

Εργασία 4^η

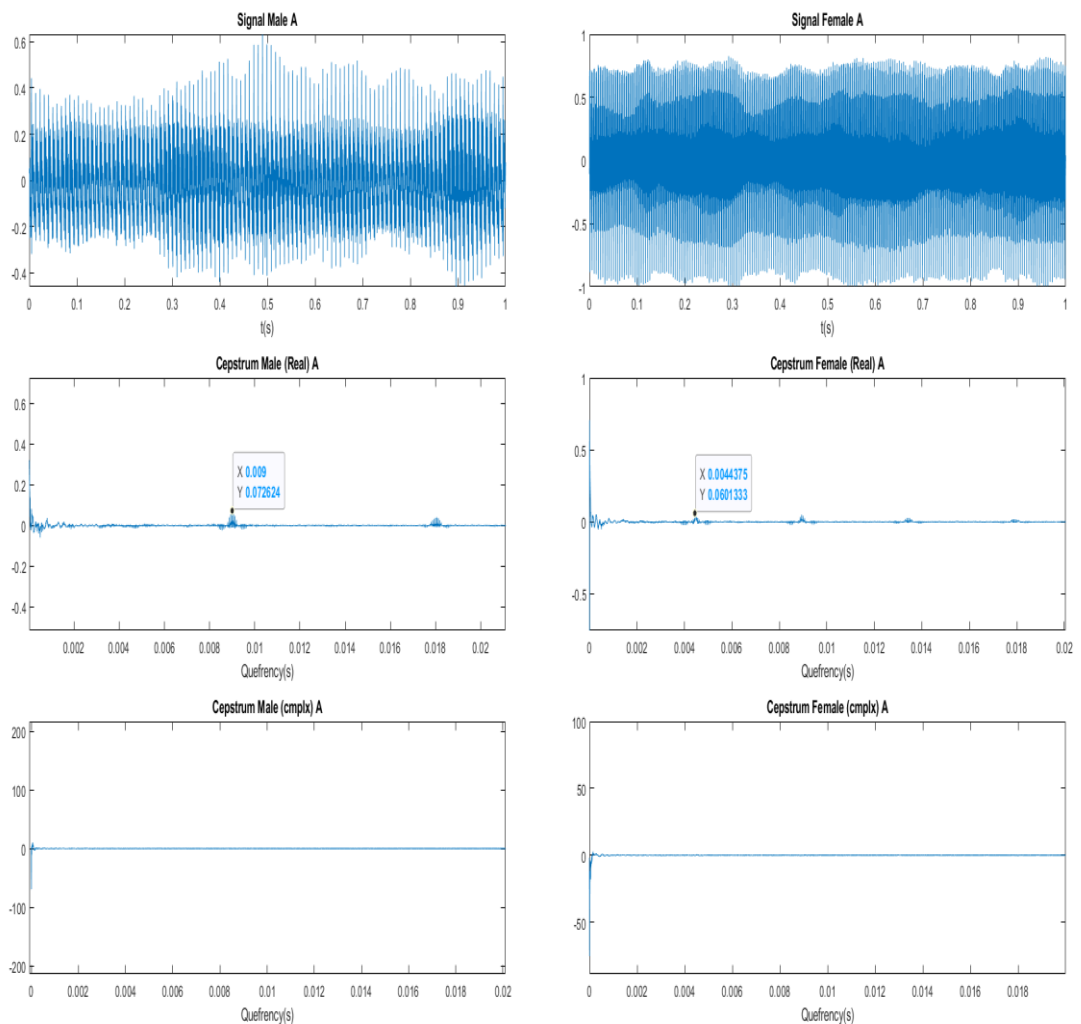
Γρηγορίου Στέργιος 9564

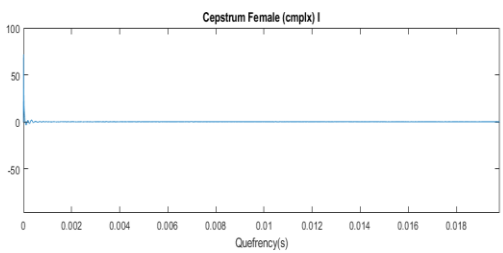
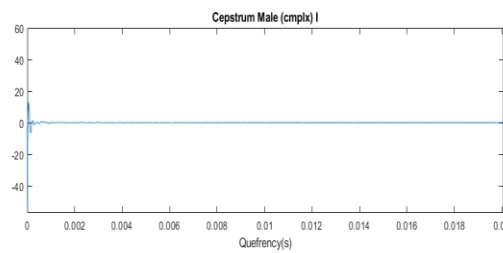
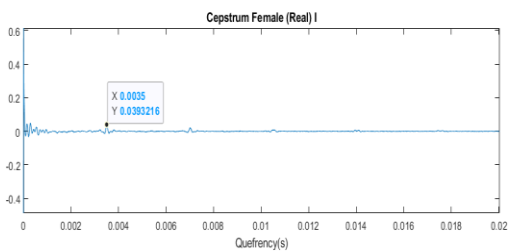
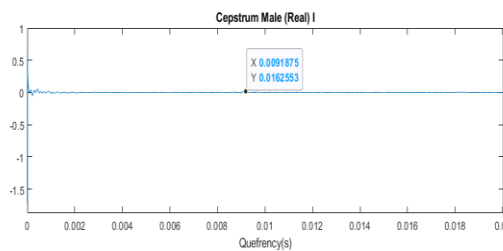
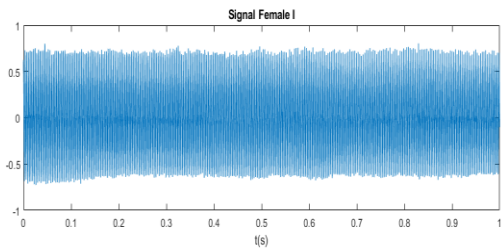
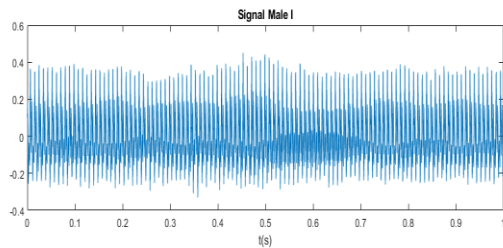
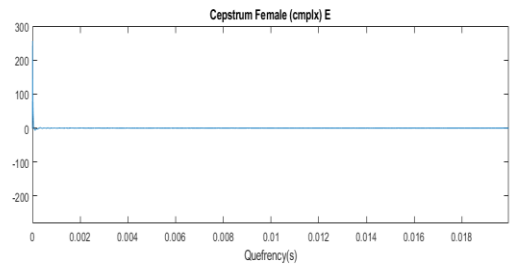
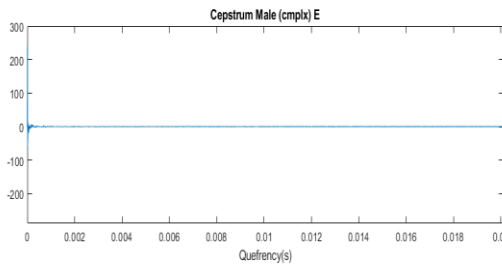
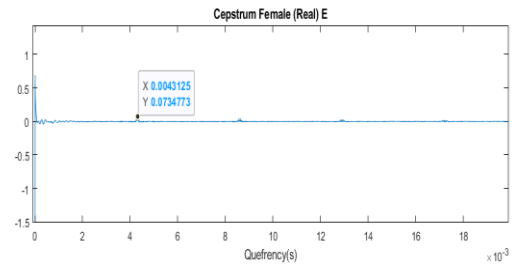
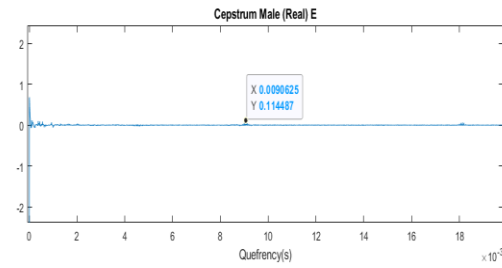
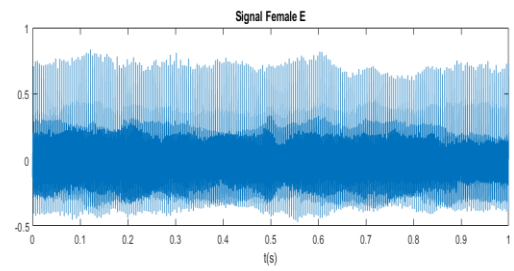
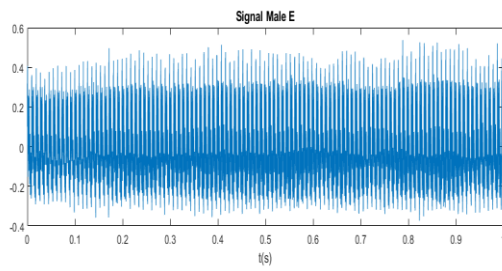
(Ο κώδικας παρατίθεται στο τέλος)

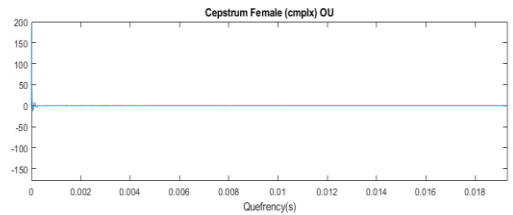
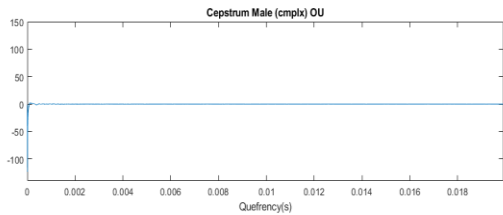
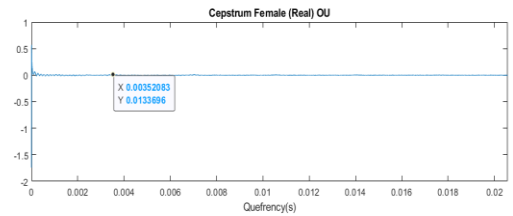
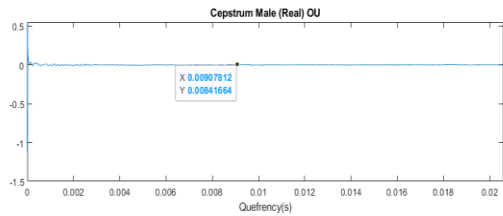
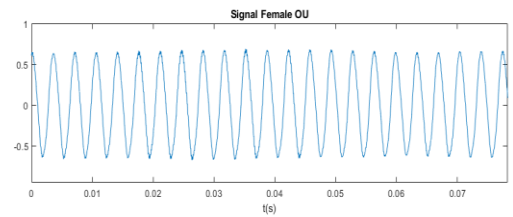
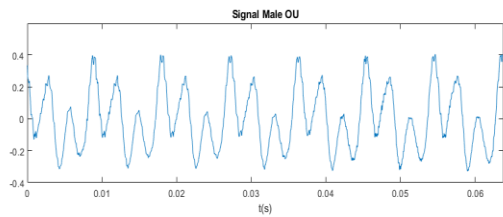
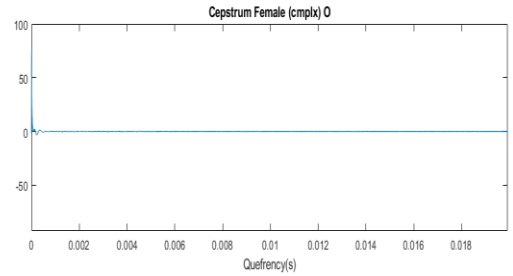
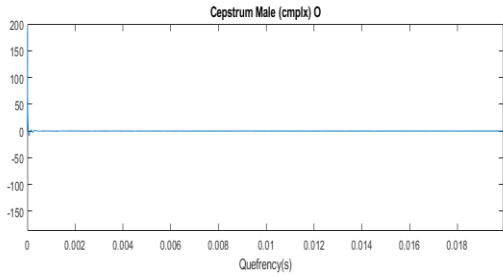
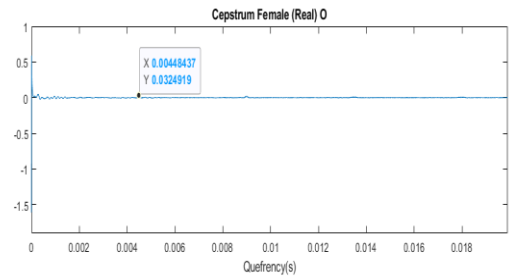
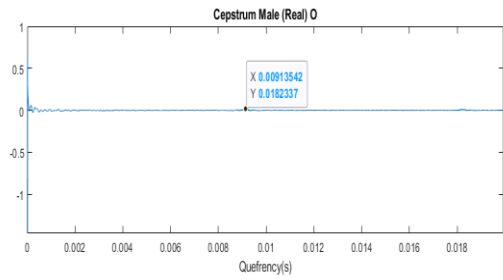
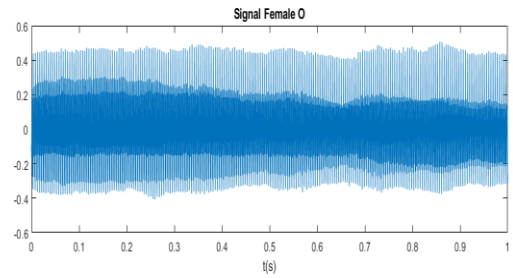
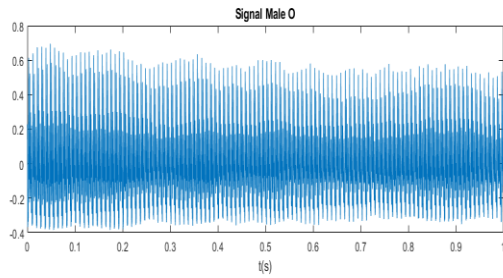
Ερώτημα 1^ο

Η ηχογράφηση έγινε στο matlab με χρήση των εντολών `audiorecorder`, `recordblock`. Έγινε χρήση μικροφώνου Realtek με πραγματική συχνότητα δειγματοληψίας 192KHz. Τα αρχικά σήματα 10 δευτερολέπτων κόπηκαν σε μικρότερα του ενός για γρηγορότερη ανάλυση. Προσπάθησα να κρατήσω διαστήματα με όσο το δυνατόν πιο σταθερή ένταση και `pitch`. Τα ανδρικά σήματα είναι επίσης όλα στο ίδιο `pitch` πάνω κάτω, ενώ τα γυναικεία ποικίλουν κι έχουν μεγαλύτερο πλάτος. Οι ιδιομορφίες αυτές όπως αποδείχθηκε δεν επηρέασαν την ανάλυση που ακολουθεί.

Ερώτημα 2^ο





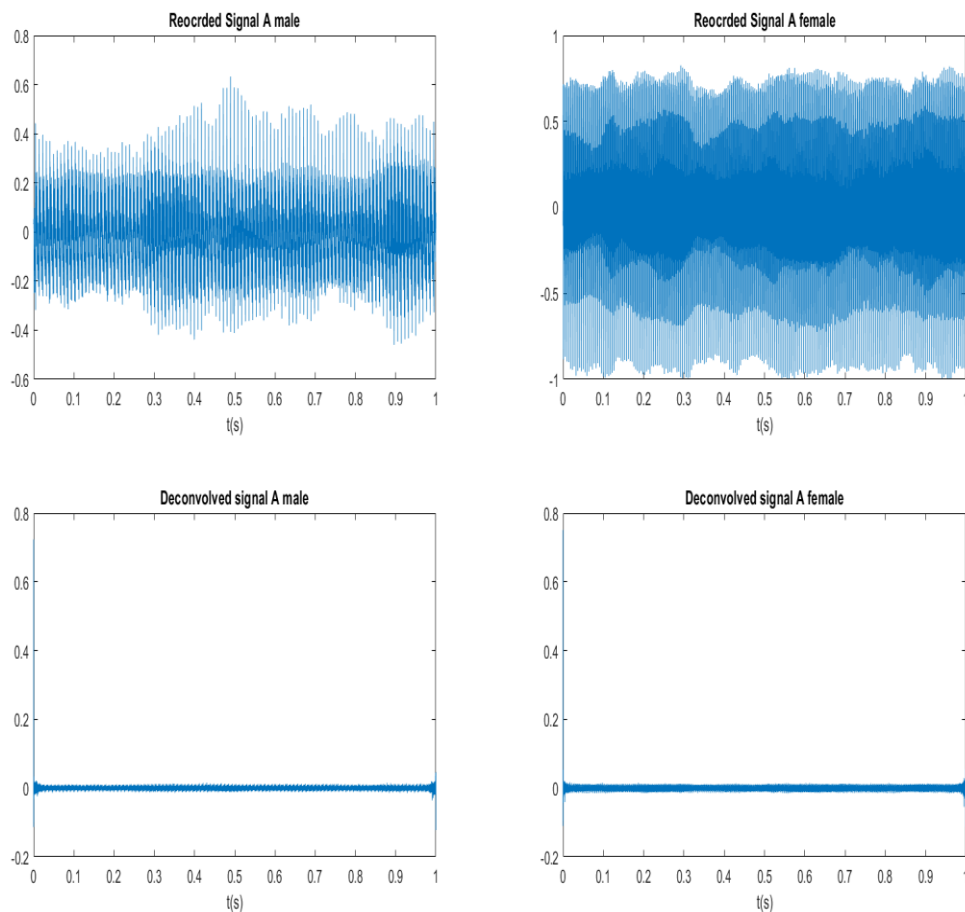


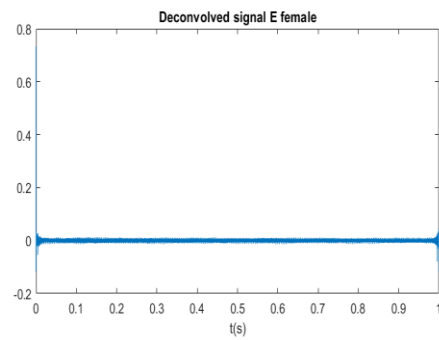
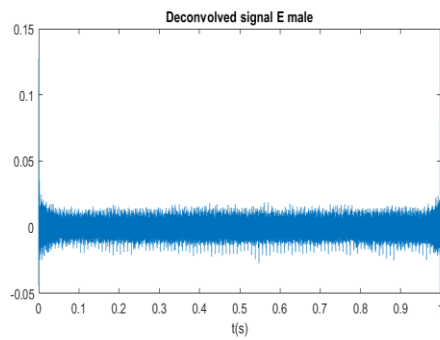
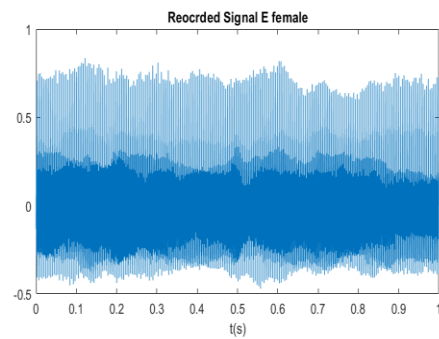
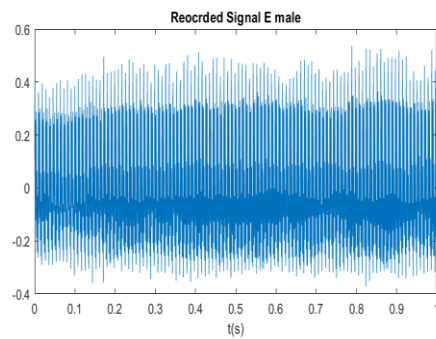
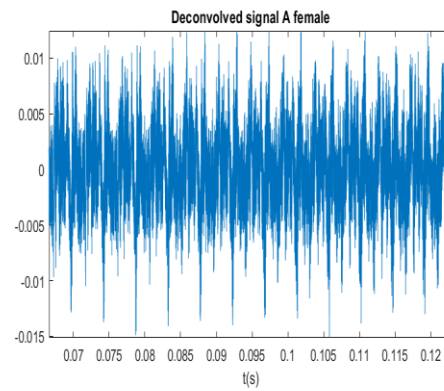
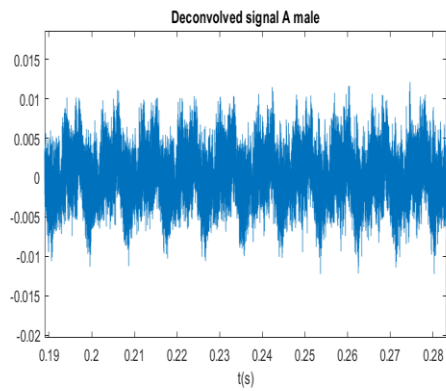
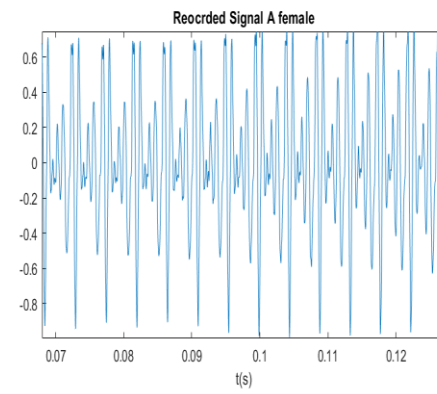
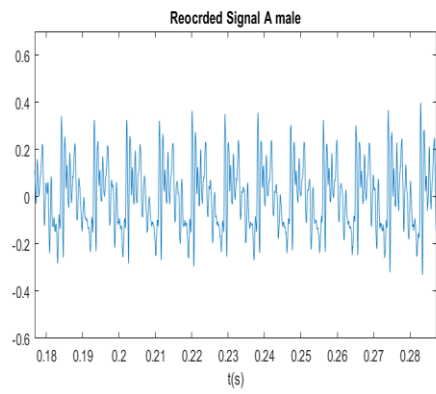
Υπολόγισα τα ζητούμενα και αποθήκευσα τα figures, ενώ τα παραθέτω σαν εικόνες εδώ. Τα παραπάνω διαγράμματα είναι στην πρώτη σειρά τα ηχογραφημένα σήματα, στην δεύτερη το πραγματικό τους cepstrum και στην τρίτη το μιγαδικό. Η πρώτη στήλη είναι για τα ανδρικά ενώ η δεύτερη για τα γυναικεία. Οι ήχοι είναι με τη σειρά (A,E,I,O,OY).

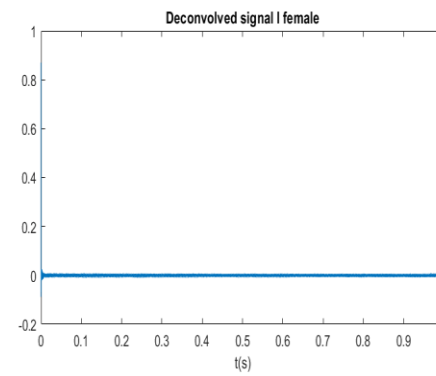
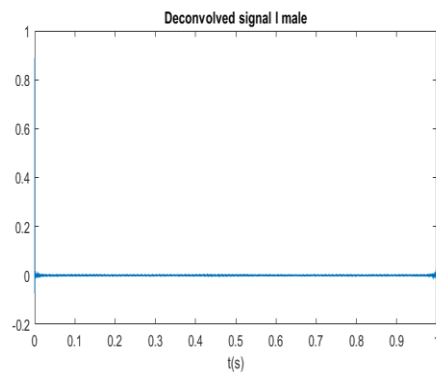
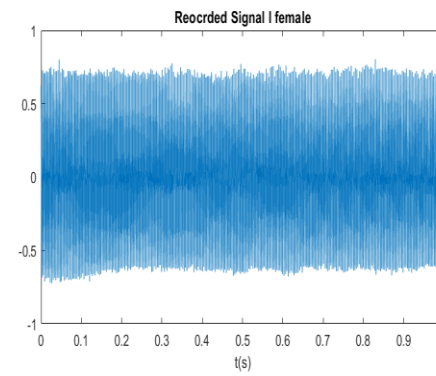
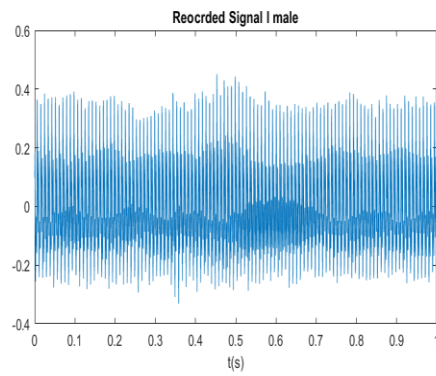
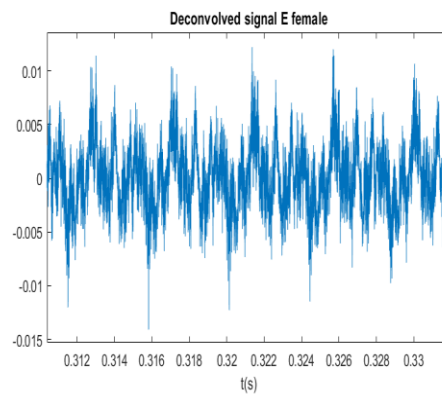
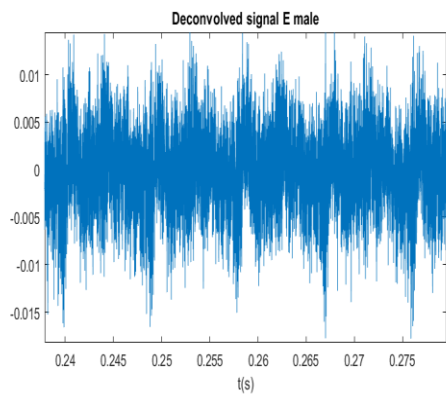
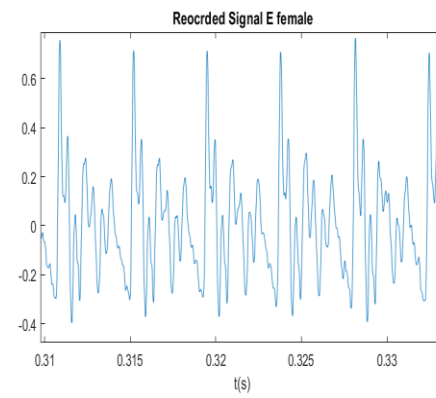
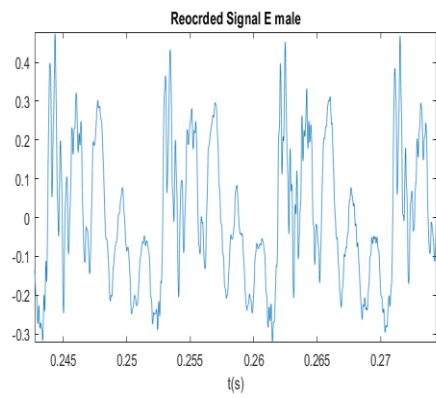
Αυτό που παρατήρησα είναι πως αρχικά τα γυναικεία σμάφата (αναγραμματισμός του φάσματα) έχουν πιο πολλές και πιο έντονες κορυφές ειδικά στο πραγματικό cepstrum (το έντονες). Αυτό οφείλεται στο γεγονός πως κατά κανόνα είναι πιο υψίσυχνα και οι κορυφές αυτές επαναλαμβάνονται ανά περίοδο του σήματος. Άρα μικρότερη περίοδος περισσότερες κορυφές.

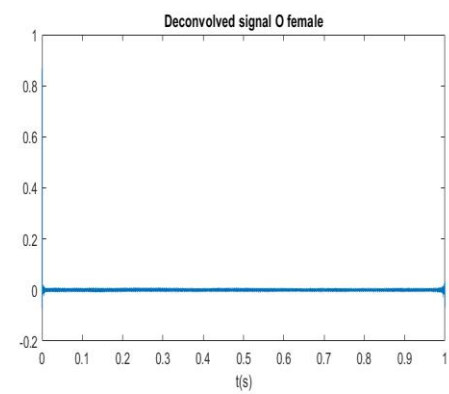
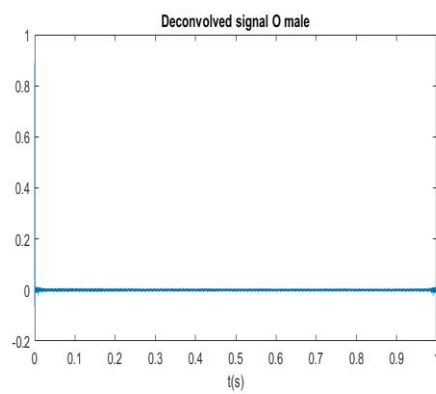
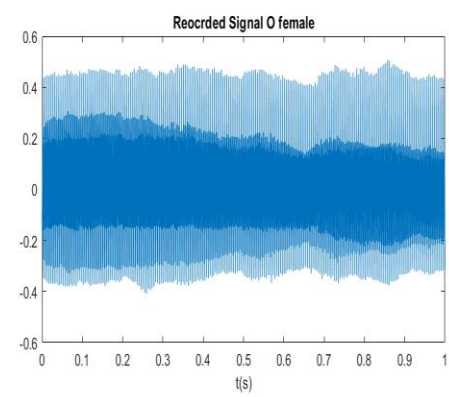
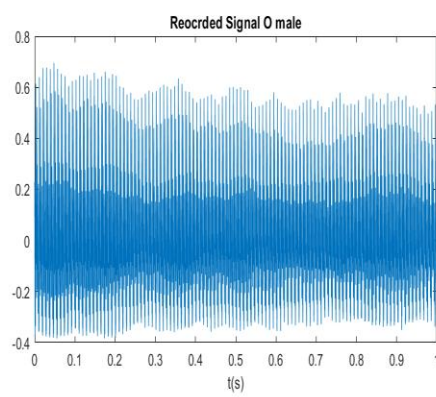
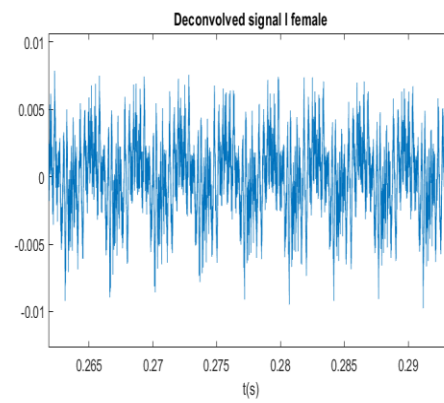
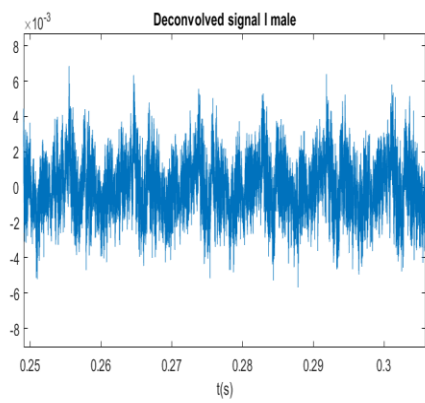
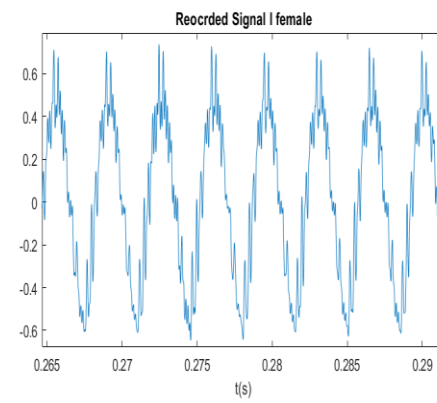
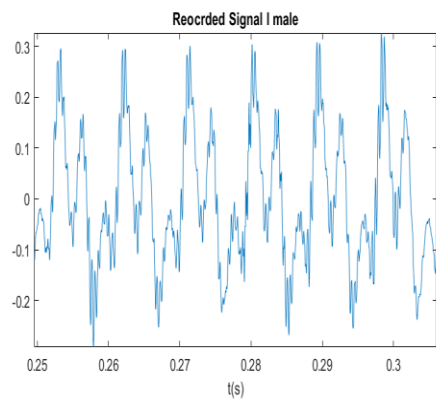
Όσον αφορά το ποιοτικό κομμάτι, τα γυναικεία φαίνεται να έχουν πιο πολλή πληροφορία και στις χαμηλές συχνότητες, πιο μαζεμένη δηλαδή την κρουστική απόκριση. Σε αντίθεση με τα αντρικά που συνήθως είναι πιο απλωμένη. Διαφορά παρατηρείται και στο ίδιο κομμάτι κυρίως στην μορφή που έχει για κάθε διαφορετικό ήχο (είναι κάτι που θα περίμενα). Χαρακτηριστικό παράδειγμα το OY.

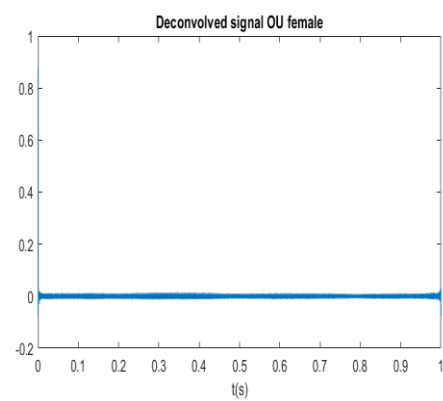
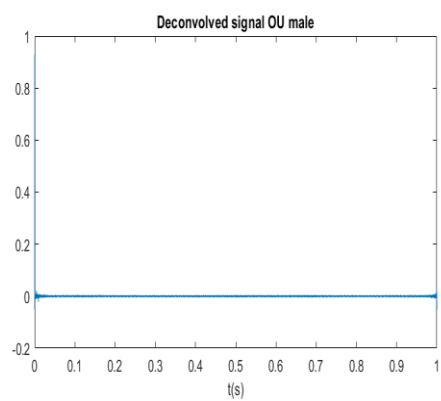
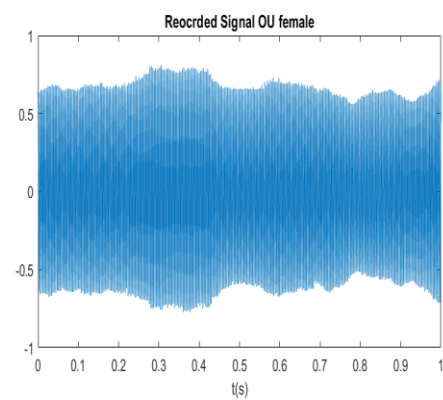
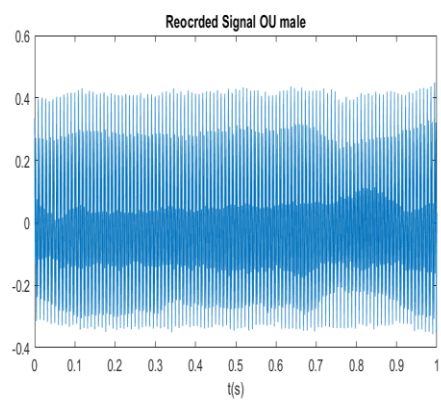
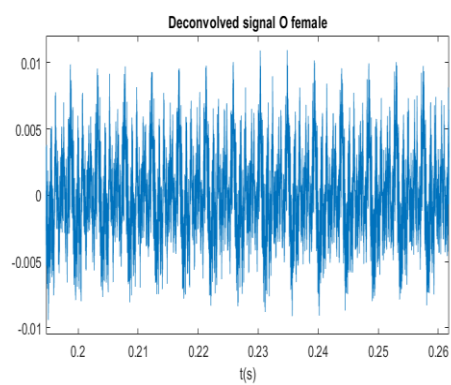
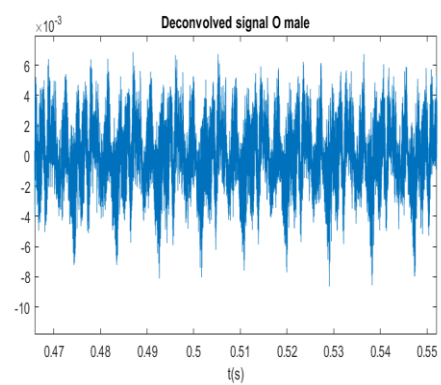
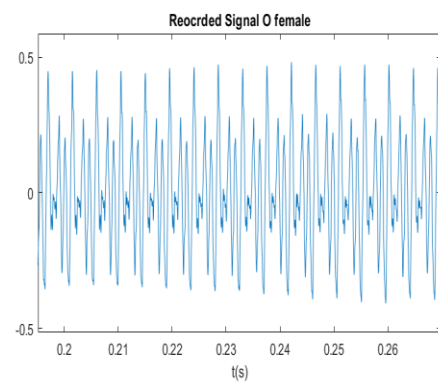
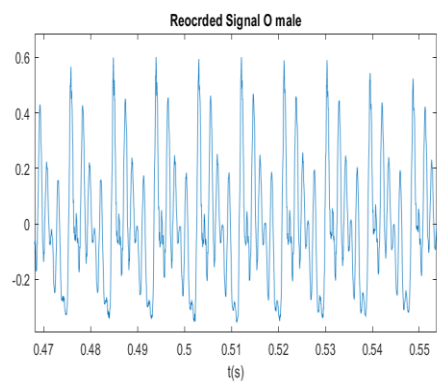
Ερώτημα 3°

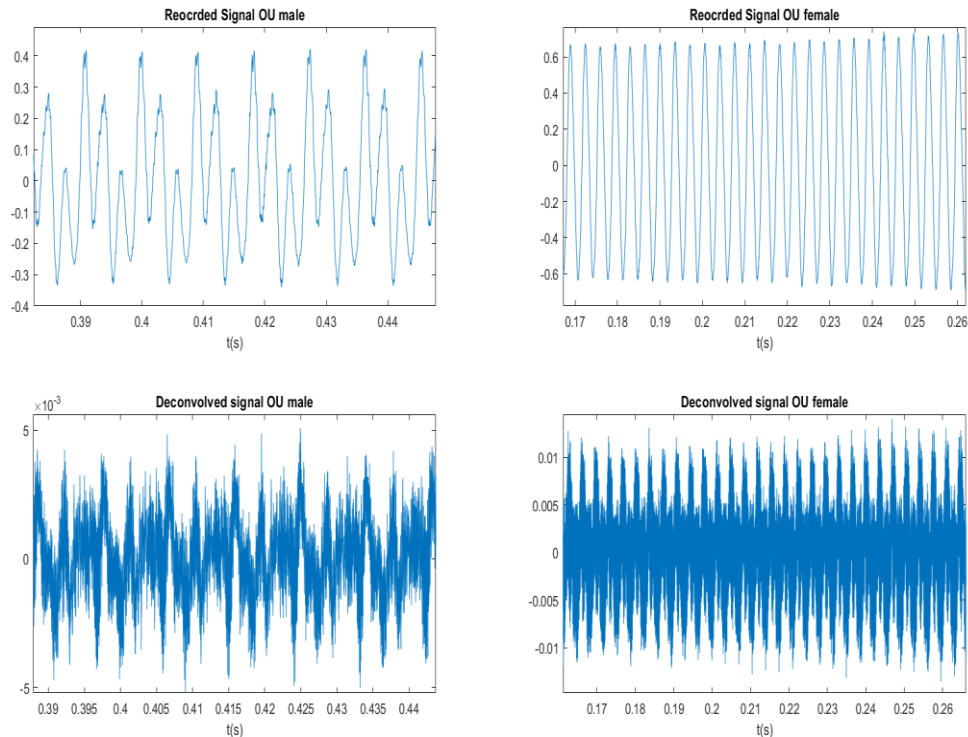








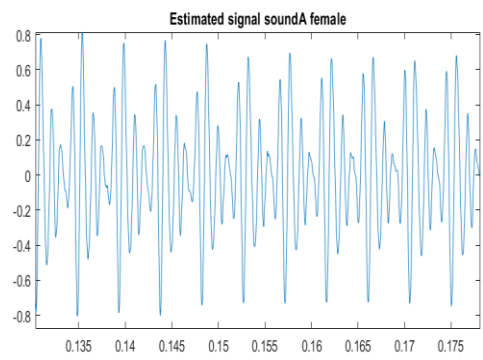
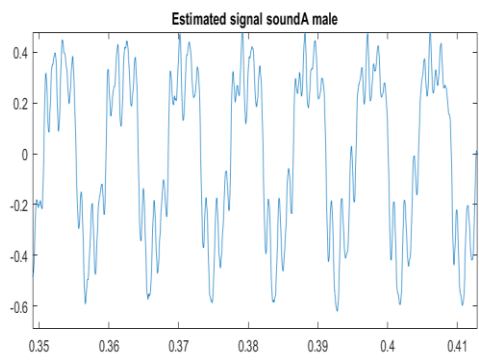
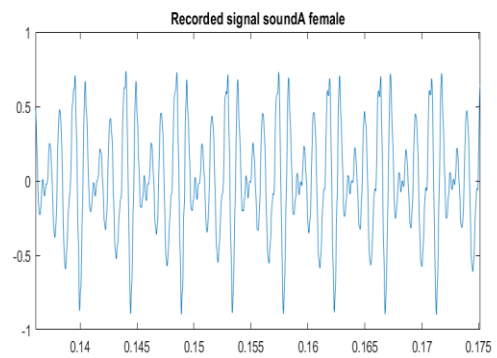
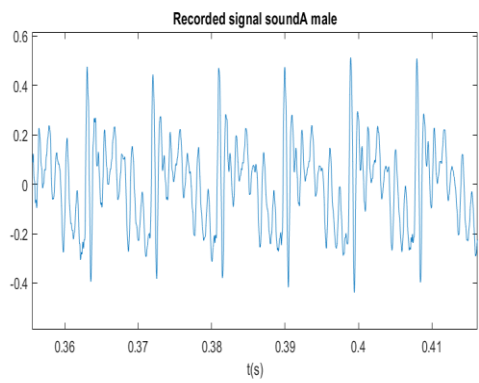
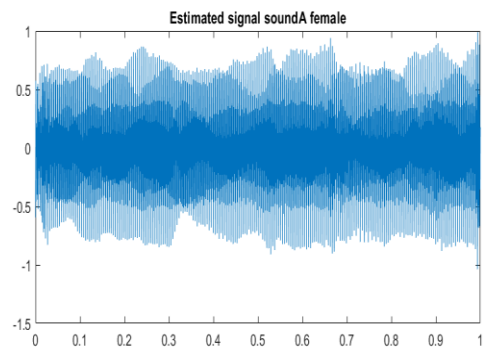
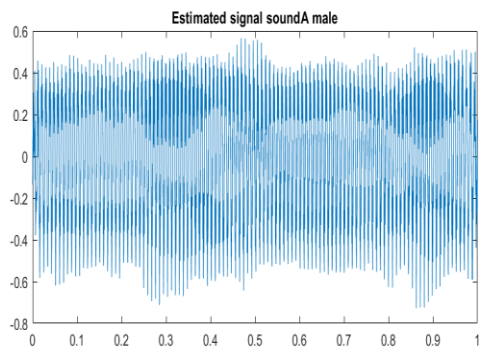
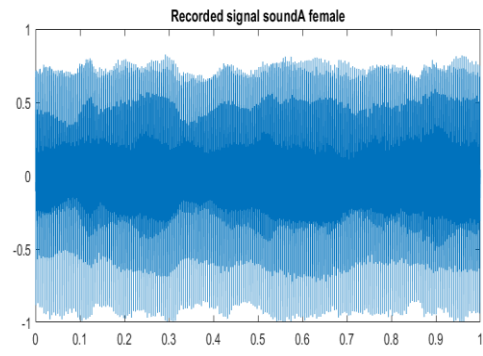
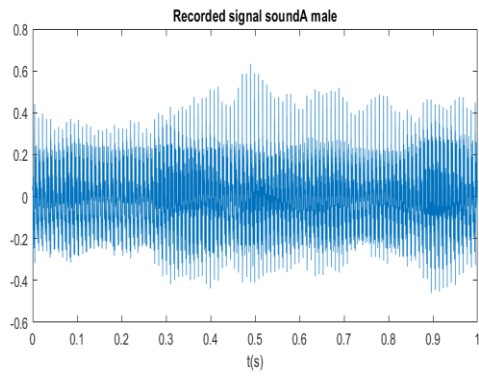


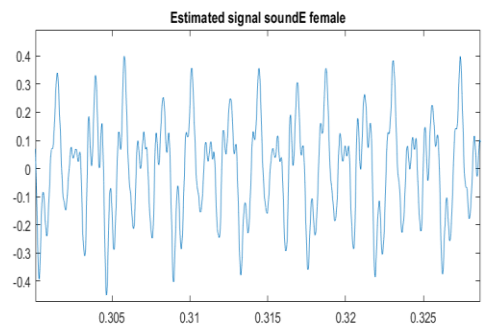
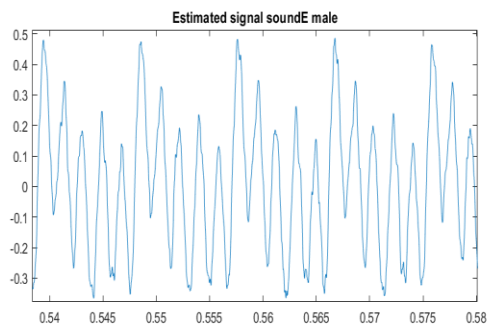
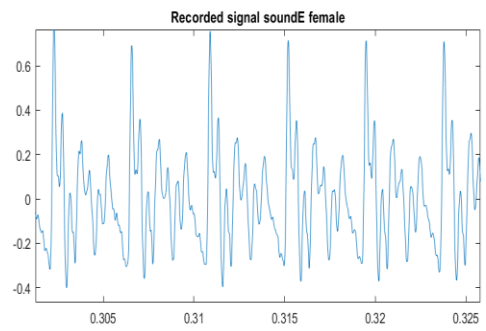
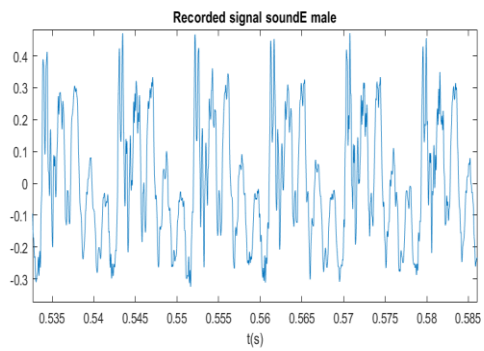
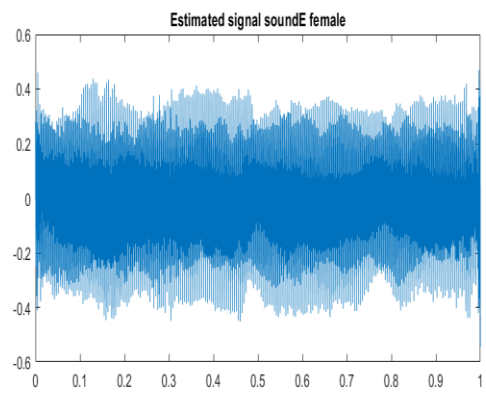
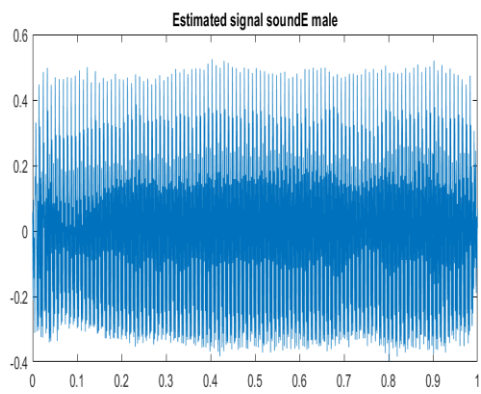
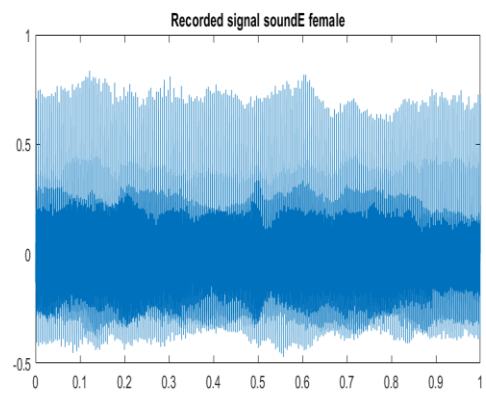
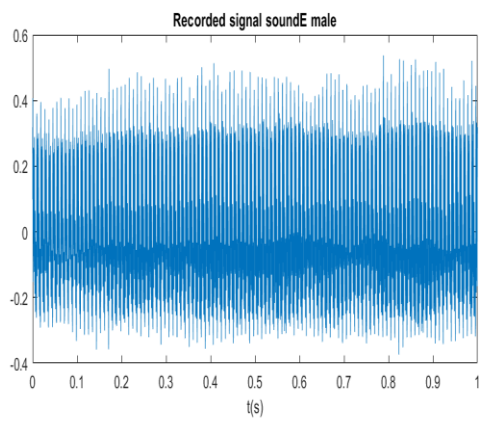


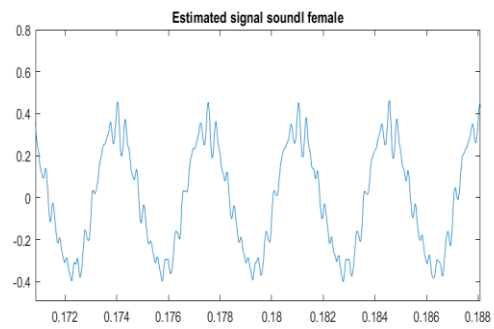
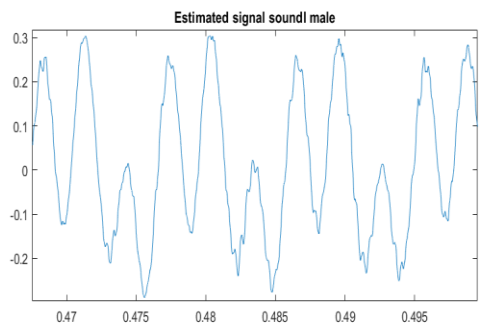
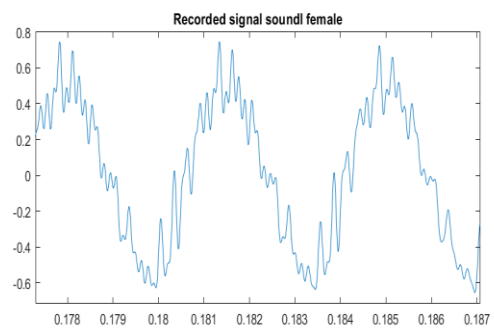
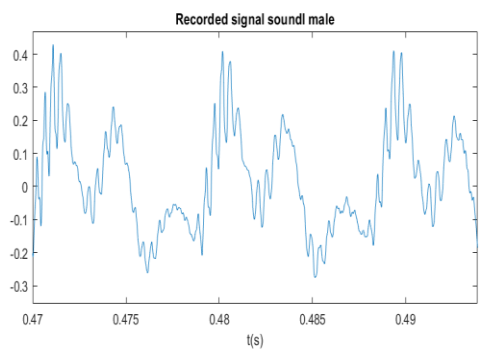
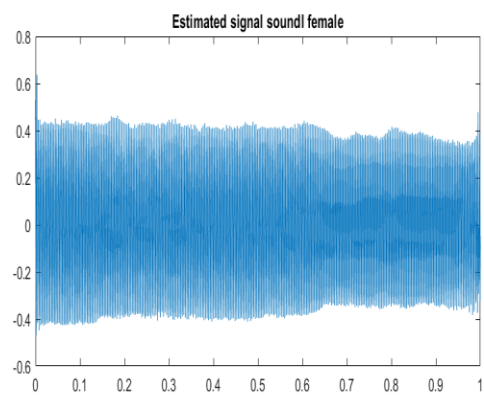
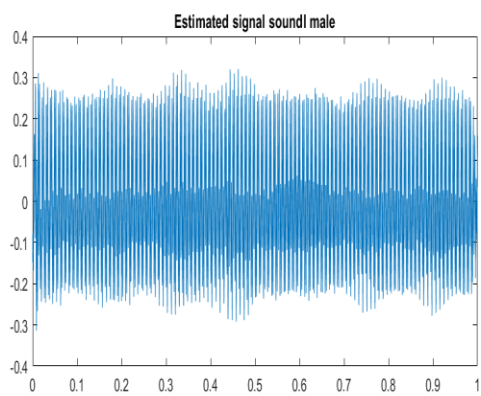
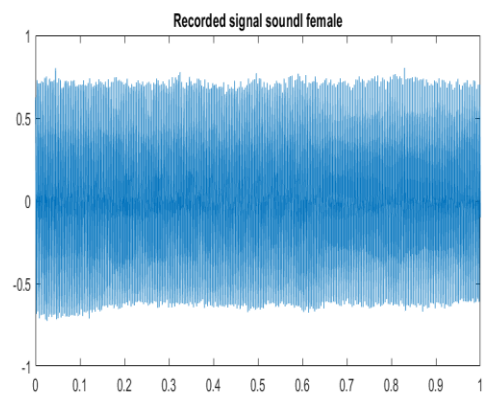
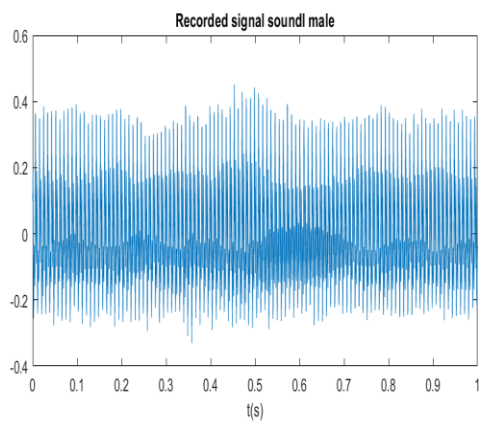
Για το lifter στο cepstrum, για να χωρίσω το voicing από την συνάρτηση μεταφοράς, ήθελα να χρησιμοποιήσω ένα HPF. Ωστόσο το αποτέλεσμα του να χρησιμοποιήσω ένα LPF και να αφαιρέσω το υπόλοιπο από το σήμα δούλεψε καλύτερα, τουλάχιστον ηχητικά. Οπότε αυτό που έκανα, ήταν να χρησιμοποιήσω ένα παράθυρο hann (αντί για τετραγωνικό $1.6(0,8)$) μεγέθους $2.8 * \text{period}$ (1,4 στην πραγματικότητα), και να αφαιρέσω το αποτέλεσμα του γινομένου από το αρχικό cepstrum. Σκεπτόμενος πως ναι μεν αυτό που θέλω να κρατήσω είναι γύρω από το $1 * \text{period}$ και μέχρι τέλους, αλλά γύρω στο πρώτη περίοδο θέλω να μειώσω όσο γίνεται την επικάλυψη.

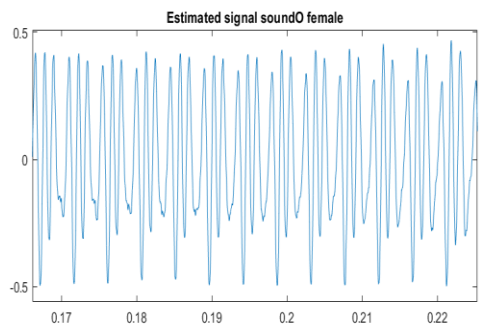
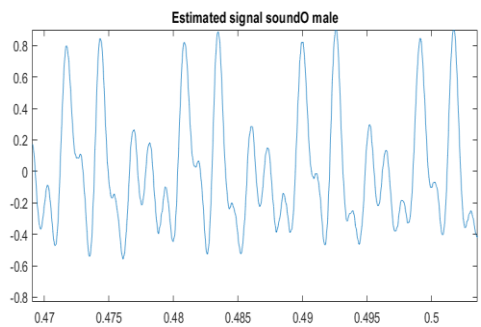
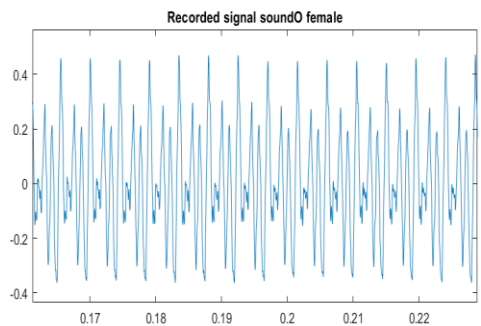
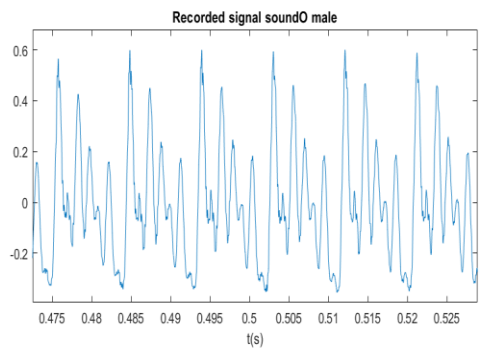
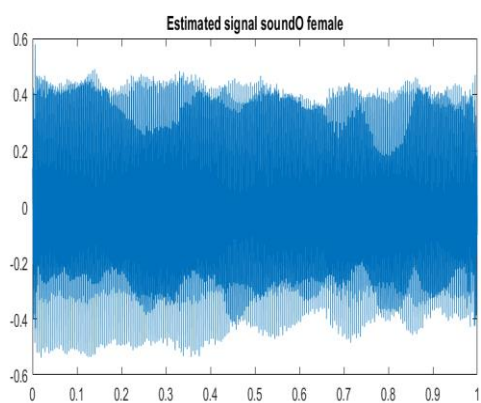
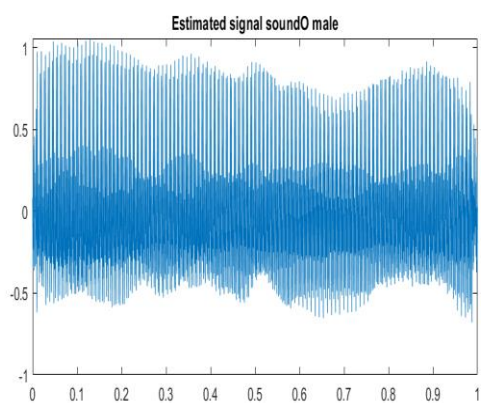
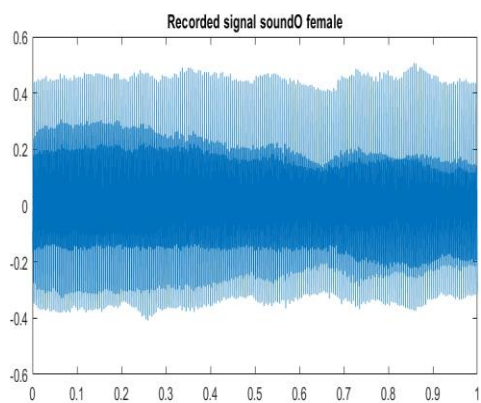
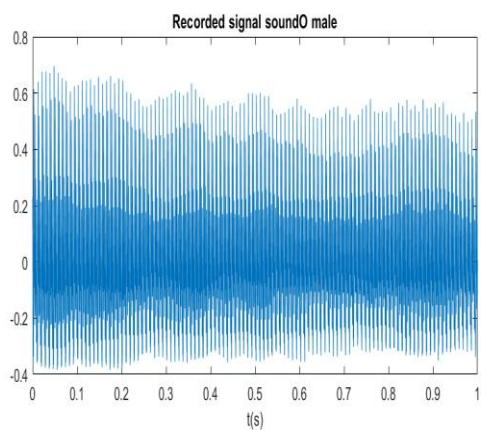
Το αποτέλεσμα όπως φαίνεται μοιάζει με το πραγματικό σήμα, αλλά με πολύ λιγότερη ενέργεια, για αυτό και ο θόρυβος είναι πιο εμφανής (αφού δεν έκοψα τις πολύ ψηλές συχνότητες στο cepstrum).

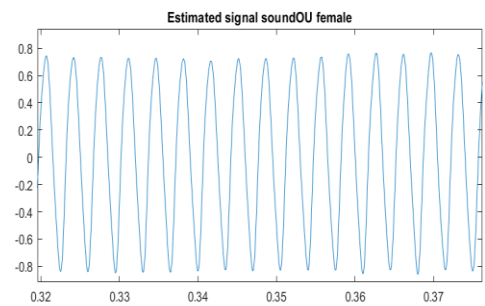
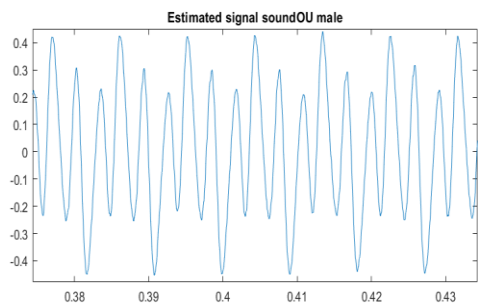
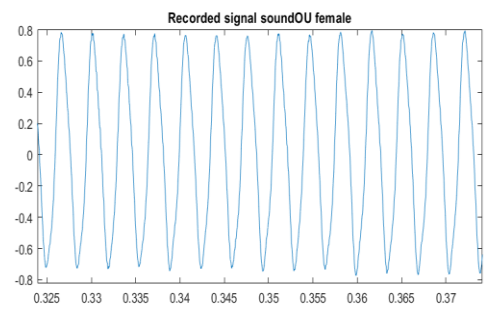
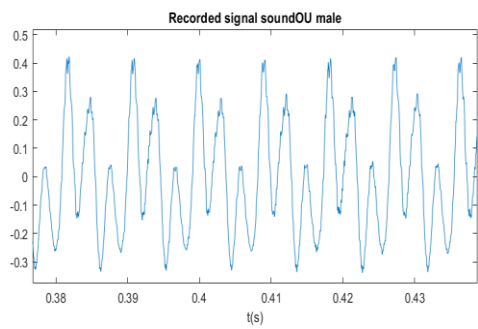
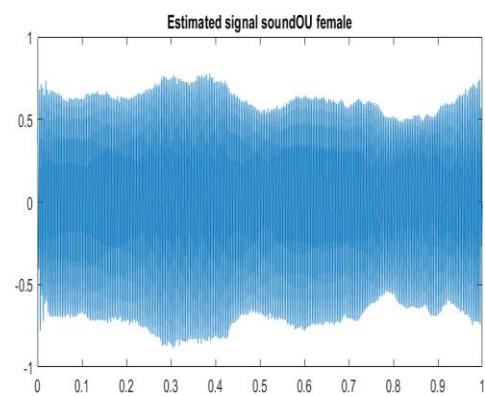
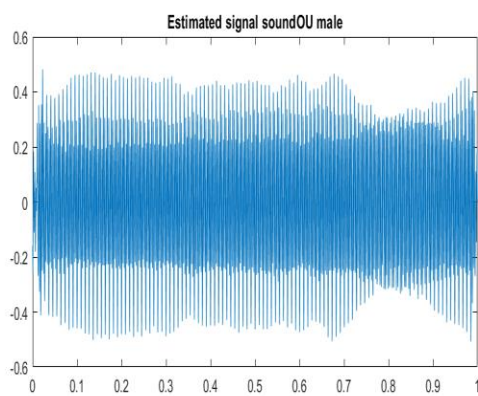
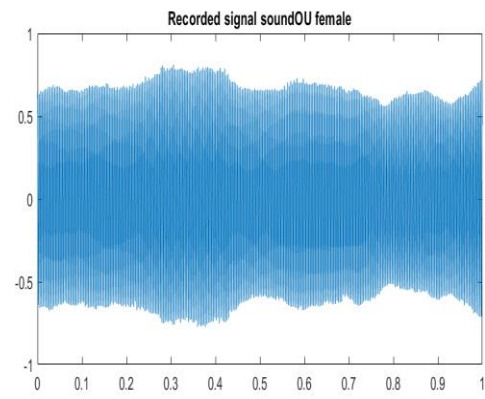
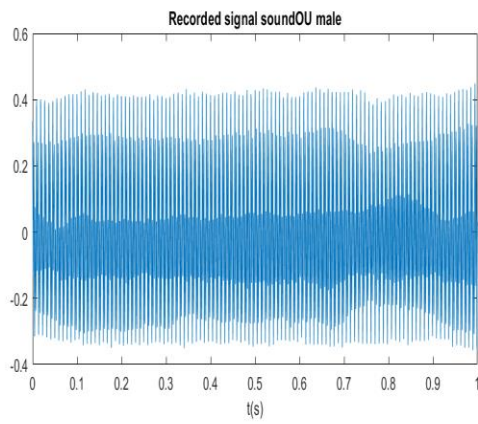
Ερώτημα 4ο











Γενικά τα αποτελέσματα ήταν καλά ακουστικά, βέβαια οπτικά παρουσιάζουν κάποιες διαφορές. Αν εξαιρέσουμε το αντρικό Α και το

γυναικείο I, τα υπόλοιπα οπτικά φαίνονται μια καλή προσέγγιση όταν τα εξετάσουμε κοντά. Σε όλες τις εκτιμήσεις, ωστόσο, είναι κατανοητός και ο φθόγγος αλλά και το ποιος το λέει. Δηλαδή δεν χάνεται η πληροφορία της κρουστικής απόκρισης.

Δοκίμασα να διεγείρω το σύστημα με συρμό ώσεων ή συρμό sinc ακόμα και συρμό από 3 περιόδους του σήματος του 3ου ερωτήματος, αλλά τα πιο κοντινά αποτελέσματα τα έδινε αυτή η υλοποίηση. Αν και σε όλες τις διεγέρσεις ήταν κατανοητός ο φθόγγος, απλά ήταν πιο buzzy.

Όσον αφορά το παράθυρο στον χρόνο (1^ο βήμα), αρχικά έκοψα κι άλλο το σήμα σε μήκος 50ms κι εκεί με κέντρο μια κορυφή και μήκος 2 και 3 pitch για τα γυναικεία και τα αντρικά σήματα αντίστοιχα, πολλαπλασίασα με ένα παράθυρο hamming. Ύστερα εφάρμοσα low pass lifter στο cepstrum ένα παράθυρο hann στο 85% του quefrency. Άρα για τα αντρικά πήρα μία περισσότερη περίοδο. Γιατί όπως είχα παρατηρήσει πριν, έχουν πιο απλωμένη την πληροφορία της κρουστικής απόκρισης στο cepstrum (ή μάλλον μεγαλύτερη επικάλυψη).

Μερικά λόγια για τον κώδικα σε σύνολο:

Για να υπολογίσω τα, cepstrum χρησιμοποίησα τις συναρτήσεις του matlab rceps/cceps/icceps. Το pitch το υπολόγισα, από την πρώτη κορυφή κάθε cepstrum κάνοντας χρήση της συνάρτησης findpeaks που εντοπίζει μερικά μέγιστα.

Ο κώδικας:

```
ex4
% allinone =
{files10.A,files10.wA;files10.E,files10.wE;files10.I,files10.w
I;
%     files10.O,files10.wO;files10.OU,files10.wOU;};
close all
clear
clc
%% Task 1 (ready)
files10 = load('allinone.mat');
files10 = files10.allinone;
fs = 192000;
sounds = {'A';'E';'I';'O';'OU'};
packs = cell(5,1);
pitches = zeros(5,2);%pitch estimation
p = zeros(fs,5,2);
h = cell(5,2);
Xest = p;
%pitches2 = zeros(5,2);
%% Task 2
for i = 1:5
    packs{i} = signalpack(files10(i,:));
```



```

    cvreal(packs{i},sounds{i},fs)
%    alternatively
%    open([sounds{i},'.fig'])
    for j = 1:2
        pitches(i,j) = mypitch(packs{i}{2,j},fs);
        p(:,i,j) = voicing(packs{i}{3,j},pitches(i,j),2.8,fs);
        if j == 1
            h{i,j}=impulrep(packs{i}{1,j},pitches(i,j),fs,3);
        else
            h{i,j}=impulrep(packs{i}{1,j},pitches(i,j),fs,2);
        end
%        [~,~,D] = alignsignals(h{i,j},p(:,i,j));%uses xcorr
        X = conv(circshift(p(:,i,j),0),h{i,j},'same');
        Xest(:,i,j) = 2.3*X(1:fs);
%        Uncomment this section to hear the signals
        (optional)
%        playX = audioplayer(packs{i}{1,j},fs);
%        play(playX)
%        pause(1)
%        playXest = audioplayer(Xest(:,i,j),fs);
%        play(playXest)
%        pause(1)
    end
    name = ['Task 3: sound ',sounds{i}];
    figure('Name',name,'NumberTitle','off')
    decvreal(packs{i}{1,1},p(:,i,1),sounds{i},'male',fs,[1,3])

    decvreal(packs{i}{1,2},p(:,i,2),sounds{i},'female',fs,[2,4])
    name = ['Task 4: sound ',sounds{i}];
    figure('Name',name,'NumberTitle','off')

    estvreal(packs{i}{1,1},Xest(:,i,1),packs{i}{1,2},Xest(:,i,2),
        ...fs,sounds{i})
end

```

```

function pt = voicing(csig,pitch,winlen,fs)
period = pitch.^(-1);
win = hann(round(winlen*period*fs)+1);
w = [win(round((end+1)/2):end);zeros(fs-
length(win),1); win(1:floor(end/2))];
env = csig.*w;
p = csig - env;
pt = ifft(exp(fft(p)));

```

```

function decvreal(X,p,letter,gender,fs,col)
subplot(2,2,col(1))
plot(0:1/fs:1-1/fs,X)

```



```

title(['Reocrded Signal ',letter,' ',gender])
xlabel('t(s)')
subplot(2,2,col(2))
plot(0:1/fs:1-1/fs,p)
title(['Deconvolved signal ',letter,' ',gender])
xlabel('t(s)')

```

```

function estvreal(X,Xest,Xf,Xestf,fs,letter)
    tiledlayout(2,2)
    nexttile
    plot(0:1/fs:1-1/fs,X)
    title(['Recorded signal sound',letter,' male'])
    xlabel('t(s)')
    nexttile
    plot(0:1/fs:1-1/fs,Xf)
    title(['Recorded signal sound',letter,' female'])
    nexttile
    plot(0:1/fs:1-1/fs,Xest)
    title(['Estimated signal sound',letter,' male'])
    nexttile
    plot(0:1/fs:1-1/fs,Xestf)
    title(['Estimated signal sound',letter,' female'])

```

```

%%
function spack = signalpack(s)
    spack = cell(3,2);
    for i = 1:length(s)
        spack{1,i} = s{i};
        spack{2,i} = rceps(s{i});%ifft(log(abs(fft(s))))
        spack{3,i} = cceps(s{i});
    end

```

```

%% pitch by cepstrum
function l = mypitch(cS,fs)
    [~,l] =
    findpeaks(cS,fs,'MinPeakHeight',0.01,'MinPeakDistance',0.002);
    l = l(l>0.0025); % < 400Hz
    l = 1/l(1);

```

```

%% plot function exe4
function cvreal(signals,letter,fs)
%signalM,rcepstrumM,ccepstrumM,signalW,rcepstrumW,ccepstrumW
datacheck = zeros(6,1);
names = {'Signal Male','Signal Female';'Cepstrum Male
(Real)',...
        'Cepstrum Female (Real)';'Cepstrum Male (cmplx)',...
        'Cepstrum Female (cmplx)'};};

for i = 1:3
    for j = 1:2
        datacheck(i*j) = max(size(signals{i,j}));
    end
end
if max(datacheck - max(datacheck))
    error('Error: Data vectors must have the same length.')
end
figure('Name',letter,'NumberTitle','off')
tiledlayout(3,2)
for i = 1:3
    for j = 1:2
        if i == 1

rplots(signals{i,j},fs,datacheck(1)/fs,names{i,j},letter,...
        't(s)')
            else

rplots(signals{i,j},fs,datacheck(1)/fs,names{i,j},letter,...
        'Quefreny(s)')
            end
        end
    end
end

function rplots(s,fs,edge,name,letter,xax)
    name = [name, ' ', letter];
    xvec = 0:1/fs:edge-1/fs;
    nexttile
    plot(xvec,s)
    title(name)
    xlabel(xax)
end
end

```

```

function h_hat = impulrep(sig,pitch,fs,winlen)
period = pitch.^(-1);
i = round(fs/3);
A100 = sig(i:i+0.1*fs);

```

```
[~,j] = findpeaks(A100, 'MinPeakDistance', period*0.9*fs);  
Aw = A100(j(2):j(round(2+winlen)));  
Aw = Aw.*hamming(length(Aw));  
cAw = cceps(Aw);  
wind = hann(round(0.85*winlen*period*fs));  
windw = [wind(round((end+1)/2):end); zeros(length(cAw)-  
length(wind),1); ...  
         wind(1:floor(end/2))];  
h_hat = cAw.*windw;  
h_hat = icceps(h_hat);
```