

# Ljudfysik- Lab 3

## Psykoakustik och ljudeditering

Namn	
Personnummer	
Datum	
Godkänd	

Rev: September 2015 Per Sandström/Peter Andersson

# 1 Introduktion

Avsikten med den här laborationen är att studera en del av de psykoakustiska fenomen som tagits upp under föreläsningar och att få prova på lite olika ljudeffekter.

Börja med att starta Matlab och koppla in hörlurar för lyssning. De flesta psykoakustiska fenomen blir tydligast i hörlurar och ljudeffekterna låter bäst där.

Matlabfunktioner finns att ladda ned från Lisam.

## 2 Några psykoakustiska fenomen

### 2.1 Hur lång måste en ton vara för att man ska uppfatta den som en ton?

Börja med att skapa en sinusvåg på 1 kHz i matlab med en samplingshastighet på 44100 Hz (CD-kvalitet) och en längd på 0.1 s. Den kan spelas upp via ljudkortet med tex kommandot wavplay. Wavplay behöver en signal med amplitud mellan -1 och 1 så kom ihåg att normera signalen om du tex adderar två vågformer. Lyssna i hörlurar

```
fs=44100;
t = [0 : 1/fs : 0.05];
y=sin(2*pi*1000*t);
quiet=zeros(1, 4410);
y1=[quiet, y, quiet];
soundsc(y1, fs)
```

Att vi måste lägga in "quiet" beror på hur datorn spelar upp ljudet. Utan pausen hörs inte korta pulser. Test med lite olika längder på tonen. Hur lång behöver den vara för att man ska uppfatta att det är en ton och inte ett klick?

Svar: \_\_\_\_\_

Är det ungefär samma för alla i gruppen?

Svar: \_\_\_\_\_

### 2.2 Ekon

Eko i dess enklaste form är en reflektion från någon yta som man uppfattar som en distinkt egen signal. Detta förutsätter signalen är tillräckligt fördröjd eftersom örat inte kan särskilja ljud som kommer för tätt. I den här övningen ska vi titta lite grann på ekon och se vid vilken tidsfördröjning örat börjar uppfatta en signal som ett eko. Blockschemat för en ekoeffekt gjord visas i figur 1.

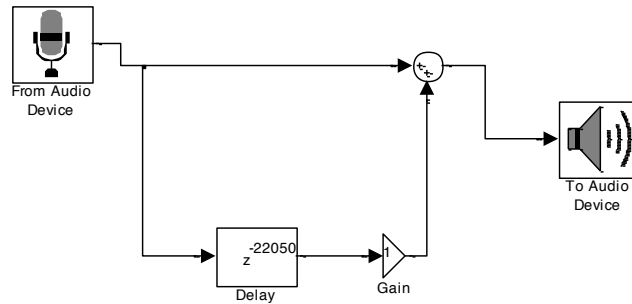


Fig. 1 Enkel eko-effekt i simulink

Differensekvationen för ett enkelt eko blir  $y(n)=x(n)+gx(n-M)$ , där  $x$  är den ursprungliga signalen,  $g$  är förstärkningen och  $M$  antalet sampel man vill fördröja signalen. Använder man transformer kan man skriva överföringsfunktionen som  $H(z)=1+gz^{-M}$ . Exemplet beskriver ett FIR-kamfilter (Finite Impulse Response), dvs impulsresponsen dör ut efter ett tag.

Skapa en sinusvåg med frekvens 1 kHz som är 50ms lång och använd kommandot `audioecho(signal, delaytid, först, fs)` för att testa lite olika ekotider.

```
function [ output] = audioecho( signal, delaytime, gain, fs)
% antal sampel att fördröja
N=round(delaytime*fs);

% skapa en tillräckligt lång utvektor
output=zeros(1,length(signal)*2+N);
% Fyll på signalen i utvektorn
for n=1:length(signal);
    output(n)=signal(n);
end
signal=[signal zeros(1,length(signal)+N)];

% Fördröjningsledningen
for n=N+1:length(output);
    output(n)=signal(n)+gain*signal(n-N);
end

% Normalisera till max 1
output=output/max(abs(output));
end
```

Hur lång behöver delaytiden vara för att ni tydligt ska uppfatta ljudpulserna som två separata?

Svar: \_\_\_\_\_

Det finns även en ljudfil "hej.wav" där examinators säger hej. Ladda in den här filen med kommandot `wavread`, eller spela om möjligt in ett eget ljud. Man behöver transponera den inlästa matrisen för att få den på samma format som använts ovan.

```
y=wavread('hej.wav');  
y=y.';
```

Undersök hur lång delaytid du behöver nu för att uppfatta den fördröjda signalen som ett distinkt eko. Är det samma tid som förut? Vad kan eventuella skillnader bero på?

Svar: \_\_\_\_\_

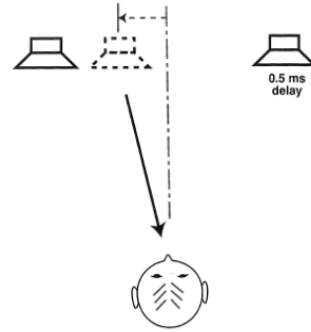
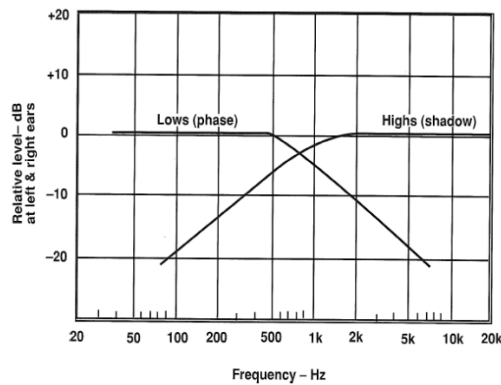
## 2.3 Panorering

När man lyssnar i stereo kan man uppfatta att ljud kommer från olika platser i rummet. Till stor del (men inte enbart) bestämmer hjärnan varifrån ett ljud kommer beroende tidskillnader och intensitetskillnader mellan öronen. Intensitetskillnaden fungerar bäst vid höga frekvenser eftersom att vid långa våglängder/låga frekvenser gör diffraktion att intensiteten blir i stort sett samma i båda öronen. Fasskillnaden fungerar bäst vid låga frekvenser, eftersom korta våglängder gör att en fasskillnad inte entydigt avgör skillnaden i gångväg (samma fasskillnad kan uppnås om vi adderar ett helt antal våglängder till ena vågen). Lokalisering med fasskillnad kallas för Haas-effekten. Figur 2 beskriver i stora drag hur det fungerar.

I mixerbord finns det rattar där de står "pan" på. Dessa sköter panorering genom att ändra intensiteten mellan de två stereokanalerna.

På datorn finns en fil med en kort snutt av en sång av en sång omgjord till mono. Matlab-programmet `panning.m` gör om filen till stereo och varierar intensiteten mellan de två hörlurarna. Studera gärna programmets uppbyggnad. Provkör och lyssna på filen i något annat program än matlab. Spelar man upp en lång wav-fil i matlab går det inte stänga av under uppspelningen. Jämför med originalet.

```
panning('x1.wav', 'newfile.wav')
```



2 a.

2 b.

Fig. 2 a. Lokalisering av ljudkällor med hjälp av faskillnader och intensitetsskillnader b. Haaseffekten. (Båda bilderna från: Stericher och Everest: The new Stereo Soundbook, AEA 1998)

Nu ska vi testa faspanorering. Funktionen phasepanning gör detta på ett simpelt sätt genom tidsfördröja den ena kanalen i förhållande till den andra genom att stoppa in 0:or i början av den ena eller andra stereovektorn. Titta gärna på filens uppbyggnad. Utgå från samma ljudfil som ovan och testa med +1 ms och -1 ms och 0 ms.

phasepanning(x1.wav', 'newfile.wav', -1) den sista siffran anger tid i ms. Testa med olika tider i intervallet 0,3 - 30 ms.

Hur förändras positionen på ljudet?

Svar: \_\_\_\_\_

Förutom positionen ändras karaktären på ljudet. Försök beskriv hur du upplever att ljudet förändras.

Svar: \_\_\_\_\_

Varför tror du att man i inspelningsammanhang oftast använder metoden baserad på intensitetsskillnader för att panorera en signal?

Svar: \_\_\_\_\_

Förläng succesivt delaytiden. Hur lång behöver delaytiden vara för att ni tydligt ska uppfatta den fördröjda signalen som ett distinkt eko utan att den sammansmälter med ursprungsljudet?

Svar: \_\_\_\_\_

## 2.4 Maskering

Maskering innebär att ett ljud blir svårare att uppfatta på grund av störande ljud. Om ljuden förekommer samtidigt kallas det simultan maskering och om de är separerade i tid och inte förekommer samtidigt kallas det temporal maskering.

Ett exempel på simultan maskering är om två toner ligger nära varandra i frekvens, då döljer den starkare den svagare.

Skapa ett några sekunder långt ljud med två frekvenser 375Hz och 400Hz sänk den högre tonen en faktor 10 i amplitud (=20dB)

```
t=[0:1/fs:4];  
y4=sin(2*pi*375*t)+0.1*sin(2*pi*400*t);  
soundsc(y4,fs);
```

Kan du särskilja de båda frekvenserna?

Svar: \_\_\_\_\_

Testa att ändra den ena frekvensen. Hur stor måste frekvensskillnaden vara för att man ska höra två distinkta toner?

Svar: \_\_\_\_\_

Är det någon skillnad på hur mycket maskering som sker om den svagare signalen har högre eller lägre frekvens än den starkare?

Svar: \_\_\_\_\_

Figur 3 visar grovt hur maskering fungerar. En stark ton gör att hörbarhetströskeln för toner med frekvenser i närheten ökar kraftigt.

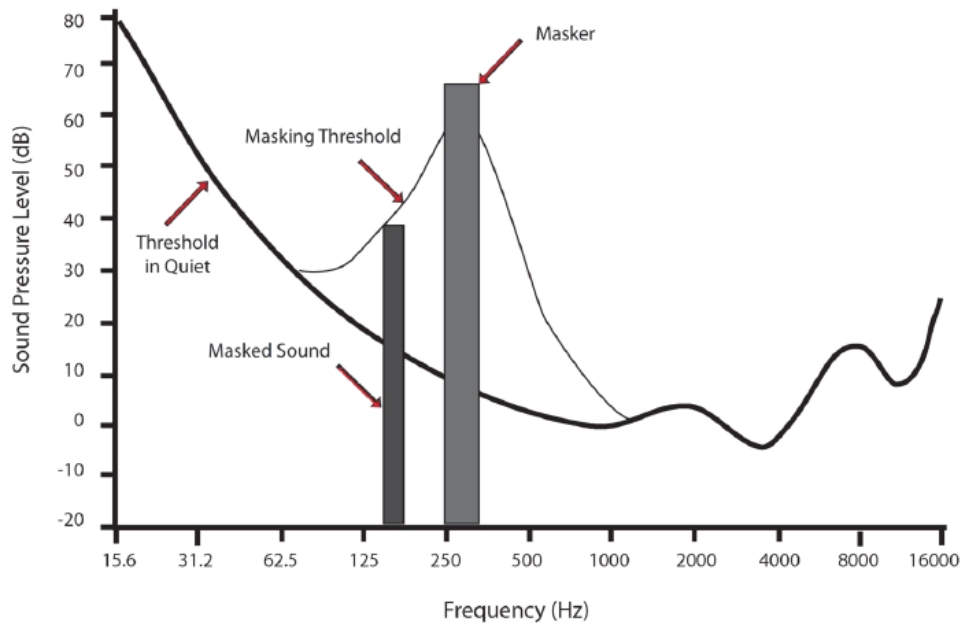


Fig. 3 Maskning (källa Wikimedia commons)

Nu ska vi titta vidare på temporal maskning. Det här experimentet är lite svårt att få bra resultat på om man vet att det finns en ton man ska lyssna på efter bruset, men vi gör ett försök. Vet man inte att ljudet finns som ofta är fallet i tal och musik uppfattar man det inte om pausen och ljudet är kort.

Skapa en brussignal på 1 s. och en tyst vektor på 0.12 s följt av en ljudpuls på 2 kHz/20ms

Slumpmässigt brus i en sekund:

```
noise=2*rand(1,44100)-1;
```

Skapa ”tyst” vektor 0.12 s

```
quiet=zeros(1, round(0.12*44100));
```

Generera en sinusvåg på 2 kHz i 15ms, 5% av max brus

```
t=[0:1/fs:0.015];
```

```
y=0.01*sin(2*pi*1000*t);
```

Skapa sekvensen och spela upp.

```
y5=[noise quiet y quiet];
```

```
soundsc(y5, fs);
```



Justera amplituden på sinustonen så att den är knappt hörbar. Minska succesivt tiden på tystnaden och lyssna om tonen blir svårare att uppfatta.

Vid hur kort uppehåll försvinner tonpulsen för örat?

Svar: \_\_\_\_\_

Ett intressant fenomen är att efterföljande starka ljud kan maskera korta/svaga toner som kommer före det i tid. Testa att byta ordning så att tonpulsen kommer före en kort paus följt av bruset.

```
y5=[y quiet noise];  
soundsc(y5, fs)
```

Börja med en paustid på ca 150ms och minska sedan volymen på tonpulsen tills den nätt och jämt hörs. Minska sedan paustiden succesivt. Kan du hitta en tid där tonen börjar maskeras?

Svar: \_\_\_\_\_

Figur 4 beskriver hur maskering kan ske både före och efter ett störande ljud. Maskeringen är effektivare efter det störande ljudet.

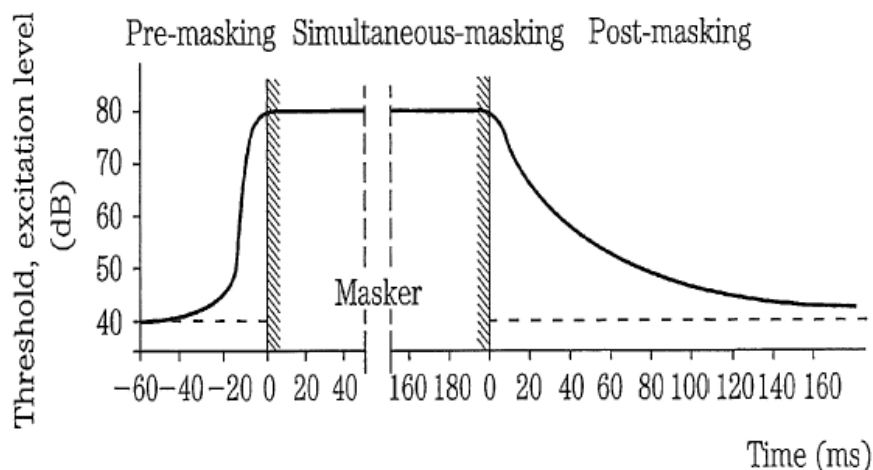


Fig. 4 Temporal maskning (Källa: Pohlmann, Principles of Digital Audio, McGraw-Hill 2011)

### 3 Några ljudsyntesexperiment

Analoga syntar använder sig av det man kallar ljudsyntes. Man utgår från några enkla vågformer och modifierar dem så de låter annorlunda. Vid ljudsyntes utgår man ofta från sinus-, triangel-, sågtands, och fyrkantspulser. Med analoga syntar var detta nästan det enda sättet men även idag förekommer det ofta att man gör på det här sättet för att skapa ljud även om man arbetar digitalt. I det här avsnittet får du prova på några vanliga metoder.

Med hjälp av kommandot "waveforms" kan du starta ett matlab-program med en del ljudsyntesmöjligheter, liknande de som används i analoga syntar. Starta det och testa lite olika parametrar. Låt ljudet spela klart mellan knapptryckningar och ändringar av reglage. Det verkar som om programmet ibland får Matlab att sluta fungera om man trycker för mycket.

Testa att spela sinus-, triangel- och fyrkantsvågor. Hur ändras övertonsinnehållet i en triangelvåg när man ändrar så att den blir mer sågtandslik? Hur ändras ljudkaraktären?

Svar: \_\_\_\_\_

Distorsionsalgoritmen överstyr vågformen. Motsvarigheten i ett analogt sammanhang är att man förstärker vågformen så att förstärkaren klipper, den kan inte ge ut högre signal än ett maximalt värde som oftast bestäms av matningsspänningen hos förstärkaren. Ju mer överstyrning desto mer fyrkantlik blir vågformen. Algoritmen i programmet implementerar en form av mjuk klippning för att minska ner diskantinnehållet i ljudet. Skruvar du upp volymen på förstärkaren där hemma så att den klipper riskerar man sina högtalarelement, särskilt diskanten. Distorsion används dock mycket i modern musik.

Testa hur det låter. Använd en ren sinusvåg och variera graden av distorsion.

Om man ringmodulerar en signal får man två frekvenskomponenter, dels skillnaden och dels summan av de två modulationsfrekvenserna. Ringmodulation är ett specialfall av det som kallas amplitudmodulation.

Vid amplitudmodulering låter man en sinusvågs amplitud (bärvågen) styras av en annan sinusvåg.  $u(t) = \hat{U} \sin \omega_1 t \cdot \sin \omega_2 t$ . Den här tekniken användes förr för att sända radio. Den används fortfarande om du till exempel lyssnar på AM-

bandet på din radio. AM står för just "Amplitude Modulation". Med amplitudmodulering fås förutom skillnads- och summa frekvenser även bärvågens frekvens.

Använd ringmodulation och ändra frekvenserna. Studera frekvensinnehållet. Hur låter det om man ställer frekvenserna ca 200 Hz isär (exempelvis 400 och 600 Hz)?

Svar: \_\_\_\_\_

Genom att fylla i kryssrutorna längst ned till vänster kan man mixa sin signal med olika vågformer, dvs signalerna summeras. Den längst till vänster mixar med en sinus, den i mitten med en triangelvåg och den till höger med en fyrkantsvåg. Testa!

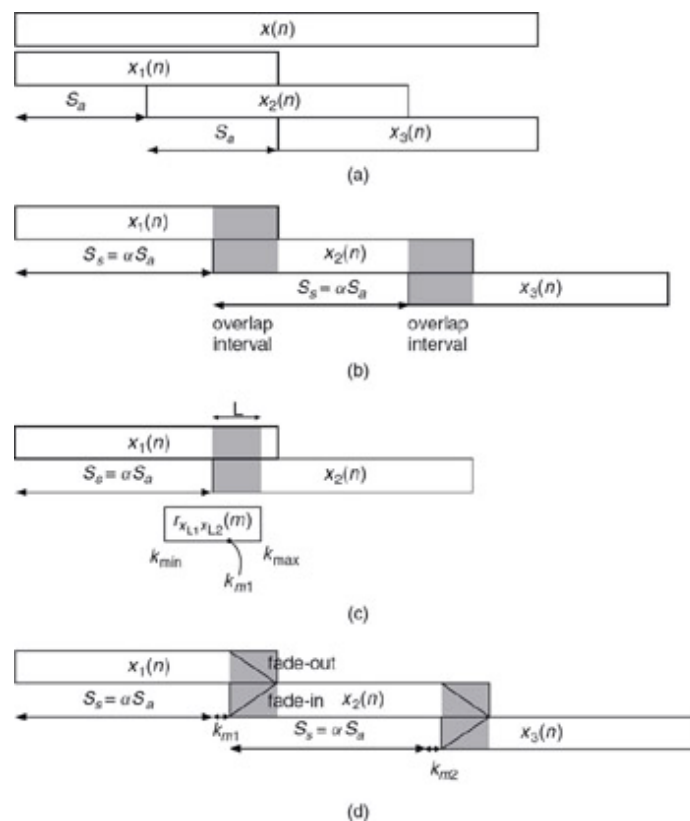
## 4 Några andra digitala ljudeffekter

Här får du prova på några andra typer av ljudmodifikationer. De flesta är hämtade från boken "DAFX - Digital Audio Effects (Issue 2)" av Udo Zölzer. Boken finns på biblioteket som e-bok och innehåller mycket om digitala ljudeffekter. På DAFX.de kan man ladda ned matlabfiler för bokens exempel om du är intresserad av att fördjupa dig.

### 4.1 Time stretch

Ändrar man hastigheten på en inspelning genom att spela upp stycket snabbare eller långsammare ändrar man även tonhöjden.

En metod att ändra längden på ett musikstycke utan att ändra frekvensinnehållet, pitch, är att använda så kallade "Synchronous Overlap and Add", eller SOLA metoden. I denna delas musikstycket in i segment. Ska stycket förlängas, repeteras vissa segment. Ska det kortas ned, tar man bort vissa segment. Se figur 5 nedan. Ett problem som uppstår är diskontinuiteter i amplitud och fas vid skarvarna. Problemet kan minskas genom att man låter segmenten "fada" in och ut ur varandra.



Figur 5. Tidstörning, källa "DAFX - Digital Audio Effects (Issue 2)" av Udo Zölzer.

Testa genom att köra programmet "TimeScaleSOLA.m" variera parametern ,alpha, som bestämmer hur mycket ljudet töjs.

## 4.2 Phase Vocoder

Ett kraftfullt verktyg för ljudeffekter är det som kallas Phase Vocoder. Vocoder står för voice coder. En vocoder gör om ljud till en kod som man kan processa. Detta låter förstås rätt allmängiltigt men är en metod som kan användas till mycket. Det finns flera olika varianter. Tex använde syntpionjärerna Kraftwerk vocodrar för många av sina ljudeffekter. I de första Star Wars filmerna gjordes Darth Vaders röst med hjälp av en vocoder. Man kan göra tonlägesändringar, roboteffekter och viskningseffekter och mycket annat.

En teknik att göra detta är att använda det som kallas "Short Term - Fast Fourier Transform" (ST-FFT) som går ut på att man delar upp sin ljudfil i flera småbitar och gör analysen på var bit för sig. Modifikation av signalen görs på varje bit i frekvensplanet och följs sedan av en invers fouriertransform som rekonstruerar signalen i tidsplanet. Se figur 9.

Här ska vi prova på några exempel som använder en "Phase Vocoder". Vi går inte in på några detaljer. Det ligger utanför den här kursen mål att programmera dessa effekter om det inte sker i projekten på eget bevåg.

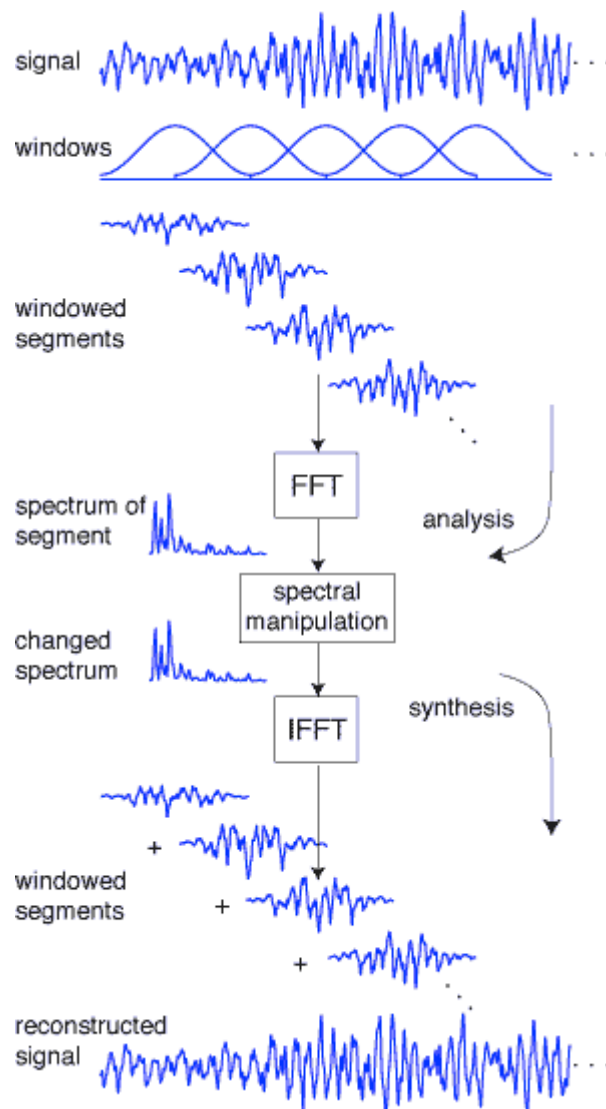


Fig. 6 Principen för en Phase Vocoder byggd på ST-FFT (Källa: <http://sethares.engr.wisc.edu/vocoders/phasevocoder.html>)

Tidstörning med "phase vocoder": En klassiskt sätt att göra detta är att hålla amplituden i signalen konstant och modifiera fasan så att alla så att alla frekvenser finns kvar i varje ögonblick. Som exempel kan man tänka sig att man spelar två toner på 1s, 1000 och 2000Hz. Tittar man på FFT:n av dessa toner ser man att innehåller då energi vid 1000Hz i 1s och vid 2000Hz i 1s. Vill man förlänga så att varje ton varar i två sekunder, kan man då i frekvensplanet se till att man har energi vid 1000Hz i 2s och vid 2000Hz i 2s. Dvs vi dubblar antalet frekvensblock.

Testa programmet, VX\_tstretch\_bank. Prova att ändra time-stretch faktorn. Låter det bättre eller sämre än SOLA metoden?

Pitch-ändring utan att ändra hastighet: Pitch-ändring innebär att man skiftar varje frekvens med en faktor. Detta är i princip enkelt att göra om man känner till frekvensinnehållet i signalen.

Testa `VX_pitch_bank` och ändra faktorn som ändra hur mycket frekvensen ändras.

Robotisering av ljud: Här sätter man fasen till 0 i varje FFT-block innan rekonstruktion. Effekten blir en constant pitch på ljudet. Det tvingar ljudet att bli periodisk så många slumpmässiga variationer omvandlas till en robot-röst. Provkör `VX_robot`.

"Whisperization": Om man gör fasen i frekvens-tidsrepresentationen slumpmässigt får man olika beteenden beroende på fönsterlängden för `ST_FFT` beräkningarna. Gör man fönstret kort kommer man att få en viskningseffekt. Provkör `VX_whisper`. Beroende på fönstrets längd `"s_win"` får man ganska olika effekt på ljudet. Testa!

## 5 Ljudmixning

Denna övning görs i MuLab som finns installerat på de bärbara datorerna. Vi skall göra en mycket enkel mix av ett förmixat musikspår och ett nästan oprocesserat sångspår. Det finns några installerade effektplugins som vi kan använda.



Öppna MuLab och filen som heter lab3-mixning.

Börja med att ställa in en någorlunda bra balans mellan musikspåret och sången genom att dra i volymreglarna som finns i den nedre delen av programmet.

### Komprimering:

Sången är mer dynamisk än musikspåret och för att alla ord skall höras tydligare provar vi med att minska dynamiken med komprimering. Lägg till en kompressor (GTA\_revC) genom att klicka på en triangel till vänster om sångregeln och välj Effects-Vst-GRA\_revC.

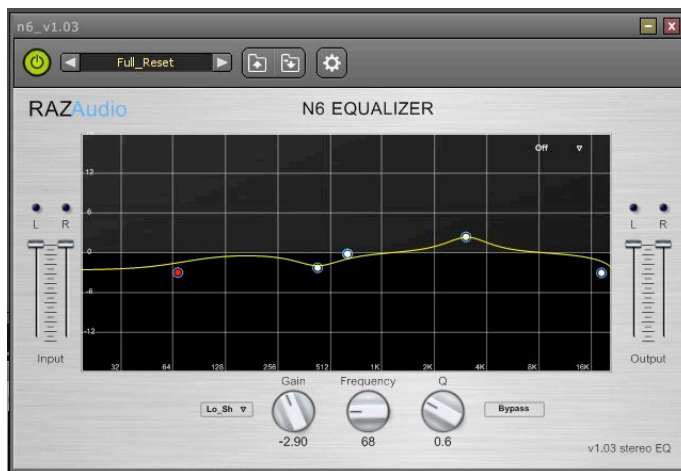




Prova lite olika parametervärden och försök få sången jämnare så att alla ord hörs tydligt. Threshold anger tröskelnivån när kompressorn börjar arbeta. Ratio anger hur mycket som ljudet skall komprimeras när det är starkare än tröskelnivån. Tidsparametrarna attack och release bestämmer hur snabbt kompressorn börjar respektive sluta arbeta. Eftersom kompressorn minskar de starkare ljuden minskar den genomsnittliga nivån, för att kompensera för detta och göra en jämförelse med det okomprimerade ljudet tydligare finns out gain, vilken höjer upp hela nivån på signalen. En kompressor ställs vanligen in för att processera allt ljud, vilket fås om vi ställer in reglaget wet på 100%. Det finns även en mätare som visar hur mycket komprimering som sker i dB. Jämför före och efter processering genom att stänga av och på effekten med den gröna runda symbolen.

### Filtrering:

Prova hur en equalizer kan användas för att få sången tydligare. I det här fallet är den redan inspelad med ett högpasfilter, men det kanske finns en aning grumlig i mellanregistret. Använd den parametriska equalisern N6. Testa om du kan få sången att framträda tydligt genom att skära bort lite grumlighet i mellanregistret vid ca 400hz och tillför lite tydlighet i taluppfattningsområdet ca 2-4kHz med ett par peakfilter och lite skär eventuellt lite av den lägsta basen och översta diskanten med ett shelvingfilter. Överdriv gärna först för att få en känsla för hur de olika frekvensområdena låter. Jämför även här före och efter processering



### Delay:

Testa att skapa ett "slapdelay" dvs ett delay som kan uppfattas som ett eko, om det mixas svagt in uppfattar vi att sången ligger lite längre bak i mixen. Prova med en delaytid i området ca 100-300 ms med maDelay. För att se tiden uttryckt i ms knappen vid notsymbolen tryckas in, annars väljs delaytiden enligt de notvärden som projektets tempo motsvarar. till maDelay simulerar ett bandeko som även har "flutter", dvs en något varierande bandhastighet vilket gör att tiden och pitchen varierar något. Ett delay skall mixas tillsammans med det oprocesserade ljudet, därför skall mixreglaget i detta fall ställas in på ett lägre

värde, starta kanske vid kl 9 på reglaget och justera så att lagom mycket delay hörs. Med feedbackreglaget kan multipla ekon skapas. Det finns även möjlighet att filtrera ekot med ett bandpassfilter och distordera ekot med kontrollen drive.



### Reverb:

För att skapa ytterligare rums känsla på sången och få den att smälta in i musiken kan vi lägga på ett reverb. Prova med REV-B och försök hitta lämpliga inställningar. Reverbet har en eq som typiskt används för att minska höga och låga frekvenser samt en dämpparameter som gör reverbet något mjukare och diffusare. Även här vill vi att oprocesserade ljudet skall blandas med det processerade så låt dry vara på 100% och ställ in wet så att lagom mycket effekt hörs.



Gå gärna tillbaks och justera de olika effekterna och ställ in sångnivån så den låter lagom starkt i mixen. Stäng nu av alla effekter och jämför din mix med hur det lät innan du startade!