

DSP - Project 2

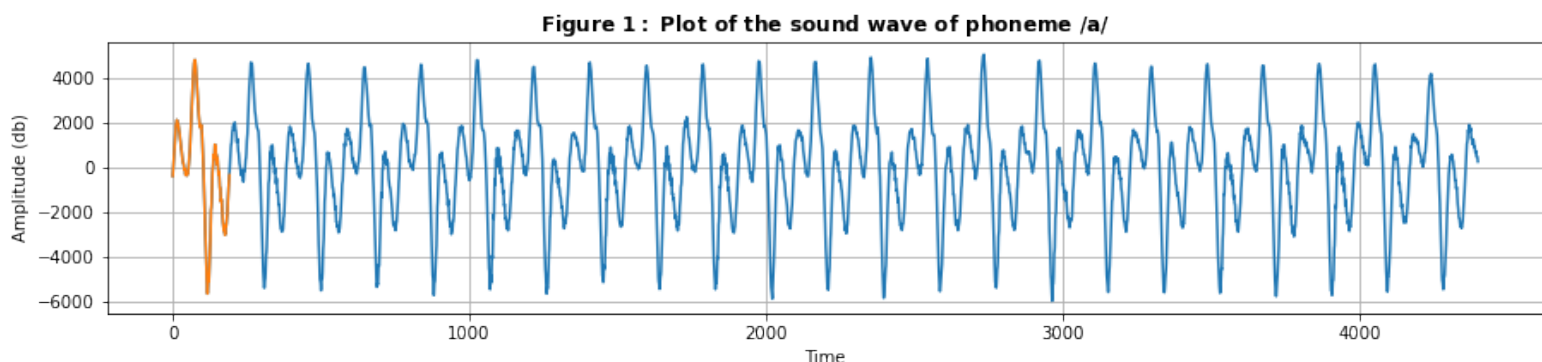
Αρχικά, χρησιμοποιώντας τα ηχητικά αρχεία της προηγούμενης άσκησης, **απομόνωσα** τμήμα **0.1 δευτερολέπτου** από το καθένα, ξεκινώντας κάθε φορά από τη μέση του αντίστοιχου σήματος (εκτός από την περίπτωση του φωνήματος / k /, όπως θα εξηγήσω παρακάτω).

Για να το υλοποιήσω, με βάση τη γνωστή συχνότητα δειγματοληψίας των παραπάνω σημάτων (44.1kHz), υπολόγισα το πλήθος των τιμών που καλύπτουν διάρκεια 0.1sec, ώστε να κρατήσω μόνο τόσες. Εάν λοιπόν η **συχνότητα δειγματοληψίας** είναι **44100Hz**, οι τιμές πλάτους που θα έχω σε 1 sec για το σήμα είναι 44100, άρα σε **0.1sec**, θα έχω 44100/10 τιμές, άρα **4410**. Έπειτα, δημιούργησα έναν νέο πίνακα με ακριβώς αυτές τις τιμές που αντιστοιχούν σε 0.1 δευτερόλεπτο του ηχητικού δείγματος, ξεκινώντας από τη μέση του προηγούμενου πίνακα που περιείχε όλες τις τιμές του σήματος, για όλη τη διάρκειά του.

Στη συνέχεια, έγινε η μέτρηση των κύκλων (με προσπάθεια για μικρό σφάλμα προσέγγισης) του κάθε σήματος με διακριτή περιοδικότητα (τα σήματα των ηχηρών φωνημάτων /a/, /o/ και /e/), τους οποίους και πολλαπλασίασα με το 10, ώστε να γίνει αναγωγή στο **1sec**, βρίσκοντας έτσι τη **συχνότητα** του σήματος. Η μέτρηση των κύκλων έγινε με 2 τρόπους · διαιρώντας το πλήθος των τιμών στο 0.1 sec (4410) με την περίοδο που έχω ήδη βρει από την προηγούμενη άσκηση για το εκάστοτε σήμα και έπειτα μετρώντας τους ολοκληρωμένους κύκλους στο διάγραμμά του.

Τέλος, ένωσα τα ισομήκη τμήματα που απομόνωσα στα παραπάνω βήματα με διάφορους τρόπους και με κάποιες επαναλήψεις αυτών και παραθέτω τις γραφικές τους παραστάσεις.

1. Phoneme / a /

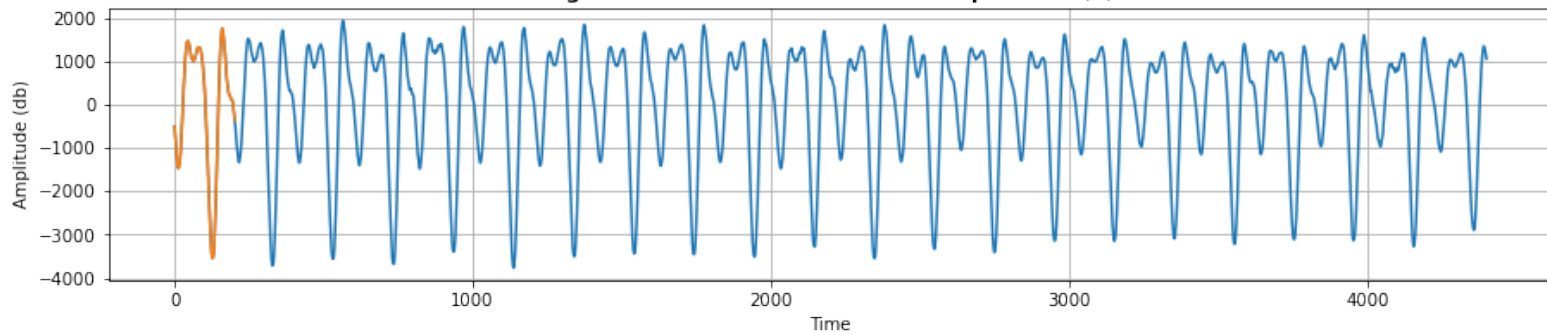


Sample rate: 44100Hz | **~Cycles in 100msec:** 23.037 | **~Frequency:** 230.37Hz ~ A₃/B₃ (233.08Hz)

Με τη διαδικασία που περιγράφηκε παραπάνω, παρατηρήθηκαν ~23.037 ολοκληρωμένοι κύκλοι στο διάγραμμα του σήματος (**Figure 1**). Συνεπώς, η συχνότητά του είναι ~230.37Hz, η οποία πλησιάζει τη συχνότητα 233.08Hz, δηλαδή εκείνη του ήχου της νότας A₃/B₃.

2. Phoneme / o /

Figure 2 : Plot of the sound wave of phoneme /o/

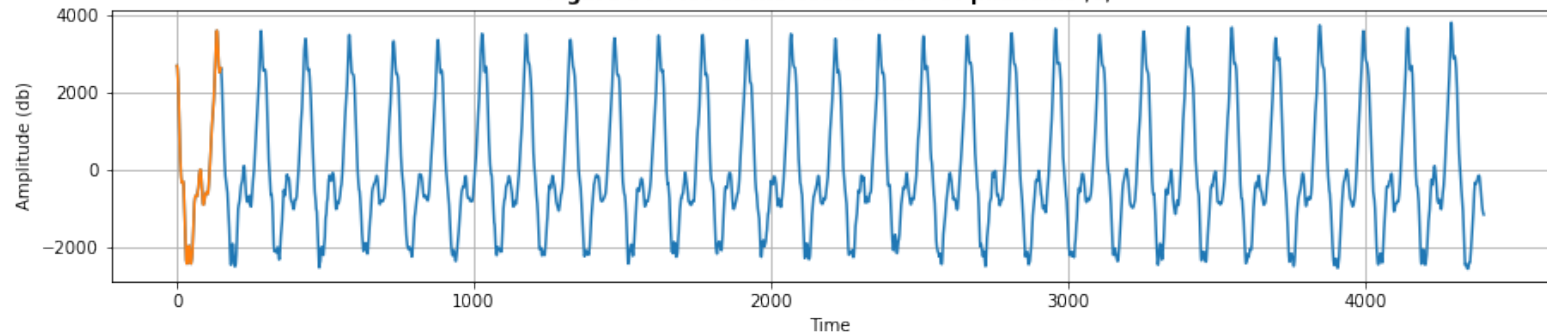


Sample rate: 44100Hz | **~Cycles in 100msec:** 21.675 | **~Frequency:** 216.75Hz ~ A₃ (220Hz)

Αντίστοιχα με την περίπτωση του φωνήματος /a/, έτσι κι εδώ, οι κύκλοι που μετρήθηκαν ήταν ~21.675 και η συχνότητα 216.75Hz, πλησιάζοντας αυτή τη φορά τη συχνότητα 220Hz, δηλαδή εκείνη του σήματος της νότας A₃.

3. Phoneme / e /

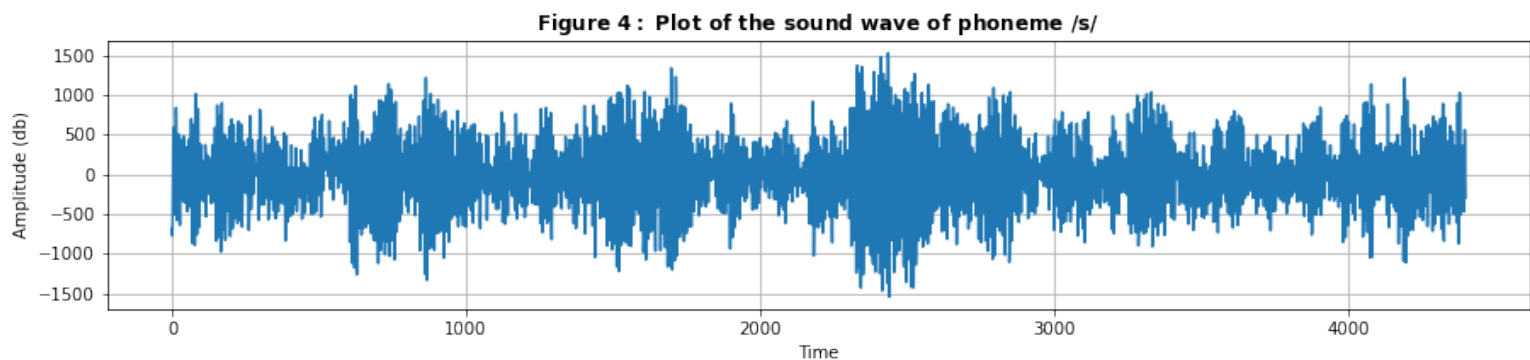
Figure 3 : Plot of the sound wave of phoneme /e/



Sample rate: 44100Hz | **~Cycles in 100msec:** 29.333 | **~Frequency:** 293.33Hz ~ D₄ (293.66Hz)

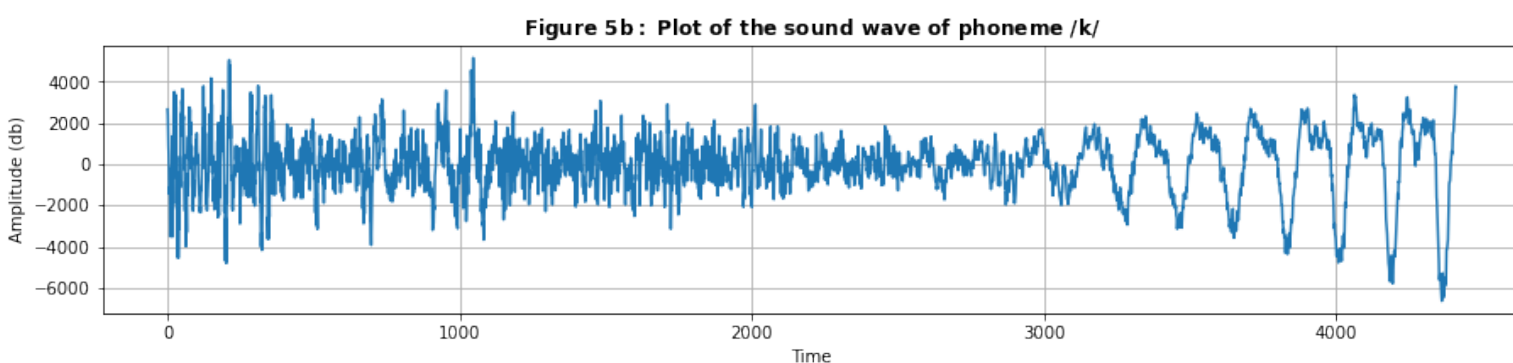
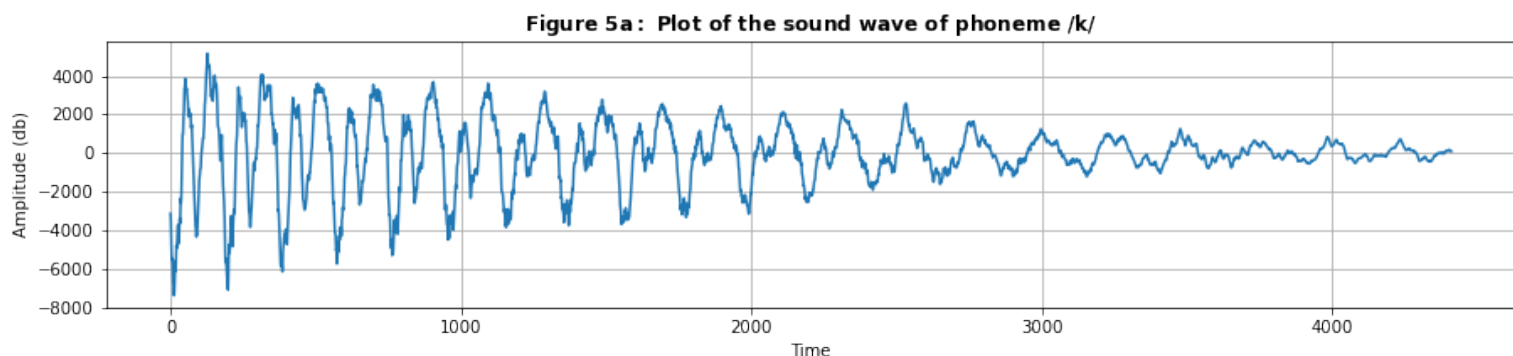
Όπως και στα παραπάνω, έτσι κι εδώ υπολογίστηκαν ~ 29.333 κύκλοι, άρα συχνότητα 293.33Hz, η οποία βρίσκεται κοντά στη συχνότητα της νότας D₄, που είναι 293.66Hz.

4. Phoneme / s /



Το **τριβόμενο** σύμφωνο /s/, όπως σχολιάστηκε και στην προηγούμενη άσκηση, πρόκειται για μη ηχηρό / άηχο (**voiceless**) φώνημα, το οποίο παράγεται με τέτοιον τρόπο και θέση/κινήσεις αρθρωτών, χωρίς τις δονήσεις των φωνητικών χορδών, με αποτέλεσμα να μην παράγεται περιοδικός ήχος, αλλά ένας **απεριοδικός θόρυβος**. Αυτή η φύση του ήχου του /s/, είναι εμφανής και από τη γραφική παράσταση του σήματός του, ακόμα και από ένα στιγμιότυπο διάρκειας 0.1 sec (**Figure 4**).

5. Phoneme / k /



Όσον αφορά το σύμφωνο /k/, το δείγμα των 100 msec το εξήγαγα όχι από τη μέση του ηχητικού αρχείου που είχα ηχογραφήσει, αλλά από προγενέστερο χρονικά σημείο, καθώς στην πρώτη περίπτωση (**Figure 5a**), το φώνημα βρίσκεται προς το τέλος της προφοράς του, με αποτέλεσμα να μην είναι χαρακτηριστικός ο ήχος του ώστε να χρησιμοποιηθεί στη συνέχεια για το μίξάρισμα των φωνημάτων. Γι' αυτόν ακριβώς το λόγο, αποκόπηκε από το αρχικό σήμα, το τμήμα του φωνήματος με γραφική παράσταση το **Figure 5b**. Όπως και στην περίπτωση του τριβόμενου /s/, έτσι και το **στιγμικό** /k/ είναι ένα **άηχο** φώνημα και όπως περιγράφηκε στην Άσκηση 1, το σήμα του οποίου πρόκειται για 2 σειριακές, ξαφνικές “εκρήξεις” ενέργειας από τους αρθρωτές (χωρίς τη δόνηση των φωνητικών χορδών όπως και στην προηγούμενη περίπτωση), με τη δεύτερη πιο έντονη από την πρώτη. Το σήμα αυτό είναι **μη περιοδικό**, εκτός από πολύ μικρές χρονικές περιόδους στις οποίες ίσως διακρίνονται σύντομα μοτίβα.

6. Mixed signal

Μετά από διάφορους συνδυασμούς μιξαρίσματος των σημάτων, αποφάσισα να δημιουργήσω την εξής πρόταση με τα παραπάνω φωνήματα:

“ Κακώς σε έσωσα Σάσα ”

Παρακάτω, παρουσιάζονται οι γραφικές παραστάσεις των λέξεων ξεχωριστά (**Figure 6**), καθώς και της πρότασης συνολικά (**Figure 7**).

Το **ηχητικό** αρχείο των λέξεων και της πρότασης, προφανώς αποτελεί μια προσπάθεια μεν, όχι πλήρως πετυχημένη δε, για τη **σύνθεση φωνής** και πιθανά μετατροπή κειμένου σε ομιλία με ένωση προηχογραφημένων φωνημάτων. Όπως είναι προφανές, χρειάζεται ειδική **ηχογράφηση**, **μίξη** πιθανώς διαφορετική ανάλογα με τις **συνθήκες** και **επεξεργασία** του αποτελέσματος αυτού, ώστε να παραχθεί **ομιλία** που να χαρακτηρίζεται από **καταληπτότητα**, **φυσικότητα** και να γίνεται τελικά **κατανοητή** η πληροφορία που μεταφέρεται.

Figure 6 : Plot of the sound waves of phonemes /a/, /o/, /s/, /e/ and /k/ concatenated in multiple order

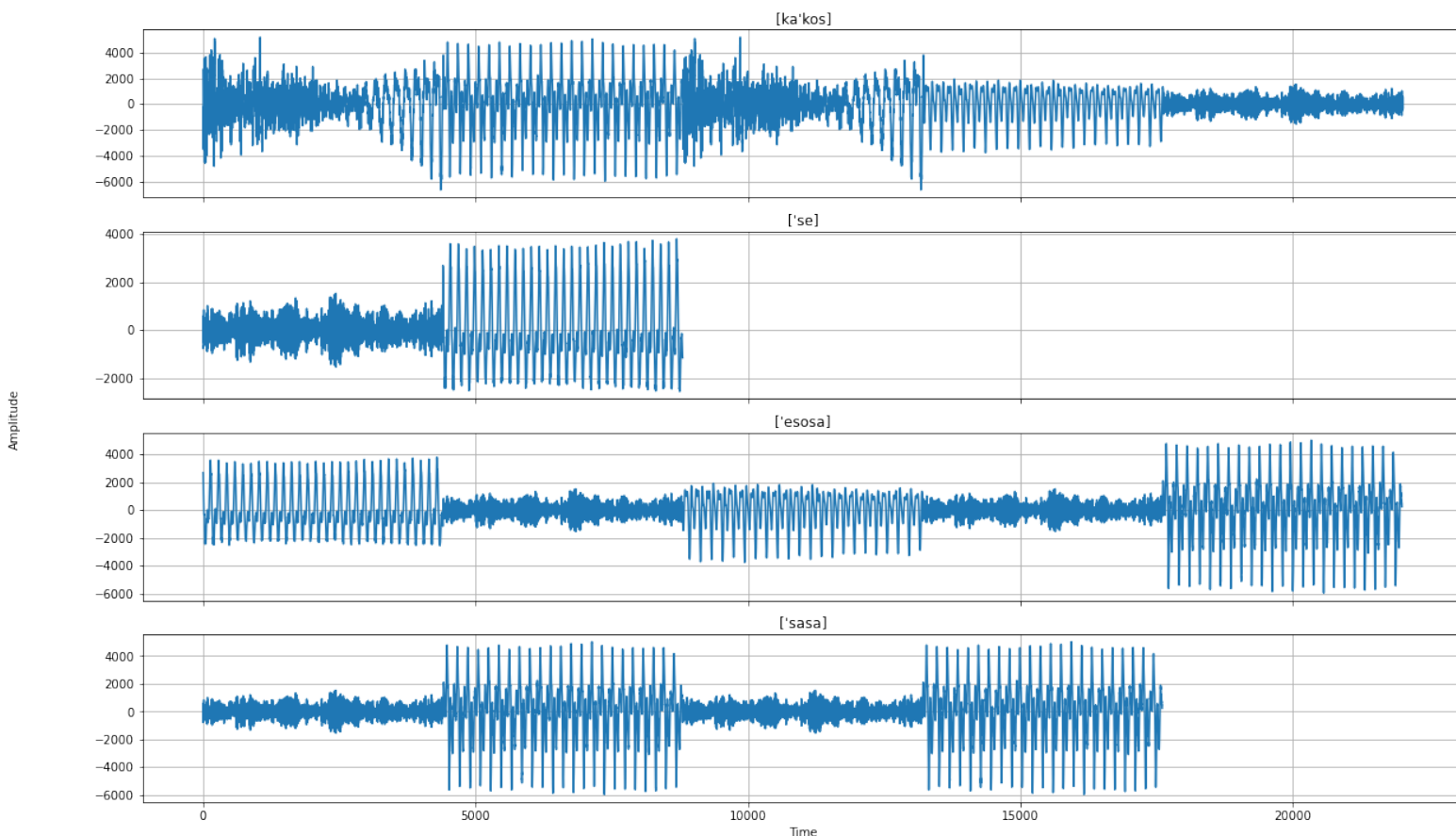


Figure 7 : [ka'kosse'esosa'sasa]

