***ΑΝΑΛΥΣΗ ΟΜΙΛΙΑΣ ΧΡΗΣΙΜΟΠΟΙΩΝΤΑΣ ΤΟ ΠΡΟΓΡΑΜΜΑ PRAAT***

***(Ένας σύντομος οδηγός από τον Pranav Jawale)***

1. **Εγκατάσταση του προγράμματος Praat:**

Οι χρήστες των Windows μπορούν να εγκαταστήσουν την τελευταία έκδοση του προγράμματος Praat από την ηλεκτρονική διεύθυνση <http://www.fon.hum.uva.nl/praat/download_win.html>

Κάντε διπλό κλικ στο αρχείο εγκατάστασης για να εγκαταστήσετε Praat. Αυτό θα δημιουργήσει ένα αρχείο praat.exe.

Οι χρήστες του Linux μπορούν να εγκαταστήσουν το πρόγραμμα Praat από την ηλεκτρονική διεύθυνση <http://www.fon.hum.uva.nl/praat/>

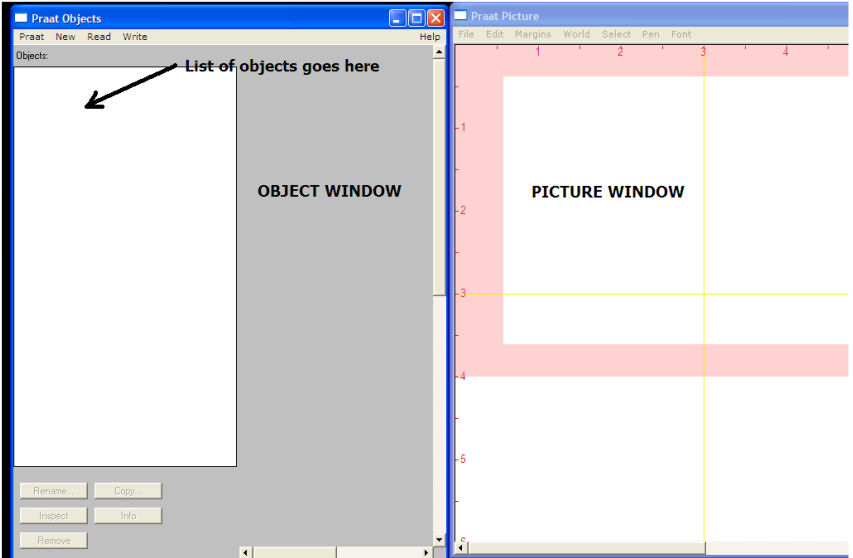
Είναι καλύτερα αν εκτελέσετε την Windows έκδοση του προγράμματος Praat κάτω από «wine» στα Linux.

**Πρόσθετες εγκαταστάσεις:**

Doulos SIL γραμματοσειρά: Έχει φωνητικά σύμβολα, μπορείτε να το κατεβάσετε από την ηλεκτρονική διεύθυνση <http://www.fon.hum.uva.nl/praat/DoulosSIL4.106.zip> ή από την <http://scripts.sil.org/cms/scripts/page.php?site_id=nrsi&item_id=DoulosSIL_download#FontsDownload>

Όταν ανοίγετε το πρόγραμμα Praat (κάνοντας δηλαδή διπλό κλικ στο αρχείο praat.exe) δύο παράθυρα θα εμφανιστούν, το Praat Objects και το Praat Image. Το παράθυρο Praat Objects περιέχει όλες τις επιλογές ανάλυσης.

Χρησιμοποιώντας το παράθυρο Praat Image, εικόνες με διάφορες πτυχές των αντικειμένων μπορούν να εξαχθούν. Προς το παρόν, το παράθυρο Praat Image δεν χρειάζεται, μπορείτε να το κλείσετε.

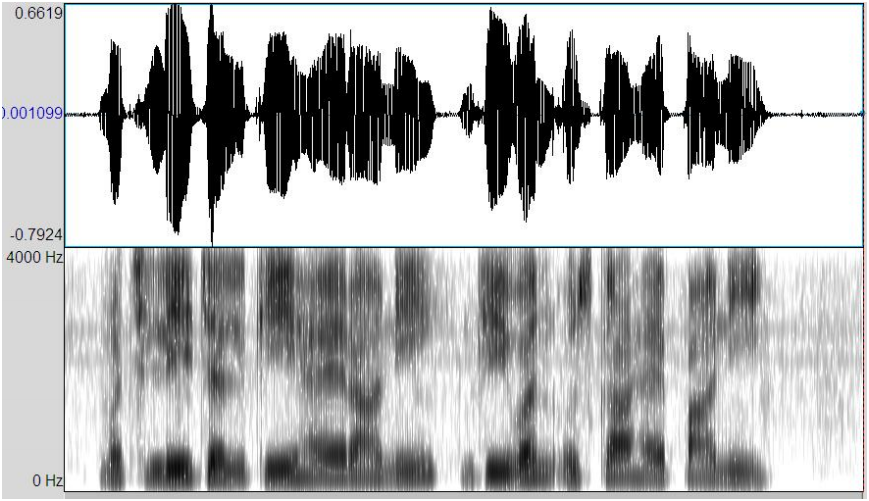


Σχήμα 1: τα παράθυρα Praat Object και Praat Picture.

Στο παράθυρο Praat Object, κάντε κλικ στο Read -> Read from file.

Τότε θα εμφανιστεί ένα παράθυρο επιλογής αρχείων. Αναζητείστε, εντοπίστε και επιλέξτε το αρχείο ήχου «machali.wav».

Κάντε κλικ στο κουμπί Open. Τονίστε το όνομά του στη λίστα των αντικειμένων και κάντε κλικ στο Edit.



Σχήμα 2: Κυματομορφή και φασματογράφημα του «machali.wav».

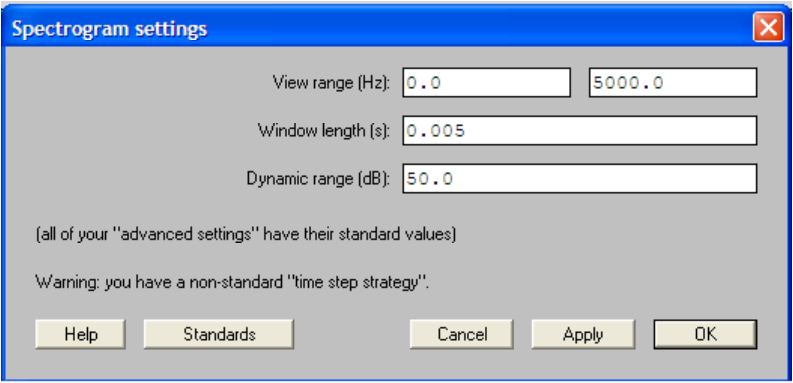
Παίξτε το αρχείο. Μπορείτε να καταλάβετε τι ακούτε;

Εστιάστε σε μικρότερα τμήματα της κυματομορφής και παρατηρήστε τις ακουστικές ιδιότητες της σε διάφορα στοιχεία της όπως είναι: η περιοδικότητα, η ενέργεια, το μέσο όριο μηδενισμού της κυματομορφής, τα formants-φωνοσυχνότητες (ιδιοσυχνότητες της φωνητικής κοιλότητας).

1. **Διάφορες ρυθμίσεις του φασματογραφήματος στο προόγραμμα Praat**

Τώρα θα εξετάσουμε τις ρυθμίσεις του φασματογραφήματος.

Κάντε κλικ στο Spectrum -> Spectrogram settings.

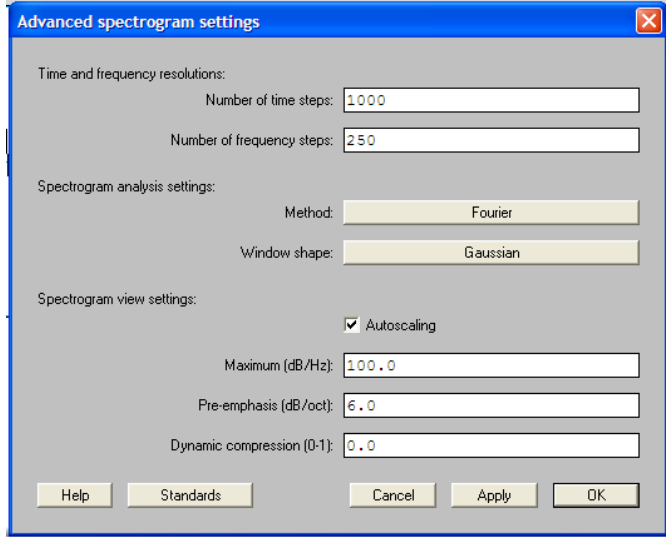


Σχήμα 3: Προεπιλεγμένες ρυθμίσεις του φασματογραφήματος.

**Επεξήγηση των παραμέτρων (από το εισαγωγικό εγχειρίδιο λειτουργίας του προγράμματος Praat[1])**

* Εμφάνιση της σειράς (Hz) - περιοχή στην οποία εμφανίζεται το φασματογράφημα. Για να δείτε το φασματόγραμμα σε όλο του το συχνοτικό φάσμα, η μέγιστη συχνότητα θα πρέπει να είναι το η μισή του ρυθμού δειγματοληψίας (8000/2 = 4000 Hz). Το φασματογράφημα θα είναι κενό για συχνότητες υψηλότερες της συχνότητας Nyquist, εφόσον εμφανίζονται.
* Μήκος παραθύρου (s) - για να υπολογίσουμε το μετασχηματισμό Fourier σε σύντομο χρονικό διάστημα το πρόγραμμα Praat χρησιμοποιεί ανάλυση παραθύρων (κουφώματα) του οποίου το κάθε μήκος ορίζεται από την παράμετρο του μήκους του παραθύρου. Για ένα μήκος παραθύρου 0.005 s, το Praat χρησιμοποιεί για κάθε μέρος του τμήματος τον ήχο που βρίσκεται μεταξύ 0,0025 δευτερόλεπτα πριν και 0,0025 δευτερόλεπτα μετά από τη μέση του τμήματος αυτού.
* Δυναμικό εύρος (dB) - όλες οι τιμές που είναι περισσότερες από το δυναμικό εύρος σε dB κάτω από το ανώτατο θα εμφανιστεί λευκό. Ενδιάμεσες τιμές διαθέτουν κατάλληλες αποχρώσεις του γκρι. Έτσι, αν η υψηλότερη κορυφή στο φασματογράφημα έχει τονικό ύψος 30 dB/Hz και το δυναμικό εύρος είναι 50 dB (που είναι η τυπική τιμή), τότε οι τιμές κάτω από τα -20 dB / Hz θα γίνουν λευκές και οι τιμές μεταξύ των -20 dB / Hz και των 30 dB / Hz θα πάρουν διάφορες αποχρώσεις του γκρι.

Τώρα κάντε κλικ στο Spectrum -> Advanced spectrogram settings.



Σχήμα 4: Προεπιλεγμένες προχωρημένες ρυθμίσεις του φασματογραφήματος.

Εξήγηση μερικών παραμέτρων (από το εγχειρίδιο intro Praat [2])

* Αριθμός των βημάτων του χρόνου - ο μέγιστος αριθμός των σημείων κατά μήκος παραθύρου του χρόνου για τα οποία το Praat πρέπει να υπολογίσει το φάσμα. Αν η οθόνη σας δεν είναι μεγαλύτερη από 1200 pixels, τότε το πρότυπο των 1000 είναι το κατάλληλο, δεδομένου ότι δεν υπάρχει κανένα πρόβλημα στο να προγραμματιστούν παραπάνω από ένα φάσματα ανά ένα pixel στην κάθετη γραμμή.
* Αριθμός των βημάτων της συχνότητας - ο μέγιστος αριθμός των σημείων κατά μήκος του άξονα της συχνότητας για τον οποίο το Praat πρέπει να υπολογίσει το φάσμα. Αν η οθόνη σας δεν είναι υψηλότερη από 768 pixels, τότε το πρότυπο του 250 είναι το κατάλληλο, δεδομένου ότι δεν υπάρχει κανένα πρόβλημα στο να προγραμματιστούν παραπάνω από ένα φάσματα ανά ένα pixel στην οριζόντια γραμμή.
* Το σχήμα του παραθύρου – η προεπιλογή είναι το Gaussian παράθυρο. Μπορείτε να το αλλάξετε σε Rectangular, Hamming, Hanning κ.λπ. (βλέπε [1] στο πλαίσιο για περαιτέρω ανάγνωση).

1. **Μεγάλου και μικρού εύρους φασματογραφήματα**

Τα φασματογραφήματα της ομιλίας ονομάζονται και ευρείας ζώνης ή στενής ζώνης βάση του μήκους του παραθύρου που χρησιμοποιούνται. Όσο μικρότερο είναι το παράθυρο, τόσο μεγαλύτερο είναι το εύρος ζώνης και το φασματογράφημα είναι ευρείας ζώνης. Ομοίως, αν το παράθυρο είναι μεγαλύτερο, το φασματογράφημα λέγεται στενής ζώνης.

Τώρα, πόσο μήκος παραθύρου αντιστοιχεί σε καθένα από αυτά; Δεν υπάρχει ξεκάθαρη όριο αλλά για ένα δείγμα ομιλίας, αν το μήκος του παραθύρου είναι περίπου 3 ως 5 ms (εύρος ζώνης ~ 200-300 Hz), το φασματογράφημα που προκύπτει ονομάζεται ευρείας ζώνης. Για το μήκος του παραθύρου που είναι περίπου 20 ως 30 ms (εύρος ζώνης περίπου 30 - 50 Hz), το φασματογράφημα ονομάζεται στενής ζώνης.

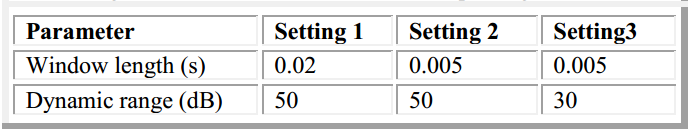
Ποια είναι η σημασία των δύο αυτών τύπων φασματογραφημάτων;

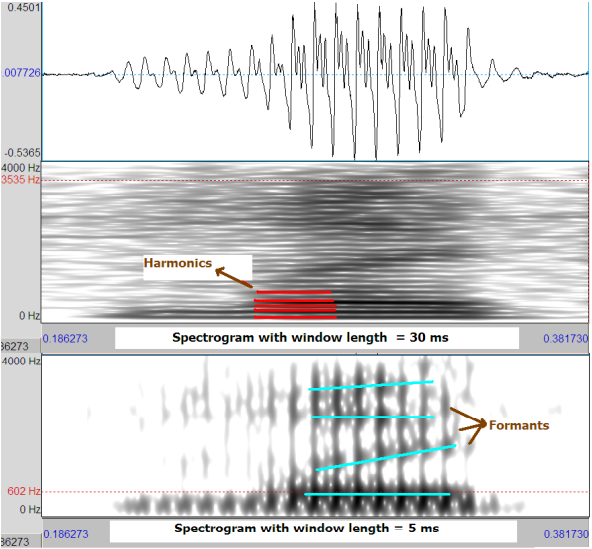
Τα ευρείας ζώνης φασματογραφήματα χρησιμοποιούνται για να παρατηρήσουμε τη δομή των formants ενώ τα στενής ζώνης φασματογραφήματα αποκαλύπτουν την αρμονική δομή (πληροφορίες για την συχνότητα-τονικό ύψος). Ας πούμε ότι χρησιμοποιείτε ένα παράθυρο μήκους 20 ms. Με αυτό το μήκος του παραθύρου, οι φωνητικές αρμονικές επιλυθούν πολύ πιο ξεκάθαρα ως ανάλυση συχνότητας μεταξύ τους είναι γενικά> 50 Hz (1 / 0.02). Αν το παράθυρο είναι μικρότερο από τα 3 ms για παράδειγμα, τότε τα φάσματα δύο ή τριών αρμονικών φωνών θα συγχωνευτούν και αντί των επιμέρους αρμονικών θα βλέπετε τη δομή των formants στο φασματογράφημα.

Ως παράδειγμα, μεγεθύνετε ένα τμήμα του machali.wav που να απεικονίζει το φωνήεν / m /. Παίξτε αυτό.

Κρατήστε τις προχωρημένες ρυθμίσεις του φασματογραφήματος στις προεπιλεγμένες τιμές τους (κάνοντας κλικ στο Standards button) και παρατηρήστε την κυματομορφή και το φασματογράφημα για αυτές τις δύο ρυθμίσεις. Σημείωση: μεταβάλλοντας το δυναμικό εύρος αλλάζουν μόνο οι ρυθμίσεις της έντασης του φασματογραφήματος. Για να δείτε καθαρά τα formants, το έχουμε μειώσει στα 30 dB.

Πίνακας 1: μήκη παραθύρου για στενής και ευρείας ζώνης φασματογραφήματα.



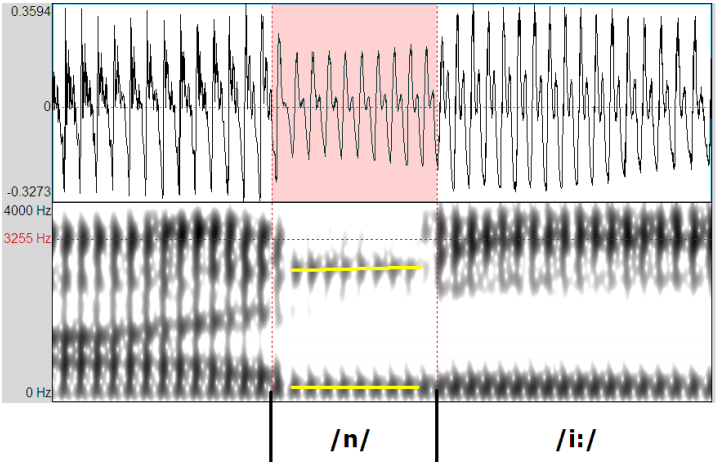


Σχήμα 5: το φασματογράφημα του / m / στο «Machali» για δύο διαφορετικά μήκη παραθύρων (30 ms και 5 ms).

Στο Σχήμα 9, όταν το μήκος του παραθύρου είναι 30 ms, μπορούμε να δούμε πολλές οριζόντιες μαύρες γραμμές στο φασματογράφημα που αντιστοιχούν στις αρμονικές των τονικών υψών-συχνοτήτων. Καθώς το παράθυρο μήκος γίνεται 5 ms, δεν μπορούμε πλέον να διακρίνουμε τίποτα μεταξύ των διαφόρων αρμονικών, αλλά μπορούμε να δούμε τα formants. Οι κάθετες σκοτεινές γραμμές στο φασματογράφημα δείχνουν γλωττιδικούς παλμούς.

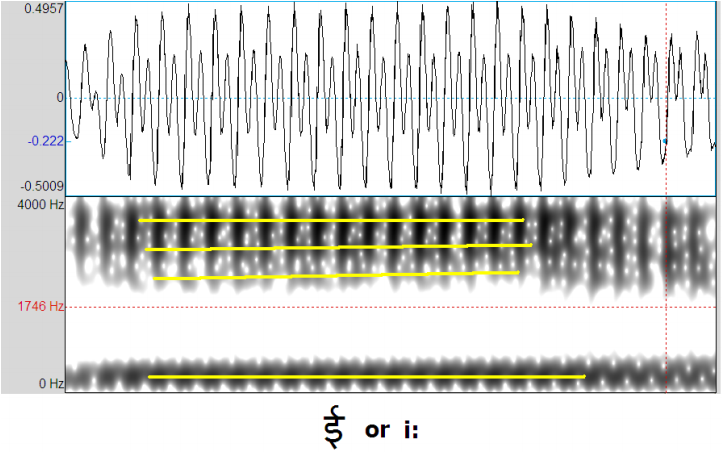
Τώρα κάντε σμίκρυνση (επιλέξτε τα όλα) και παρατηρήστε ολόκληρο το φασματογράφημα και για τις δύο ρυθμίσεις στον Πίνακα 4.

1. **Ανάλυση διαφορετικών φωνημάτων**
   1. **Η παρατήρηση των formants**



Σχήμα 6: η κυματομορφή και το ευρείας ζώνης φασματογράφημα του / N / στη «paani». Το μέγεθος του παραθύρου είναι 5 ms.

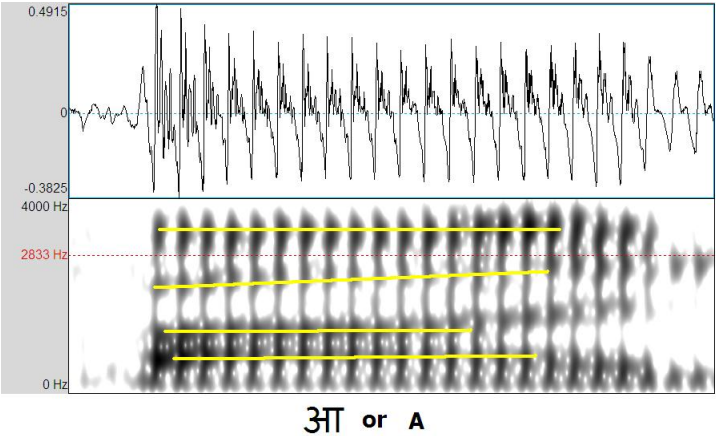
Κάντε μεγέθυνση στο μέρος της κυματομορφής που βρίσκεται το / N / στην «paani». Χαρακτηρίζεται από απότομες ασυνέχειες στην αρχή και στο τέλος του φάσματος. Σημειώστε την απώλεια της έντασης στο φασματογράφημα λόγω των αντηχήσεων. Τα Nasals είναι κυματομορφές διαφοροποιημένες από τα κατώτερα πλάτη τους (συγκρινόμενες με τα φωνήεντα).



Σχήμα 7: η κυματομορφή και το ευρείας ζώνης φασματογράφημα του / i: / στο «paani».

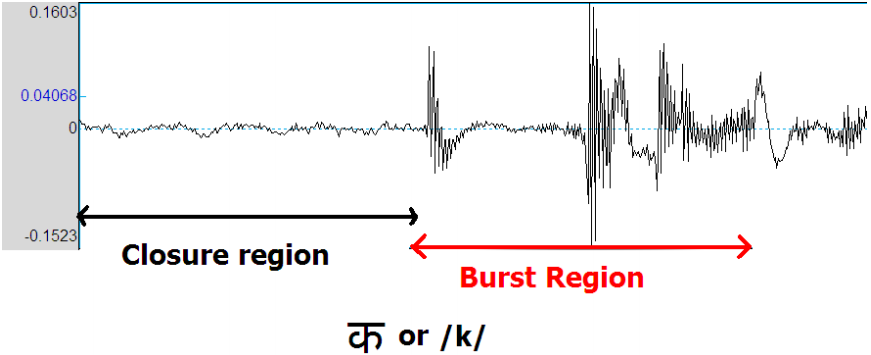
Τώρα παρατηρούμε το / i: / στο «jal ki». Παρατηρήστε τα διαφορετικά formants. Ποιες είναι οι συχνότητες τους;

Το φωνήεν, / i: /, έχει τη μέγιστη απόσταση μεταξύ των F1 και F2 όλων των φωνηέντων.



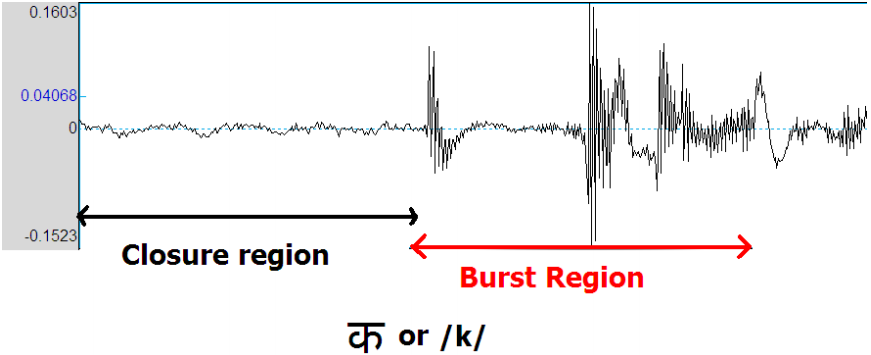
Σχήμα 8: η κυματομορφή και το ευρείας ζώνης φασματογράφημα του Α στο «paani».

Τώρα κοιτάξτε το φασματογράφημα του / Α / στο «paani». Πού βρίσκονται τα τρία πρώτα formants της;



Σχήμα 9: το τέλος και οι δυνατές-έντονες περιοχές του / k / στο «jal ki».

Ας ρίξουμε μια ματιά στην κυματομορφή του / k / στο «jal ki». Μπορούμε το να χωρίσουμε σε δύο περιοχές. Πρώτα το κλείσιμο και μετά η έντονη περιοχή. Σε αυτό το παράδειγμα μπορούμε να δούμε δύο εντάσεις μαζί με κάποια ήρεμα σημεία στην περιοχή της έντασης.



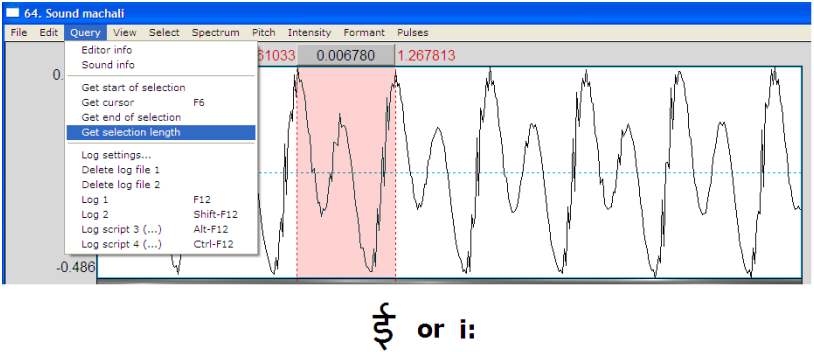
Σχήμα 10: Το τέλος και οι δυνατές-έντονες περιοχές του / p / στο «paani».

Κάντε μεγέθυνση στη κυματομορφή του / p / στο «paani». Όπως φαίνεται στο Σχήμα 10, το / p / έχει μια πολύ σύντομη αύξηση του πλάτους του.

**5.2 Βρίσκοντας το Pitch**

Τώρα θα δούμε πώς να καθορίζετε το τονικό ύψος-συχνότητα από την κυματομορφή και από το φασματογράφημα. Πρώτα ας μάθουμε το τονικό ύψος του / i: / στο «paani». Παρατηρήστε την περιοδικότητα της κυματομορφής του.

Υπολογίστε το ύψος του βρίσκοντας το αντίστροφο του διαστήματος της διάρκειας μετά το οποίο η κυματομορφή επαναλαμβάνεται. Μπορείτε να χρησιμοποιήσετε τη λειτουργία Get selection length για να πάρετε ακριβώς την χρονική διάρκεια της κάθε επιλογής (βλέπε Σχήμα 15).

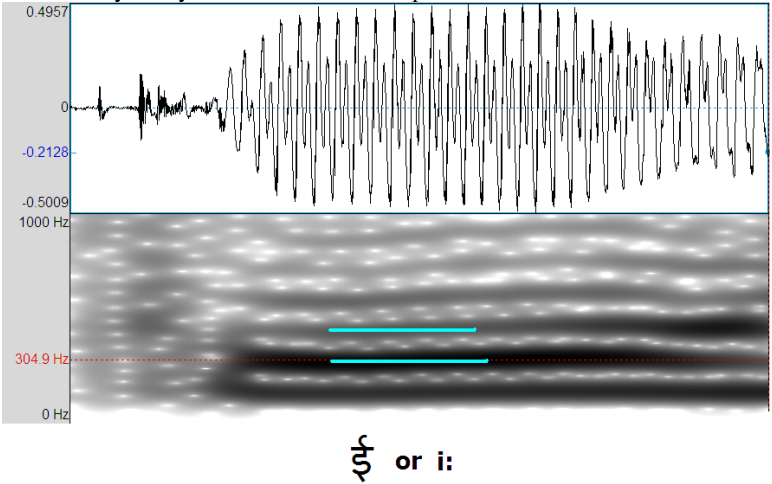


Σχήμα 11: Εύρεση του μήκους της επιλογής στο τμήμα της κυματομορφής που βρίσκεται το / i: / στο «paani».

Τι αξία βρήκατε ότι έχει το τονικό ύψος του (δηλαδή η συχνότητά του);

Ας μην αλλάξουμε τις ρυθμίσεις του φασματογραφήματος έτσι ώστε να δείχνει τις αρμονικές. Αλλάξτε το μήκος του παραθύρου στα 0,03 s. Η διαφορά σε Hz μεταξύ των γειτονικών αρμονικών είναι το ύψος. Μπορείτε να βρείτε μια τιμή κατά προσέγγιση του ύψους με τη χρήση αυτής της μεθόδου. Είναι η ίδια όμως με αυτή που βρήκατε από την κυματομορφή ανάλυση;

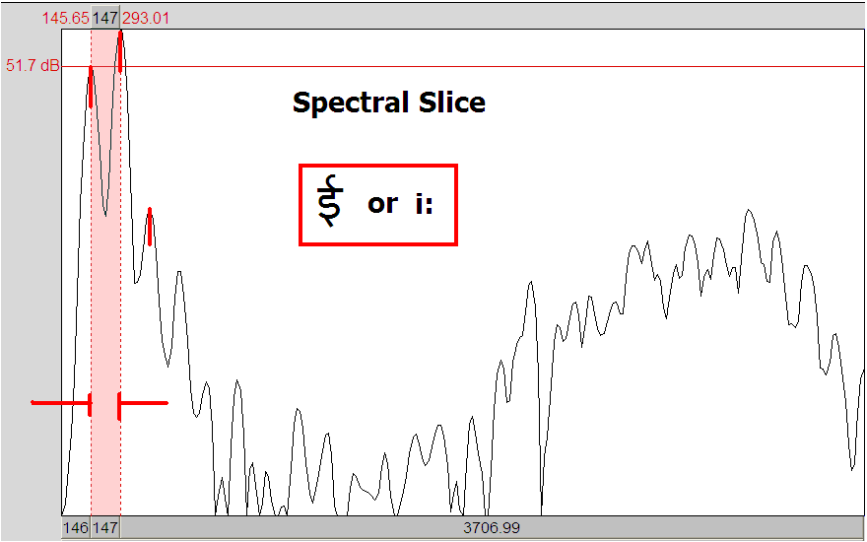
Παρόλα αυτά, σημειώστε ότι το τονικό ύψος είναι ίσο με τη διαφορά σε Hz μεταξύ των αρμονικών μόνο όταν όλες οι αρμονικές είναι παρούσες στο δείγμα ήχου. Αν λείπουν ορισμένες αρμονικές, θα υπολογίσετε λάθος τιμή για το τονικό ύψος. Η ανάλυση των κυματομορφών μπορεί σχεδόν πάντα να σας δίνει την σωστή αξία του συχνοτικού εύρους.



Σχήμα 12: Η κυματομορφής και το στενό φασματογράφημα του / i: / στο «paani».

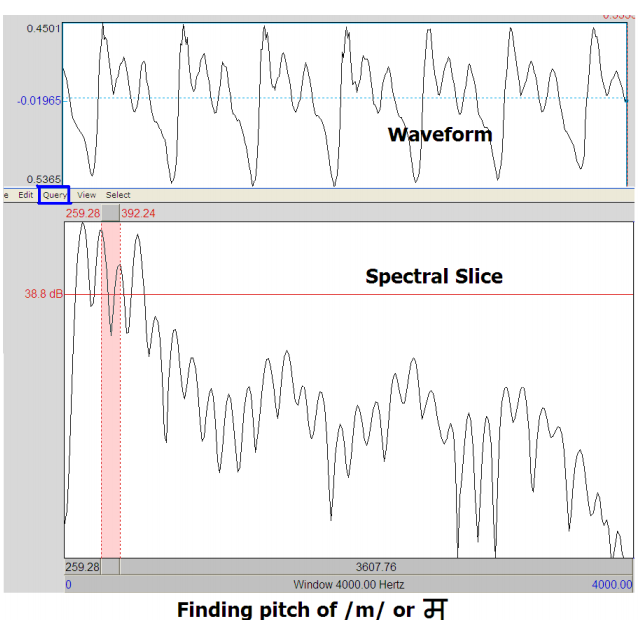
Ένας πιο ακριβής τρόπος για να βρεθεί το τονικό ύψος από το φάσμα: Επιλέξτε δύο/τρεις περιόδους της κυματομορφής και κάντε κλικ στο πλήκτρο Ctrl + L (ή πηγαίνετε στο Spectrum -> View spectral slice). Αυτό δείχνει ένα μέγεθος αντίθετο του συχνοτικού εύρους που εμφανίζεται μετά τον υπολογισμό FFT.

Η διαφορά σε Hz ανάμεσα σε κάποιες υψηλής ενέργειας γειτονικές κορυφές θα σας δώσει την αξία του τονικού ύψους.



Σχήμα 13: Το φασματικό κομμάτι του / i: / στο «paani» χρησιμοποιώντας το φασματικό εργαλείο κοπής.

Τώρα χρησιμοποιήστε την κυματομορφή και φασματικές μεθόδους κοπής για να μάθετε το τονικό ύψος του / m / στο «machali» (Σχήμα 18).



Σχήμα 14: Η κυματομορφή και το φασματικό κομμάτι του / m / στο «machali».

Μετρήστε το τονικό ύψος σε διάφορα σημεία των τμημάτων της φωνής στη φράση. Δείτε πώς μεταβάλλεται σε συνάρτηση με τον χρόνο σε όλη την έκφραση.

**Αναφορές:**

1. <http://www.fon.hum.uva.nl/praat/manual/Intro_3_2__Configuring_the_spectrogram.html> από το εισαγωγικό εγχειρίδιο λειτουργίας του Praat που είναι διαθέσιμο στην ιστοσελίδα <http://www.fon.hum.uva.nl/praat/manual/Intro.html>
2. <http://www.fon.hum.uva.nl/praat/manual/Advanced_spectrogram_settings___.html>
3. Ομάδα μελών Praat -http://uk.dir.groups.yahoo.com/group/praatusers/message/2431

Περαιτέρω ανάγνωση:

1. Harris, FJ "On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform." Proceedings of the IEEE. Vol. 66, No. 1 (January 1978)
2. Rob Hagiwara, How to read spectrograms? <http://home.cc.umanitoba.ca/~robh/howto.html>