

# Προηγμένες Τεχνικές Επεξεργασίας Σήματος

## Εργασία 4η: Τεχνικές Ανάλυση Ανθρώπινης Φωνής και Αποσυνέλιξη Σημάτων με Ομομορφικό Φιλτράρισμα

Ονοματεπώνυμο: Χριστόδουλος Μιχαηλίδης

Διεύθυνση Ηλεκτρονικού Ταχυδρομείου: [chrdimmic@ece.auth.gr](mailto:chrdimmic@ece.auth.gr)

Ημερομηνία: 21 Ιουλίου 2022

<b>Εκφώνηση Εργασίας</b>	<b>3</b>
<b>Ηχογράφηση Δειγμάτων Ανθρώπινης Ομιλίας</b>	<b>4</b>
<b>Ποιοτικές Διαφορές των κυματομορφών στο πεδίο του χρόνου</b>	<b>4</b>
<b>Ποιοτικές διαφορές των κυματομορφών στο πεδίο του cepstrum</b>	<b>7</b>

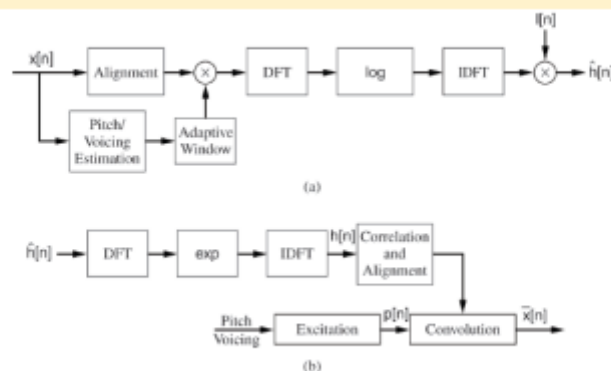
## Εκφώνηση Εργασίας

1. Acquire voice samples. In this part, please either record or find at least 10 voice samples of a male and female individual making the five vowel sounds – “a”, “e”, “i”, “o”, “u”. If you are going to record them yourself or using a friend, then exaggerate the sounds a little and keep your voice extended for a while. Please make note of the conditions used to obtain the voice samples (e.g., what smartphone, what type of speaker – built-in or microphone, using which software program, or where or from whom the files were obtained). You should have 10 files in the end. If you want to be keen and impressive, you can get more than one male and female voice to obtain a better understanding of differences between both signals in the cepstral domain. Have fun with this.

2. Compute the cepstrum of each voice signal and discuss any difference qualitatively and quantitatively amongst male and female voices in general and amongst the different vowel sounds. This is an important component of the project, so please be creative and as comprehensive as possible. Your report should provide figures with original time-domain signals as well as cepstrum signals. Female voices should generally have more peaks than male voices in the cepstrum domain. You should discuss why you think this would be the case.

3. Lifter the cepstrum domain signals. Design a window (length is an important design parameter and you should discuss how and what you select – it can be the same or different for each speech sample depending on what you would like to experiment with) to remove the transfer function dependency. Then, compute the time domain signal of the corresponding windowed result to obtain the deconvolved signal. Plot the deconvolved result. Is there anything you can say about the signal and its difference from the original time domain recorded sample? Again, your discussion is an important part of the report.

4. Try to synthesize back the voiced signals as follows and comment on your results.



# Ηχογράφηση Δειγμάτων Ανθρώπινης Ομιλίας

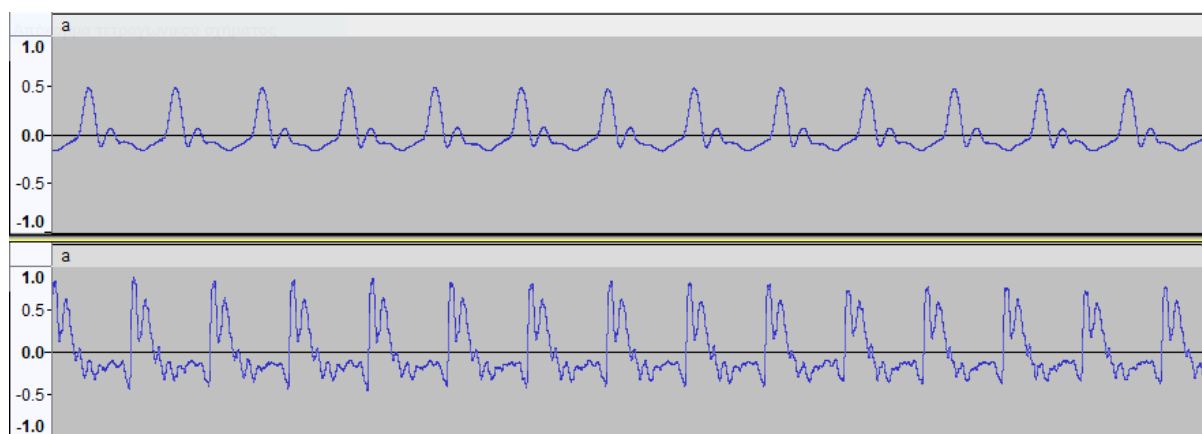
Στα πλαίσια αυτής της εργασίας ηχογραφήθηκαν τα πέντε φωνήεντα του ελληνικού αλφαβήτου, όπως αυτά προφέρονται από δύο διαφορετικά άτομα (άνδρας ηλικίας 24 ετών και γυναίκα ηλικίας 30 ετών). Η ηχογράφηση έγινε χρησιμοποιώντας το μικρόφωνο ενός κινητού τηλεφώνου (μοντέλο Xiaomi Redmi 9) χρησιμοποιώντας την προεγκατεστημένη εφαρμογή ηχογράφησης του λειτουργικού συστήματος Android.

Η ηχογράφηση πραγματοποιήθηκε σε δωμάτιο με καλή ηχομόνωση, έτσι ώστε να ελαχιστοποιηθούν οι παρεμβολές από εξωτερικούς ήχους. Το μικρόφωνο τοποθετήθηκε σε απόσταση περίπου 20 εκατοστών από τον ομιλητή. Τέλος, ανάμεσα στον ομιλητή και στο μικρόφωνο τοποθετήθηκε μία χαρτοπετσέτα, η οποία χρησιμοποιήθηκε ως αντιανέμιο (pop filter), προκειμένου το μικρόφωνο να μην ηχογραφήσει τον ήχο της ανθρώπινης αναπνοής.

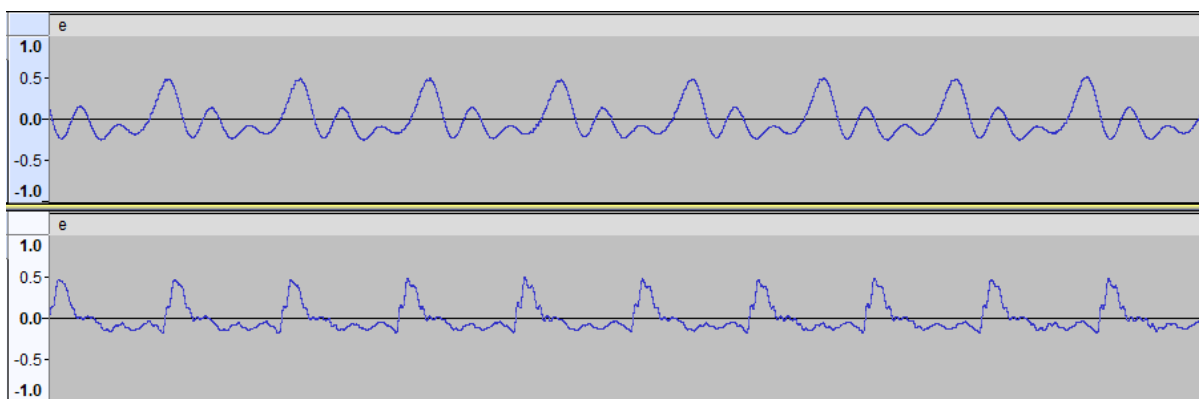
Τα δείγματα ήχου αποθηκεύτηκαν σε αρχείο επέκτασης wav και ηχογραφήθηκαν σε συχνότητα δειγματοληψίας 192kHz. Στη συνέχεια, χρησιμοποιώντας το λογισμικό Audacity τα δύο κανάλια της κάθε ηχογράφησης συγχωνεύτηκαν σε ένα κανάλι. Κάθε κυματομορφή πέρασε από ένα στάδιο ενίσχυσης, προκειμένου όλες οι ηχογραφήσεις να έχουν περίπου το ίδιο πλάτος. Τέλος, αφαιρέθηκε η αρχή και το τέλος κάθε ηχογράφησης, έτσι ώστε να αφαιρεθούν μεταβατικά φαινόμενα από την ανθρώπινη ομιλία. Το τελικό αρχείο αποθηκεύτηκε και πάλι ως αρχείο wav, χρησιμοποιώντας κωδικοποίηση float-64bit. Δηλαδή κάθε δείγμα αναπαρίσταται ως ένας αριθμός κινητής υποδιαστολής και διπλής ακρίβειας.

## Ποιοτικές Διαφορές των κυματομορφών στο πεδίο του χρόνου

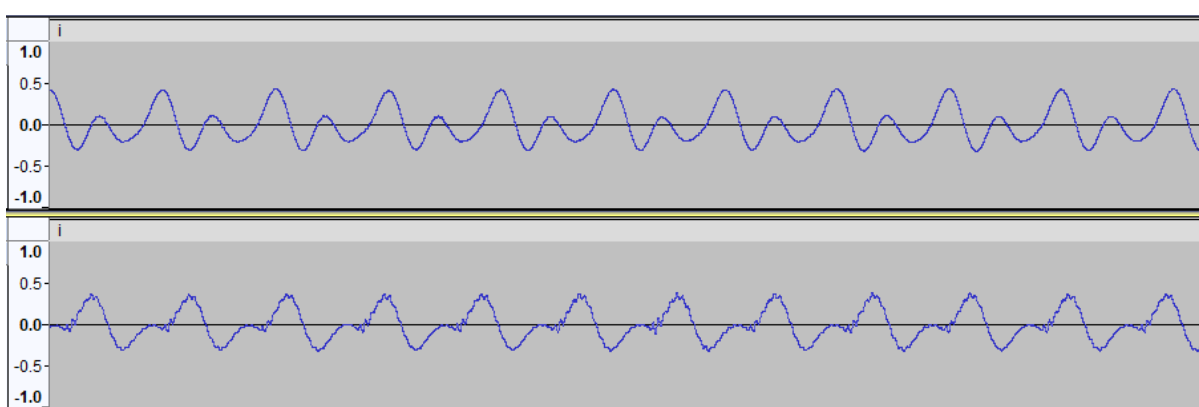
Στα επόμενα 5 διαγράμματα βλέπουμε αποσπάσματα από τις κυματομορφές με χρονική διάρκεια μερικές δεκάδες χιλιοστά του δευτερολέπτου. Όπως αυτά πάρθηκαν από το GUI του λογισμικού Audacity.



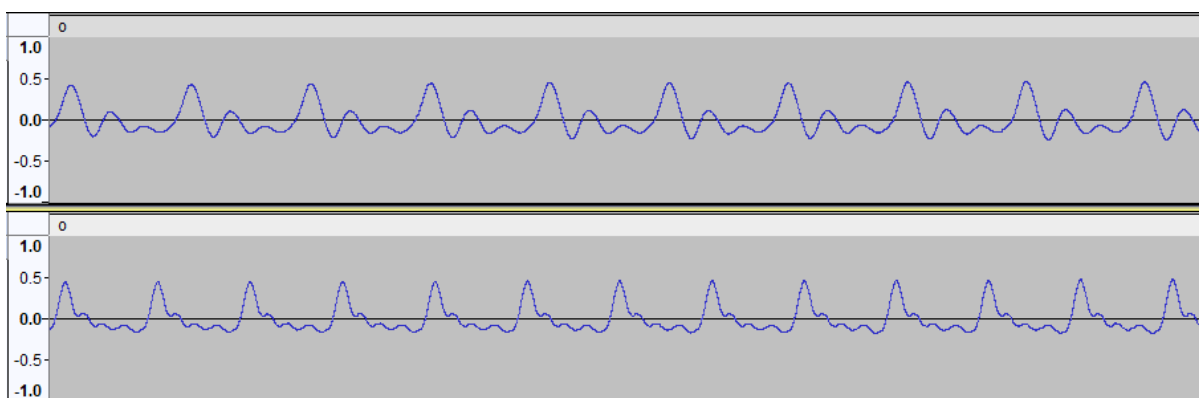
Διάγραμμα 1) Στην άνω κυματομορφή φαίνεται το φωνήεν “α” όπως αυτό προφέρεται από έναν άνδρα. Στην κάτω κυματομορφή φαίνεται η προφορά μιας γυναίκας για το ίδιο φωνήεν.



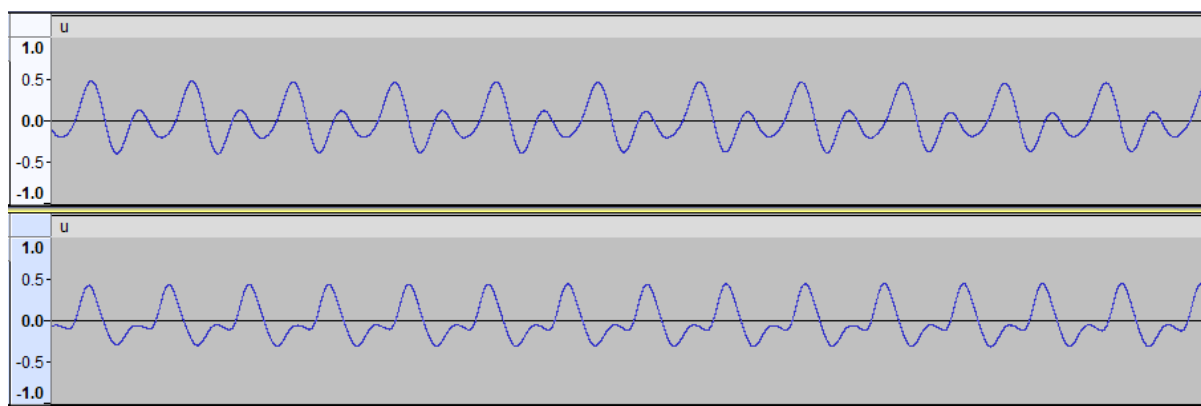
Διάγραμμα 2) Στην άνω κυματομορφή φαίνεται το φωνήεν “ε” όπως αυτό προφέρεται από έναν άνδρα. Στην κάτω κυματομορφή φαίνεται η προφορά μιας γυναίκας για το ίδιο φωνήεν.



Διάγραμμα 3) Στην άνω κυματομορφή φαίνεται το φωνήεν “ι” όπως αυτό προφέρεται από έναν άνδρα. Στην κάτω κυματομορφή φαίνεται η προφορά μιας γυναίκας για το ίδιο φωνήεν.



Διάγραμμα 4) Στην άνω κυματομορφή φαίνεται το φωνήεν “ο” όπως αυτό προφέρεται από έναν άνδρα. Στην κάτω κυματομορφή φαίνεται η προφορά μιας γυναίκας για το ίδιο φωνήεν.



Διάγραμμα 5) Στην άνω κυματομορφή φαίνεται το φωνήεν “ου” όπως αυτό προφέρεται από έναν άνδρα. Στην κάτω κυματομορφή φαίνεται η προφορά μιας γυναίκας για το ίδιο φωνήεν.

Παρατηρούμε ότι για μικρά χρονικά παράθυρα οι κυματομορφές μπορούν να θεωρηθούν περιοδικά σήματα, επειδή σε κάθε κυματομορφή βλέπουμε επαναλήψεις του ίδιου μοτίβου. Με εξαίρεση το φωνήεν “ου” όπου οι κυματομορφές του άνδρα και της γυναίκας φαίνεται να μοιάζουν αρκετά μεταξύ τους, σε όλα τα άλλα φωνήεντα το περιοδικό μοτίβο είναι διαφορετικό στα δύο φύλα. Γενικά, τα περιοδικά μοτίβα του άνδρα είναι πιο ομαλά, ενώ στα περιοδικά μοτίβα της γυναίκας φαίνεται να έχει επικαθίσει και ένα σήμα υψηλής συχνότητας, το οποίο κάνει τις κυματομορφές να δείχνουν πιο “γωνιώδεις”.

Επειδή σε κάθε φωνήεν ο άξονας του χρόνου είναι κοινός για τα δύο φύλα, μπορούμε να κάνουμε μια γρήγορη εκτίμηση για την θεμελιώδη συχνότητα κάθε κυματομορφής μετρώντας τις επαναλήψεις του περιοδικού φαινομένου.

Τα αποτελέσματα φαίνονται στον παρακάτω πίνακα

φωνήεν\φύλο	Άνδρας	Γυναίκα
“α”	13 επαναλήψεις	15 επαναλήψεις
“ε”	8 επαναλήψεις	10 επαναλήψεις
“ι”	10 επαναλήψεις	12 επαναλήψεις
“ο”	10 επαναλήψεις	13 επαναλήψεις
“ου”	11 επαναλήψεις	15 επαναλήψεις

Για κάθε φωνήεν παρατηρούμε ότι στο ίδιο χρονικό διάστημα η γυναικεία φωνή παρουσιάζει περισσότερες επαναλήψεις του περιοδικού φαινομένου σε σχέση με την ανδρική. Αυτό αποτελεί μια ισχυρή ένδειξη ότι η γυναικεία φωνή έχει υψηλότερη θεμελιώδη συχνότητα από την ανδρική φωνή (τουλάχιστον στο συγκεκριμένο δείγμα).

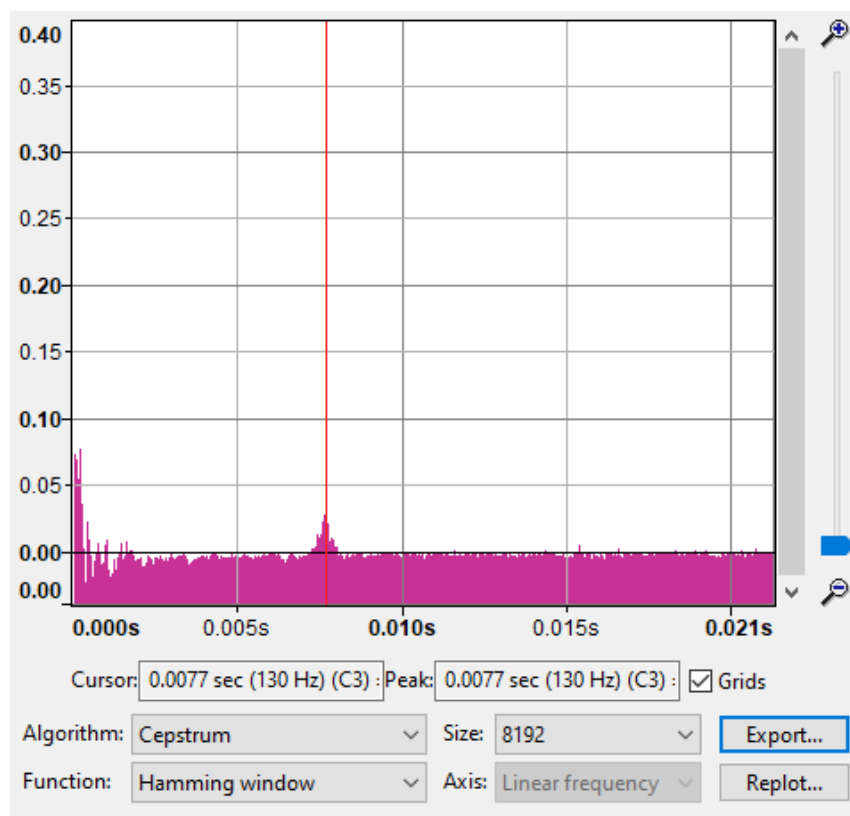
## Ποιοτικές διαφορές των κυματομορφών στο πεδίο του cepstrum

Προκειμένου να υπολογιστεί με ακρίβεια η θεμελιώδης συχνότητα της κάθε κυματομορφής υπολογίζεται το cepstrum του κάθε δείγματος.

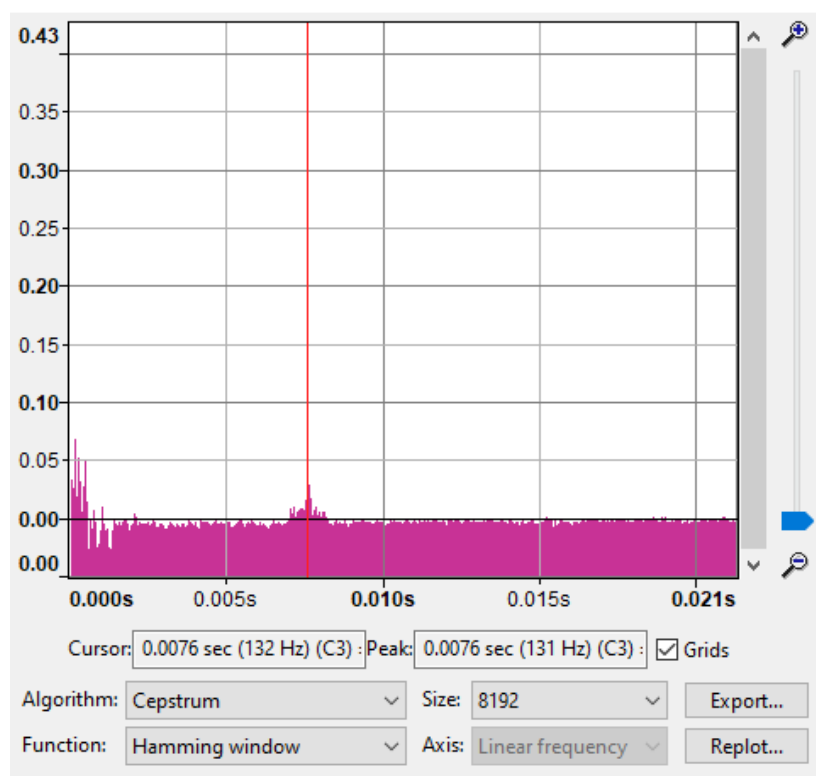
Πιο συγκεκριμένα το cepstrum υπολογίστηκε με τον εξής τρόπο για όλες τις περιπτώσεις:

- Για κάθε κυματομορφή επιλέχθηκε ένα απόσπασμα στο οποίο παρατηρούνται 26 επαναλήψεις του περιοδικού φαινομένου.
- Το κάθε απόσπασμα διαμερίστηκε σε μη επικαλυπτόμενα υποσύνολα μήκους 8192 δειγμάτων. Για συχνότητα δειγματοληψίας 192kHz αυτό αντιστοιχεί σε χρονικό διάστημα μήκους 42.6 msec.
- Σε κάθε επιμέρους υποσύνολο εφαρμόστηκε ένα παράθυρο Hanning
- Για κάθε παραθυρωμένο υποσύνολο υπολογίστηκε το cepstrum.
- Η τελική εκτίμηση προκύπτει ως ο μέσος όρος όλων των υποσυνόλων

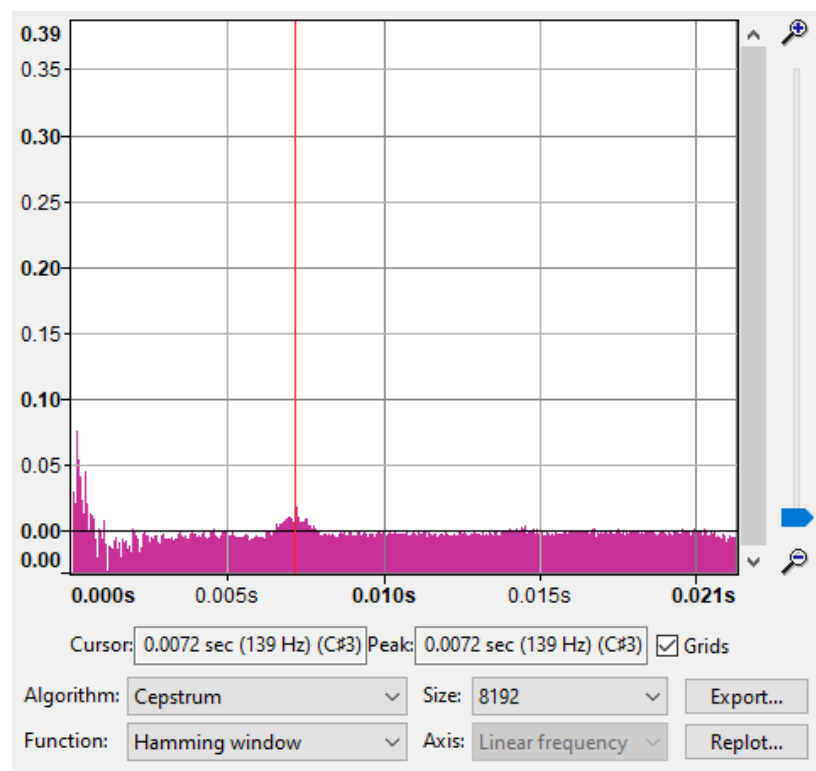
Στα επόμενα διαγράμματα φαίνονται οι εκτιμήσεις των cepstra, όπως αυτές υπολογίστηκαν μέσω του λογισμικού Audacity χρησιμοποιώντας τον παραπάνω αλγόριθμο.



Διάγραμμα 6) Cepstrum ανδρικής φωνής για το φωνήεν “α”

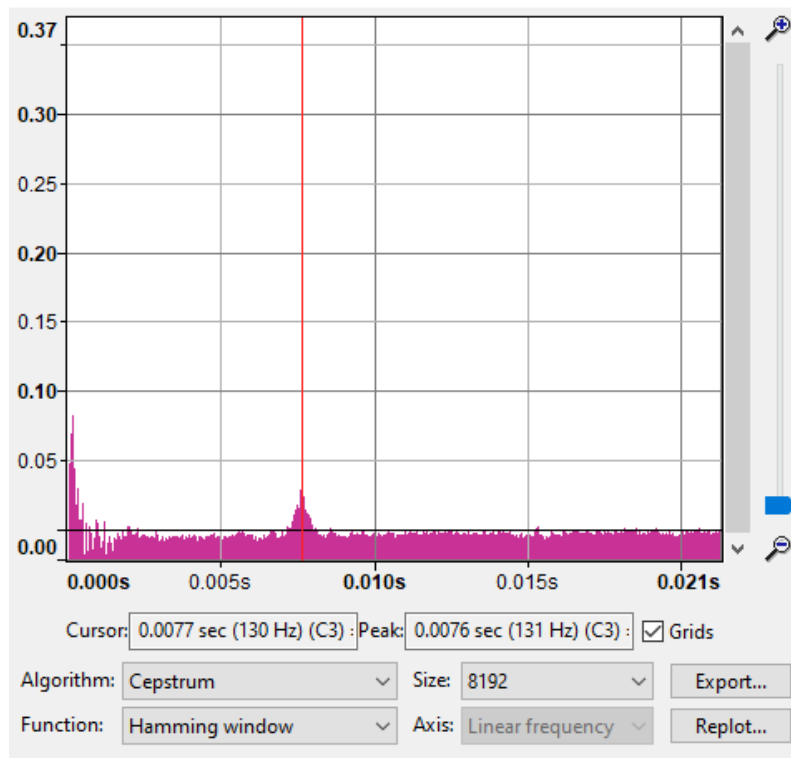


Διάγραμμα 7) Cepstrum ανδρικής φωνής για το φωνήεν “ε”

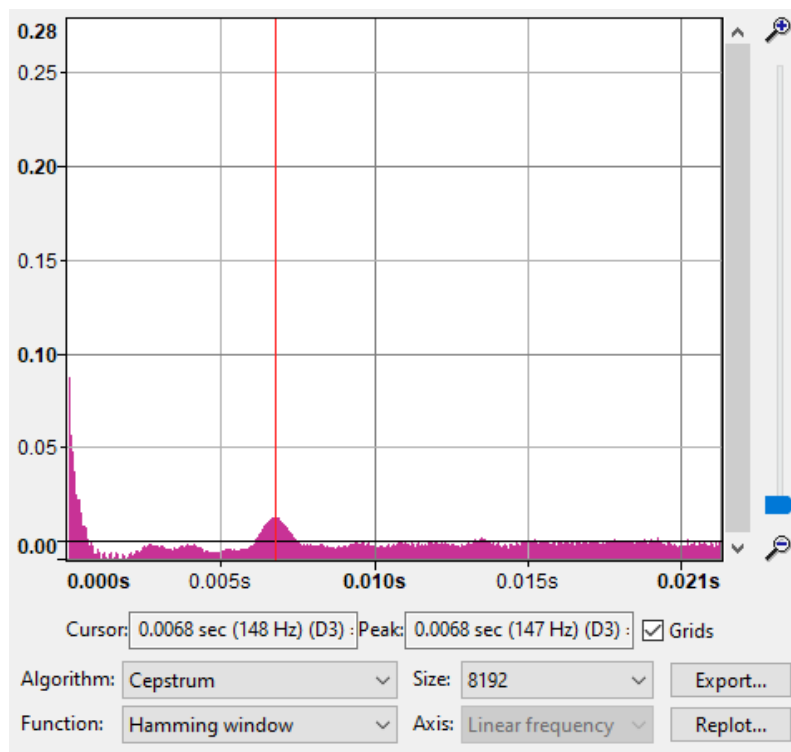


Διάγραμμα 8) Cepstrum ανδρικής φωνής για το φωνήεν “ι”

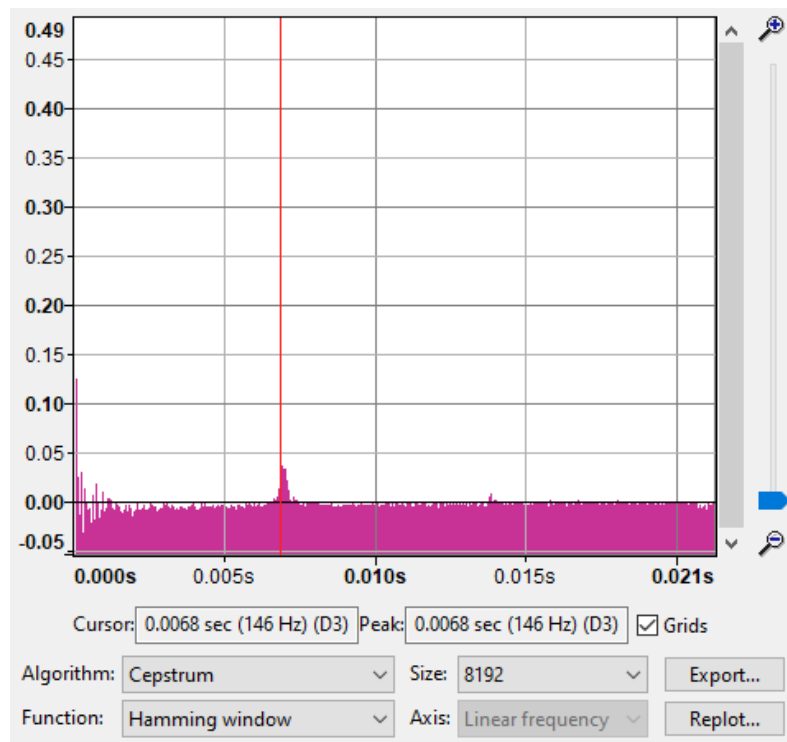




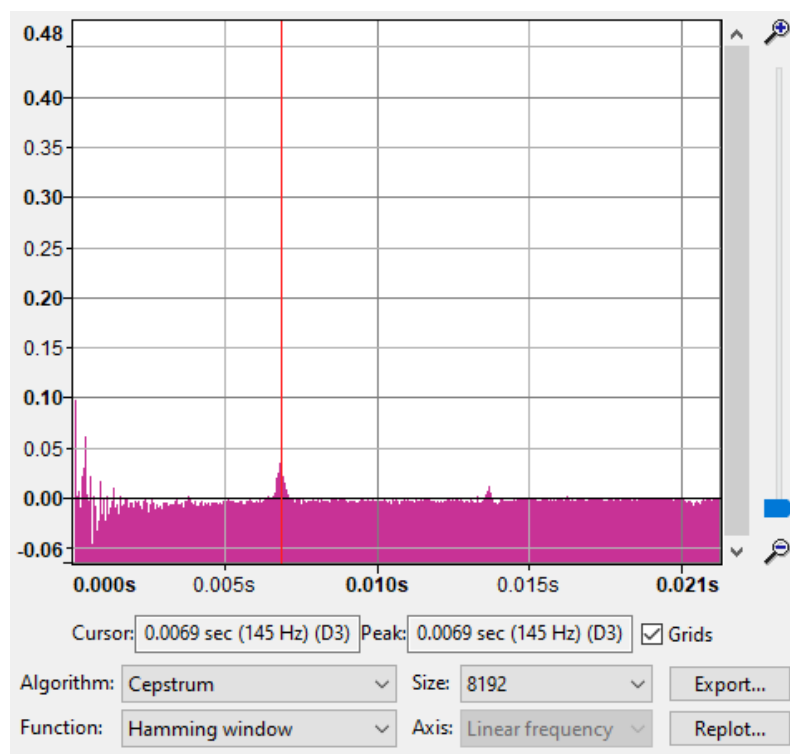
Διάγραμμα 9) Cepstrum ανδρικής φωνής για το “ο”



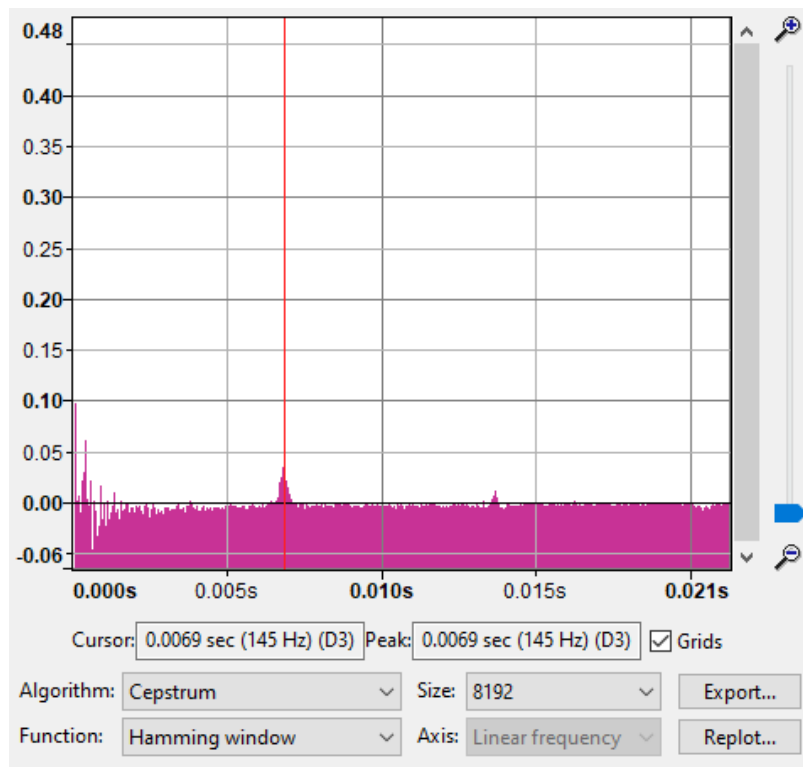
Διάγραμμα 10) Cepstrum ανδρικής φωνής για το φωνήεν “ου”



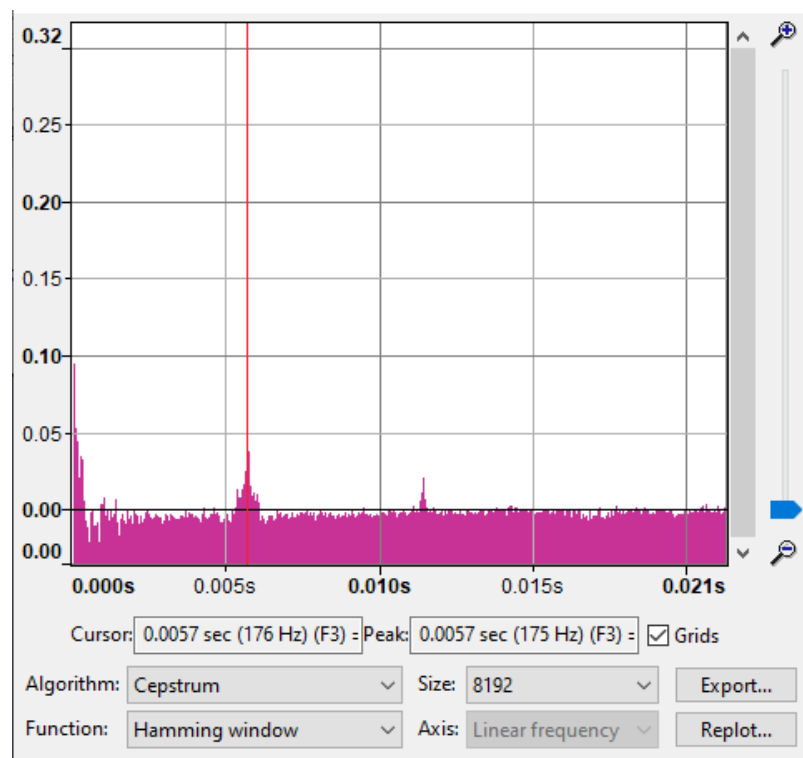
Διάγραμμα 11) Cepstrum γυναικείας φωνής για το φωνήεν “α”



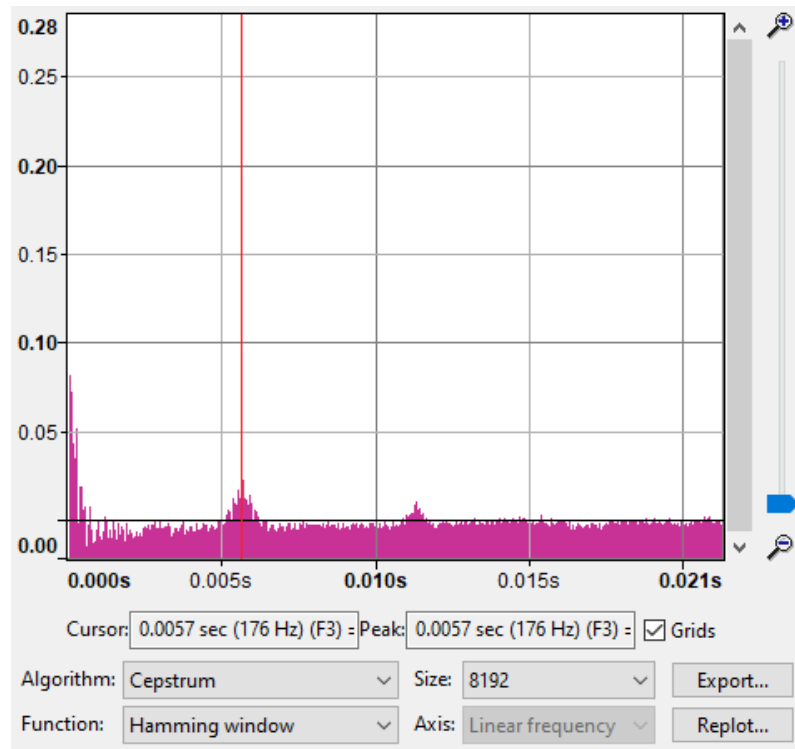
Διάγραμμα 12) Cepstrum γυναικείας φωνής για το φωνήεν “ε”



Διάγραμμα 12) Cepstrum γυναικείας φωνής για το φωνήεν “i”



Διάγραμμα 13) Cepstrum γυναικείας φωνής για το φωνήεν “o”



Διάγραμμα 14) Cepstrum γυναικείας φωνής για το φωνήεν “ου”

Η πρώτη κορυφή που συναντάμε στο cepstrum αντιστοιχεί στην θεμελιώδη συχνότητα της υπό-μελέτη κυματομορφής. Τα αποτελέσματα παρουσιάζονται στον ακόλουθο πίνακα:

φωνήεν/φύλο	Άνδρας	Γυναίκα
“α”	7.7msec, 130Hz	6.8msec, 146Hz
“ε”	7.6msec, 131Hz	6.9msec, 145Hz
“ι”	7.2msec, 139Hz	6.9msec, 145Hz
“ο”	7.6msec, 131Hz	5.7msec, 175Hz
“ου”	6.8msec, 147Hz	5.7msec, 176Hz

Παρατηρούμε ότι σε κάθε φωνήεν η θεμελιώδης συχνότητα της ανδρικής φωνής είναι χαμηλότερη από την θεμελιώδη της γυναικείας. Πέρα από τη διαφορά στην θεμελιώδη συχνότητα παρατηρείται ακόμα μία σοβαρή διαφορά ανάμεσα στο cepstrum της ανδρικής και της γυναικείας φωνής. Στην περίπτωση της γυναικείας φωνής παρατηρείται για κάθε φωνήεν και μία δεύτερη μικρότερη κορυφή στο cepstrum, η οποία απουσιάζει πλήρως από το cepstrum της ανδρικής φωνής.