till 8 kHz. Oktavbandens nivåer är anpassade så att signalens spektrum liknar vanligt tal. Bruset amplitudmoduleras med modulationsgrad 1 och fjorton modulationsfrekvenser mellan 0,63 Hz och 12,5 Hz. Oktavbanden och modulationsfrekvenserna bildar tillsammans en matris med 98 mätpunkter. Signalen i oktavband k kan för en given modulationsfrekvens F skrivas som

$$I_k(t) = \bar{I}_k(1 + \cos 2\pi Ft) \quad (1)$$

där I_k är signalens momentana intensitet och \bar{I}_k är signalens medelintensitet.

På mottagarsidan kommer bakgrundsbrus och efterklang att orsaka en minskning av modulationen, dvs. en utslätning av signalen. Den mottagna signalen kan skrivas

$$I_k(t) = \bar{I}_k(1 + m \cos 2\pi Ft) \quad (2)$$

där m är modulationsindex som ges av

$$m = \frac{\bar{I}_{\text{test}}}{\bar{I}_{\text{test}} + \bar{I}_{\text{brus}}} \quad (3)$$

där \bar{I}_{test} är den mottagna testsignalens medelintensitet, och \bar{I}_{brus} är den uppkomna störsignalens medelintensitet innefattande både buller och efterklang.

STI tar också hänsyn till maskeringen från ett lägre frekvensband till ett högre. Ett ekvivalent maskeringsbrus ($\bar{I}_{M,k}$) adderas till bruset i oktavband k . Maskeringsbrusets intensitet sätts till

$$I_{M,k} = 0,0003 \bar{I}_{k-1} \quad (4)$$

eller -35 dB per oktav. \bar{I}_{k-1} är medelintensiteten i oktavband k-1.

Ur modulationsindex kan ett signal-brusförhållande beräknas:

$$S/N = 10 \log \frac{m}{1-m} \quad (5)$$

S/N truneras till $-15\text{dB} < S/N < +15\text{dB}$, och ett medelvärde bildas för alla modulationsfrekvenser till ett S/N-värde för varje oktavband. S/N-medelvärdet normaliseras till ett talöverföringsindex för varje oktavband genom

$$STI_{\text{okt}} = \frac{\overline{S/N} + 15}{30} \quad (6)$$

Till sist vägs de olika STI_{okt} -värdena samman till ett STI. Värdena vikts efter deras betydelse för taluppfattbarheten.

Empiriskt har ett samband mellan AI_{cons} och STI befunnits vara

$$AI_{\text{cons}} = 10^{-2,3STI + 2,2} \quad [\%]. \quad (7)$$

**Denna rapport hänför sig till forskningsanslag 880855-1
från Statens råd för byggnadsforskning till DNV Ingemansson AB,
Stockholm.**

R22: 1990

ISBN 91-540-5174-6

Statens råd för byggnadsforskning, Stockholm

Art.nr: 6801022

**Abonnemangsgrupp:
Z. Konstruktioner och material**

**Distribution:
Svensk Byggtjänst
171 88 Solna**

Cirkapris: 41 kr exkl moms