DSP HW2-2
Speech Analysis

Outline

1. Introduction
2. Praat
3. Homework Problems
4. Submission Requirements

Introduction

- Analyze speech signal from spectrogram
- Try to distinguish different initials(聲母) and finals(韻母) on spectrogram.
- Right-Context-Dependent Initial Final (RCDIF)

  t_i for 去 followed by finals starting with ㄧ

  ex 1：去ㄧ = t_i i

  ex 2：去ㄚ = t_a a

Introduction

- classification of consonants

Plosive/Stop	爆破音/塞音	ㄅㄆㄉㄊㄍㄎ
Fricative	擦音	ㄈㄏㄒㄕㄙ
Affricate	塞擦音	ㄐㄑㄓㄔㄗㄘ
Nasal	鼻音	ㄇㄋ

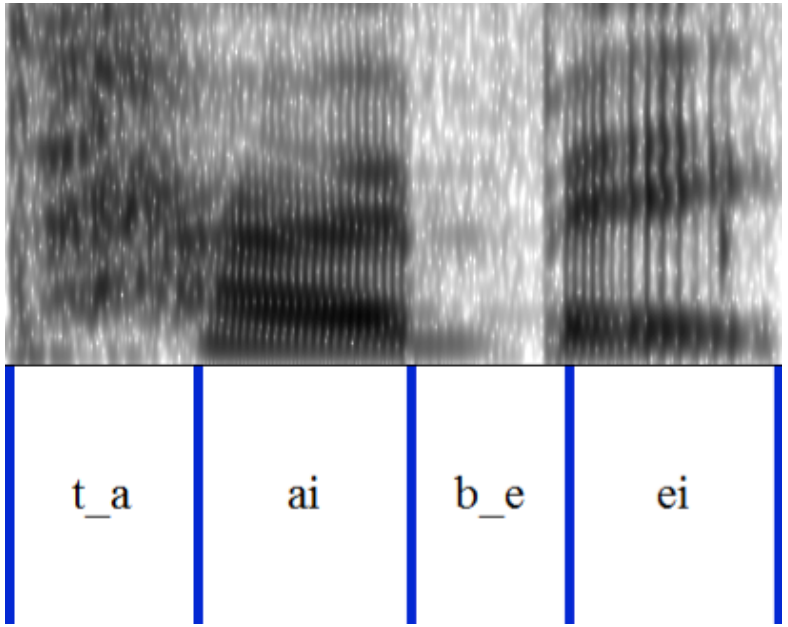
- classification of vowels

Monophthong	單母音	ㄧㄨㄩㄚㄛㄜㄦ
Diphthong	雙母音	ㄞㄟㄠㄡ

Introduction

Some useful information about labeling.

- ***"sil"*** for silence.
- ***"sp"*** for short pause.
- fricative/affricate initials do not contain voicing parts.
- plosive initials contain closure or aspiration period.



Some files you need

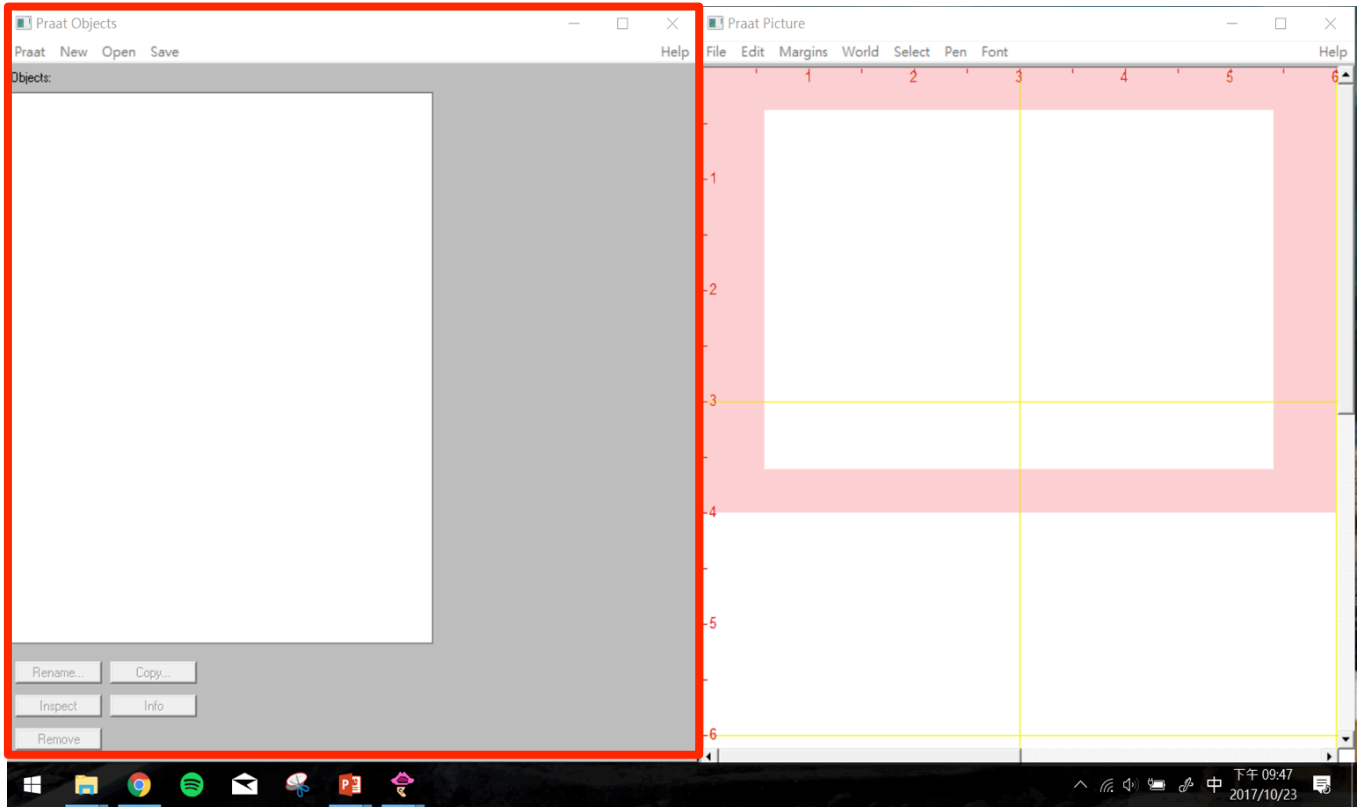
1. Phonetic class table (聲韻母表):
   http://speech.ee.ntu.edu.tw/homework/DSP_HW2-2/phonetic_class.pdf
2. Syllable table (標註模式):
   http://speech.ee.ntu.edu.tw/homework/DSP_HW2-2/syllable.txt
3. Audio data & FAQ:

   http://speech.ee.ntu.edu.tw/homework/DSP_HW2-2/

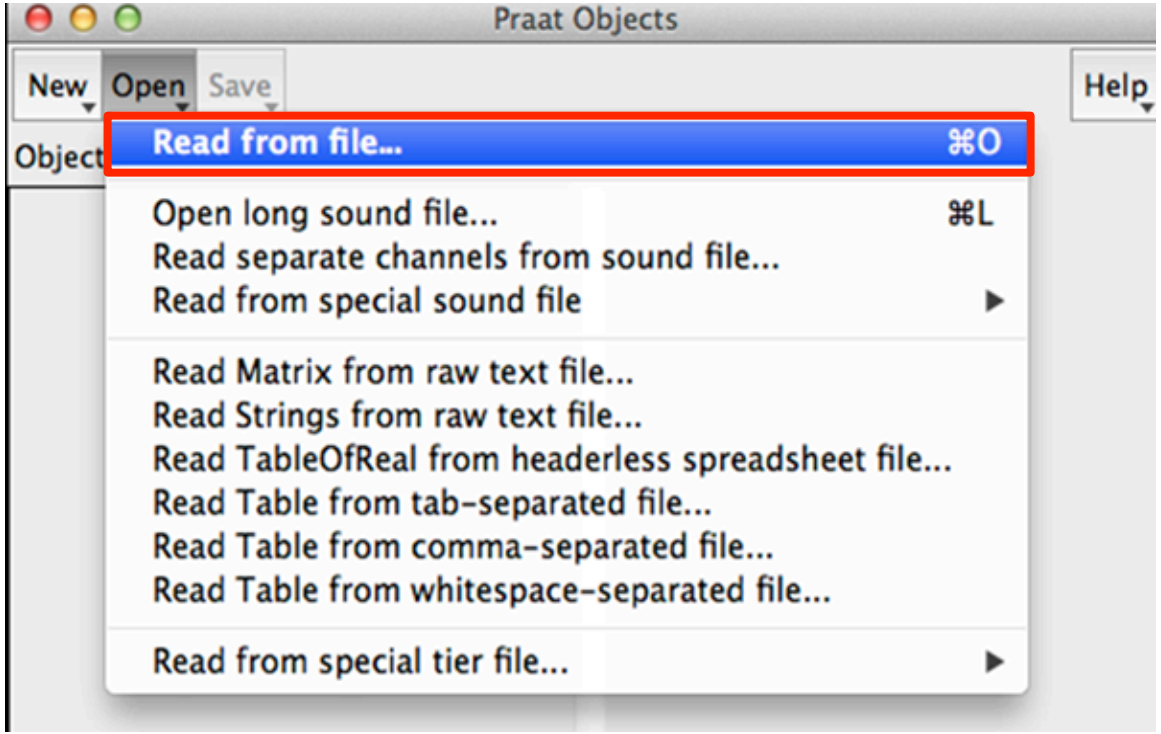
Praat

1. Download
   http://www.fon.hum.uva.nl/praat/
2. How to read a wave file
3. How to use it
4. How to label

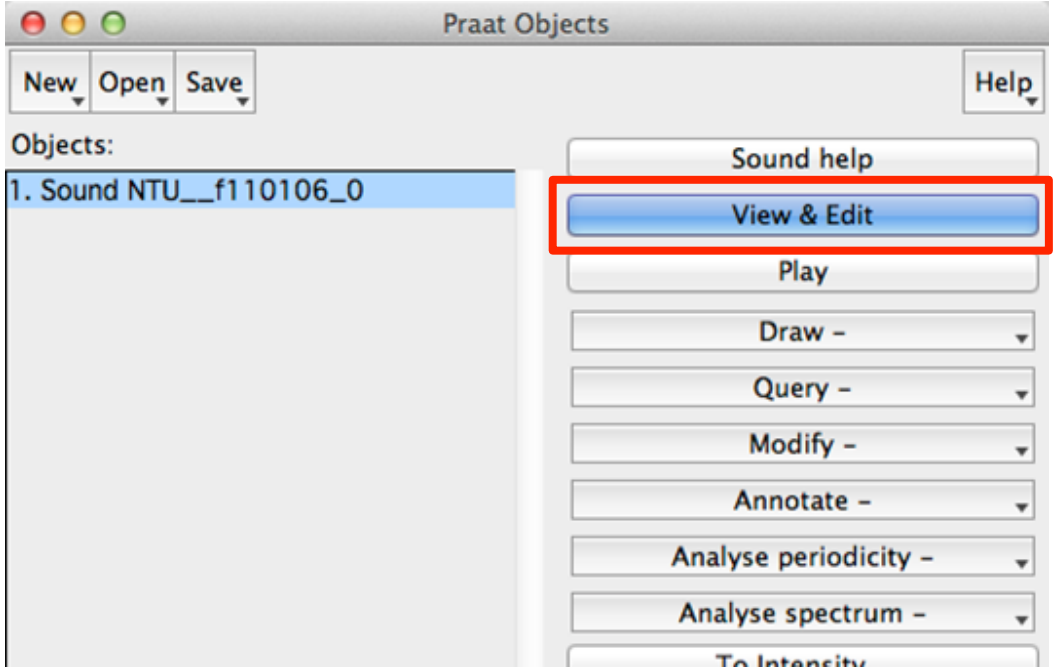
Praat


Praat Objects
Praat New Open Save Help
Objects:
Rename...
Copy...
Inspect
Info
Remove
Praat Picture
File Edit Margins World Select Pen Font Help
1 2 3 4 5 6
1 2 3 4 5 6
下午 09:47
2017/10/23

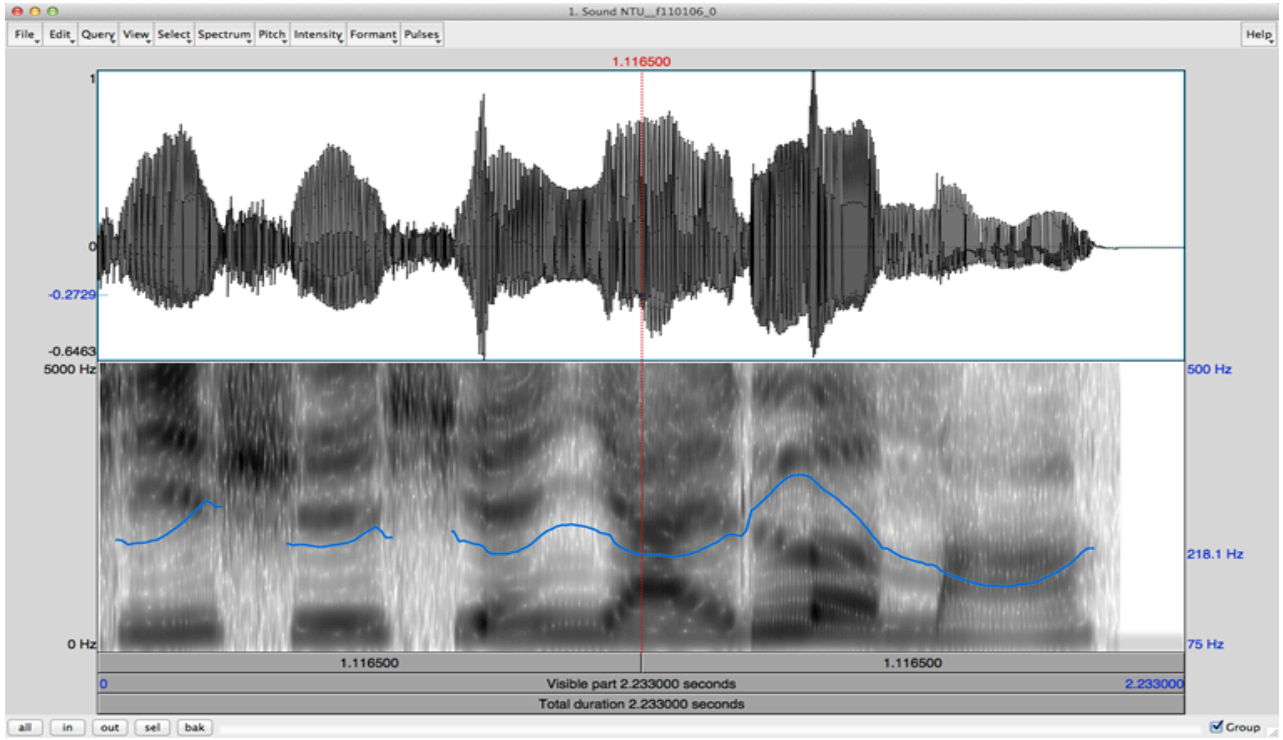
Praat - Read from file (.wav file)



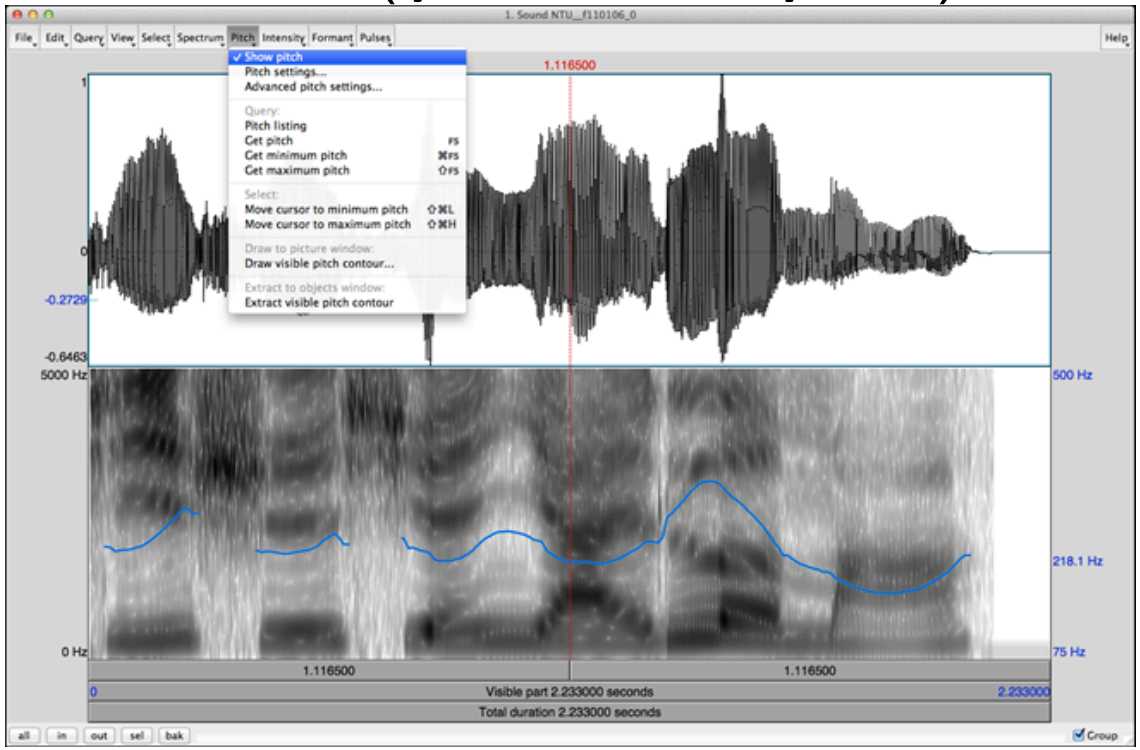
Praat - click ***View & Edit***



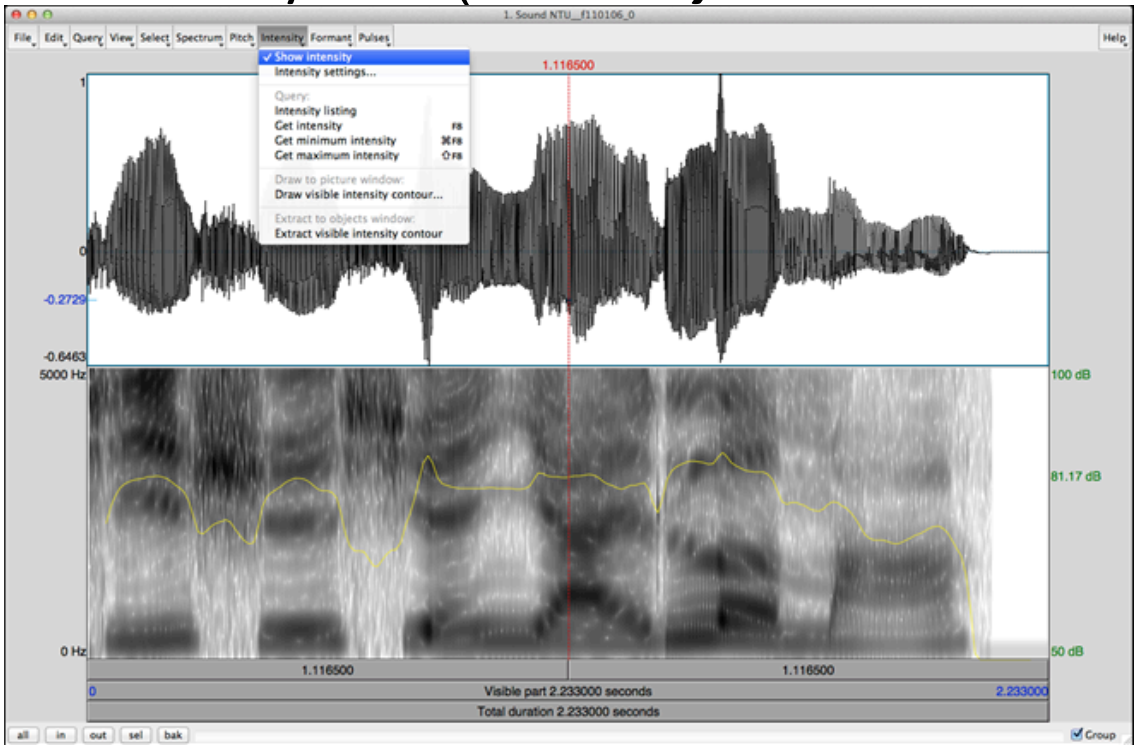
Praat - Time and Frequency Domain



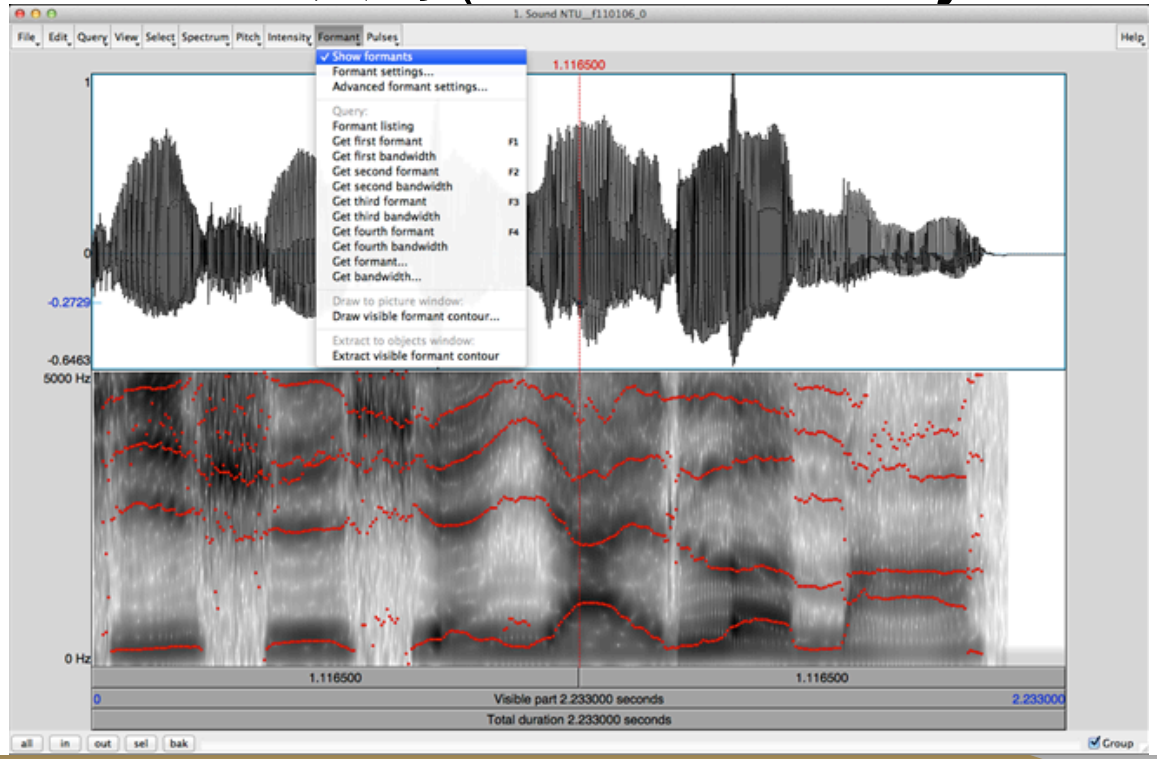
Praat - Pitch 音高（***pitch*** -> ***Show pitch***）



Praat - Intensity 音量(***Intensity ->Show Intensity)***



Praat - Formant 共鳴 (***Formant*** -> ***Show formants***)



Praat - Reminder

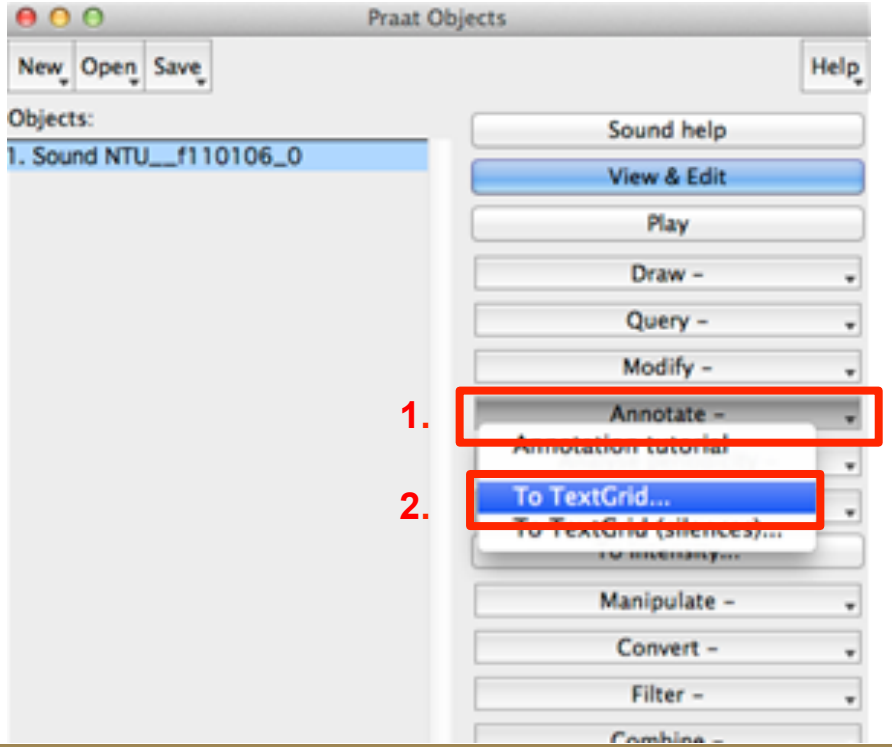
1. ***Intensity:*** power of all frequency components

   Two acoustic signals may have the same intensity but different frequency components.

2. ***Formant:*** acoustic resonance, measured by the peak in the frequency spectrum

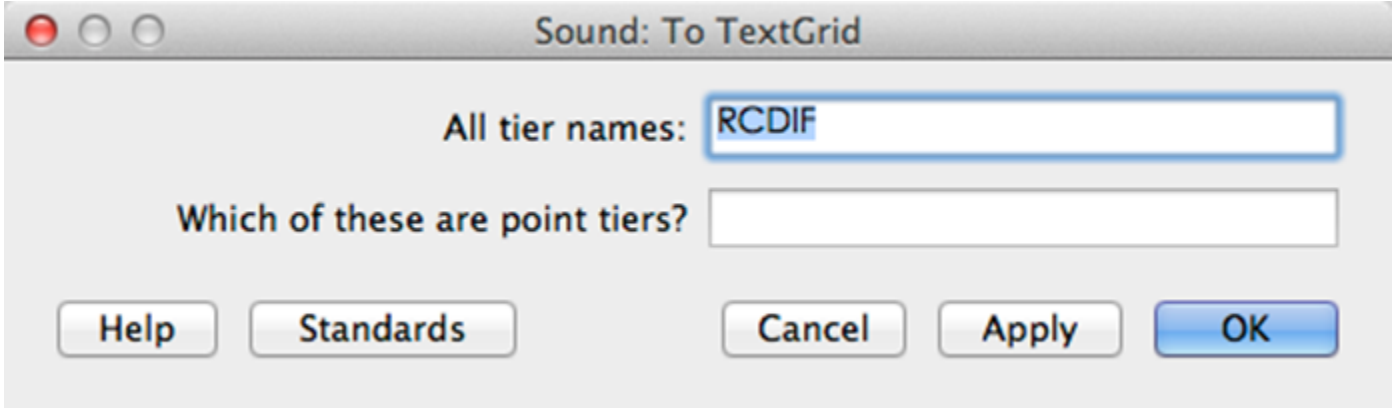
   You should not trust the formant detection output for unvoiced initials.

Praat - Label a wave file (***Annotate*** -> ***To TestGrid***)



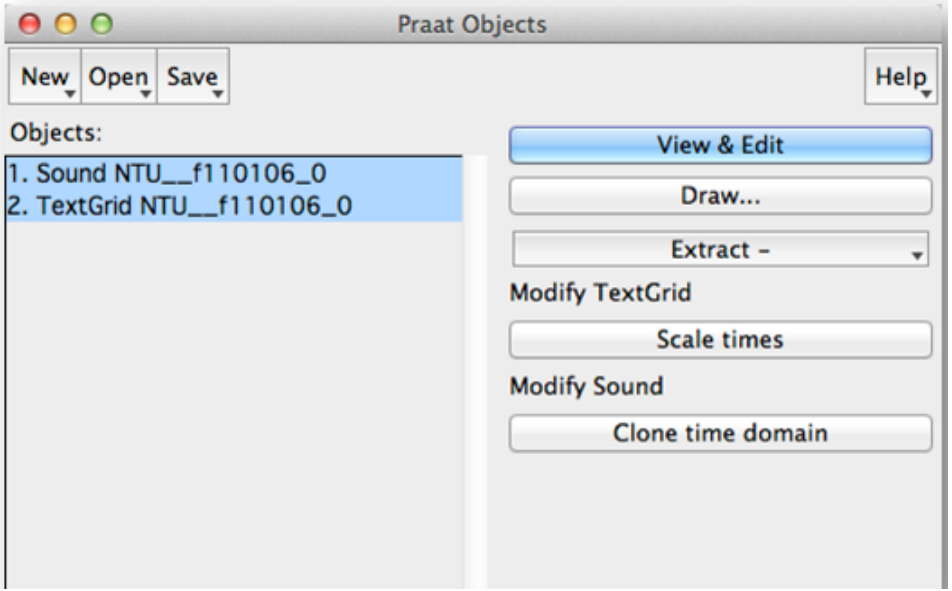
Praat - Label a wave file

- Create one interval tier named RCDIF
- No point tiers

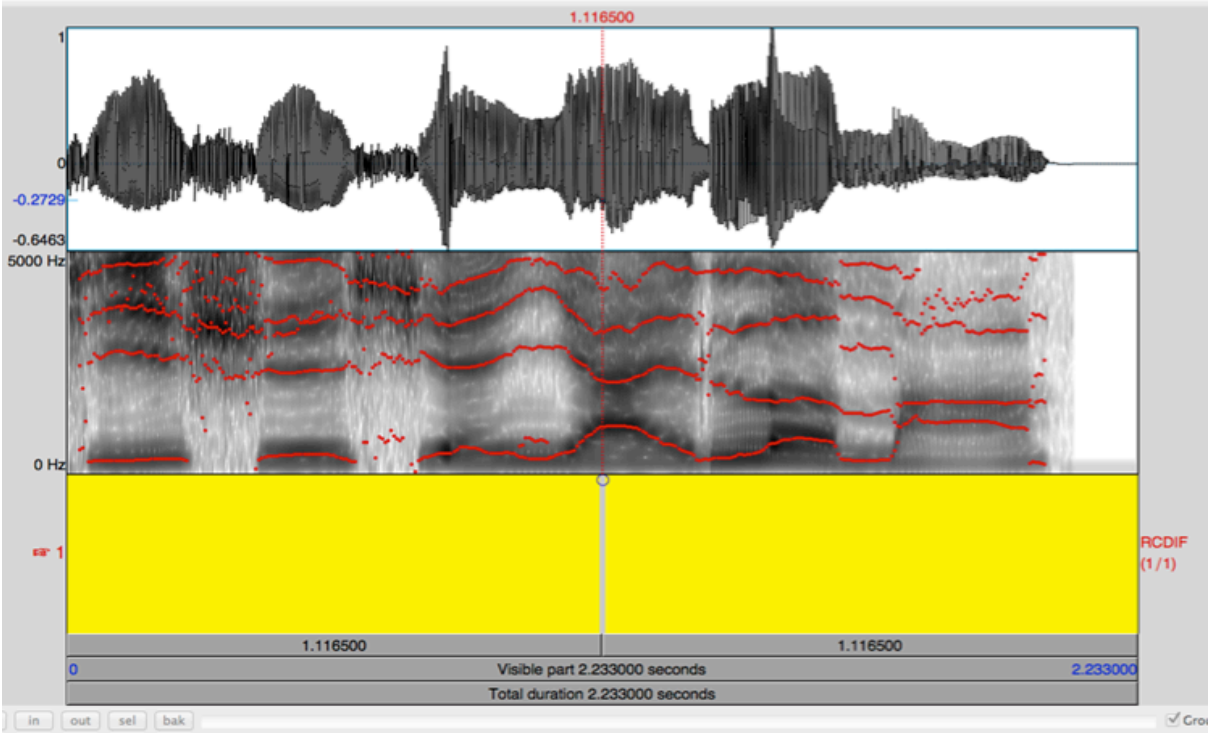


Praat - Label a wave file

- With **BOTH** objects selected
- click ***View & Edit***



Praat - Label a wave file



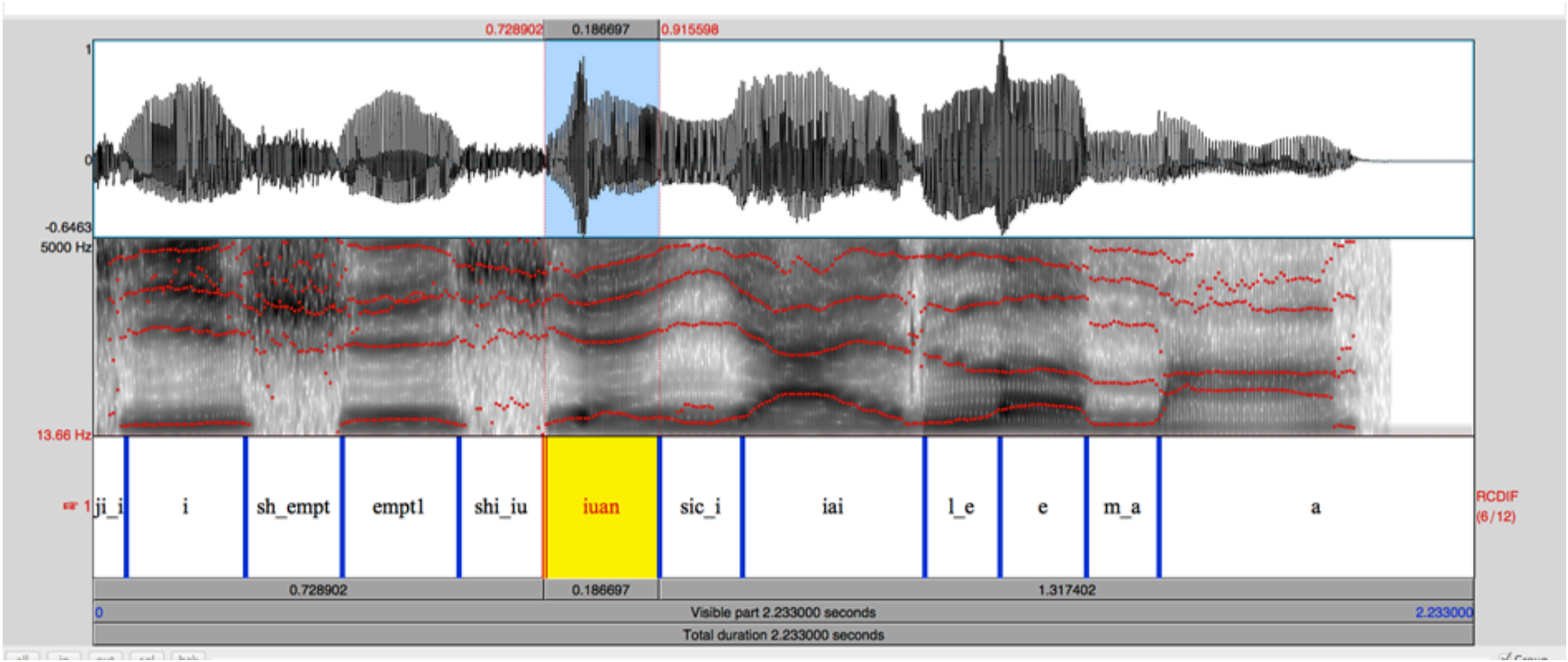
Praat - Label a wave file

- Click on spectrogram for your boundary
- Add the boundary by clicking the small circle

  Remove by choosing “Boundary/Remove”
- Drag you boundaries to be more accurate
- Click between your boundary and type in your label (according to the “Syllable table”)

  Listen to your label by clicking the number (interval time) below it

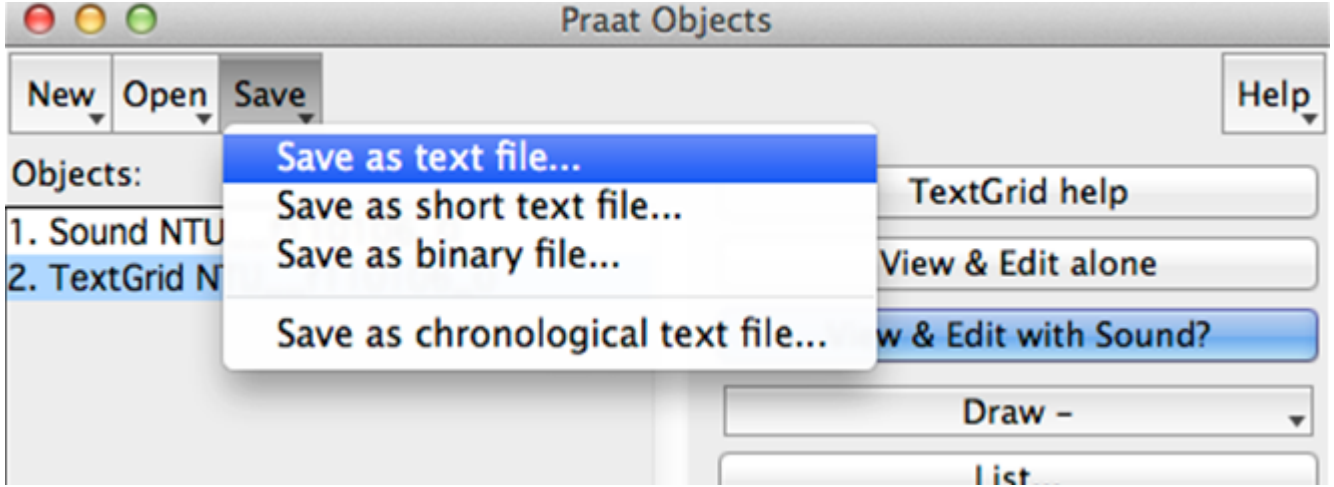
Praat - After labeling



Praat - Save your Label file

- Save your TextGrid object as short text file

  File should be “.TextGrid” not “.Collection”



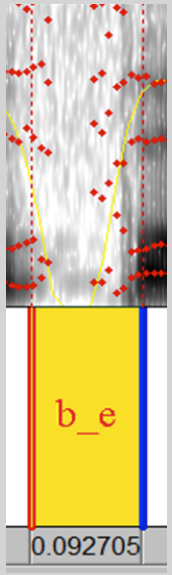
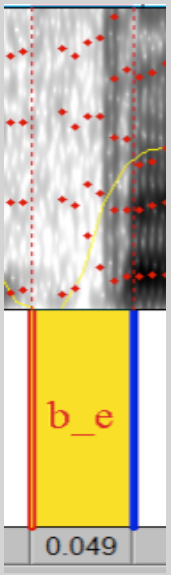
Report - Part 1 (20%)

- Choose your wave files from directories according to your student ID (https://goo.gl/ero6Ka).
- You must submit at least 5 fully labeled TextGrid files (**along with their wave files**).
- These 5 files should contain the initial/final labels you use in part 2.

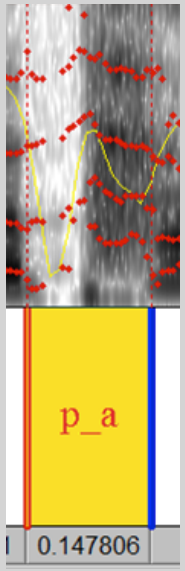
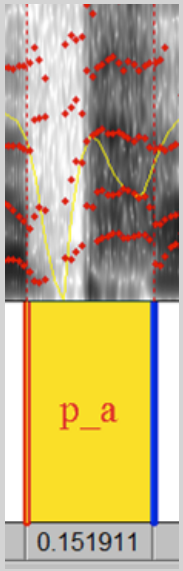
Report - Part 2 (30%)

- Choose at least 2 initials from the 4 classes (Plosive, Fricative, Affricate, Nasal)
- For each of these 8 initials, create a table that contains at least 2 screenshots of its label.
- Please show intensity and formant.

Part 2 - example：Plosive b (ㄅ)

Phonetic Class			
Plosive	b(ㄅ)		

Part 2 - example：Plosive p (ㄆ)

Phonetic Class			
Plosive	p(ㄆ)		

Part 2 - Useful tips

- Zoom in and Zoom out.
- show all or selection part in Praat by clicking the buttons on the lower-left corner of spectrograms.
- In your chosen directory.

  “NTU_XXXXX_phn2file” lists all files containing each phone

  “NTU_XXXXX_file2phn” lists all phones contained in each file

Report - Part 3 (50%)

1. (20%) What are the consistencies of the spectrogram in each phonetic class? (Plosive, Fricative, Affricate, Nasal)

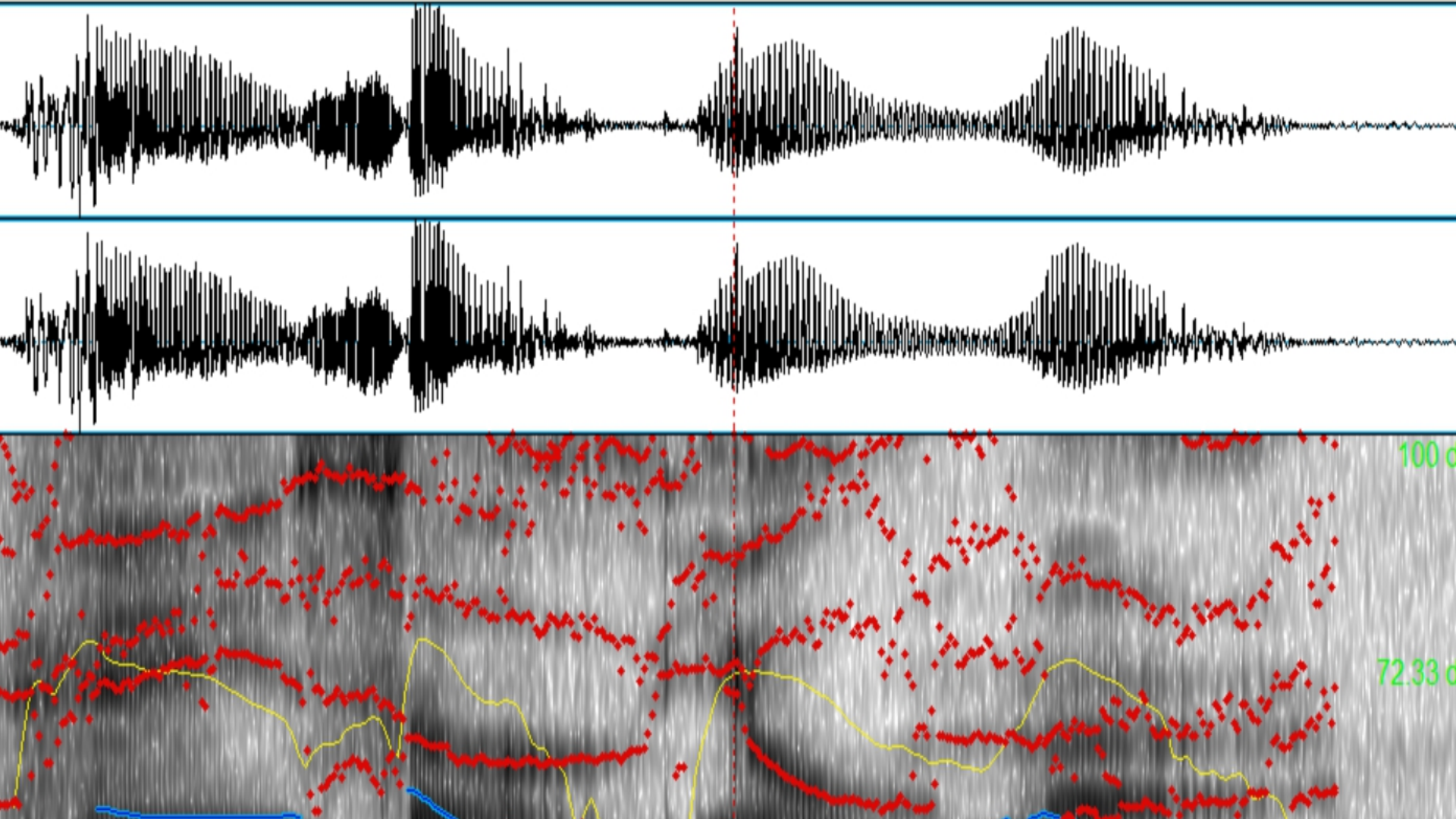
2. (10%) Is the boundary between neighboring initial and final clear? What is the benefit of using “right-context dependent” initial model (ex: sh_a) instead of pure initial model (ex: sh) to model initials?

Report - Part 3 (50%)

3. (10%) What are the differences when pronouncing ㄅ & ㄆ? How can you tell the differences in spectrogram for ㄅ & ㄆ? (You may also want to compare ㄉ & ㄊ, ㄍ & ㄎ respectively)
4. (10%) Take a look at the spectrogram of finals. Is there any simple rules to discriminate initials from finals provided only spectrogram?

Report - Bonus (10%)

- The following is a speech analysis plot for a Chinese word composed of 4 characters. Each character is composed of an initial and a final.
- Guess what the word is and describe your reasoning.
  (Score: reasoning 8%, correct answer 2%)
- If you cannot figure out the word, you can guess the phonetic class or initial/finals.
  For example, your answer can be “l_i, i, sic_a, au” or “plosive, diphthong, plosive, monophthong”.


100 d
72.33 d

Submission Requirements

1. 5 TextGrid files (each along with its wave file).
   the ".TextGrid" & ".wav" filenames should be the same.

2. 1 report (in PDF format).
   the filename should be hw2-2_bXXXXXXXX.pdf (your student ID).

Submission Requirements

3. Put those 11 files in a folder, compress the folder to 1 zip file and upload it to ceiba.
   - Folder name should be hw2-2_bXXXXXXXX (e.g. hw2-2_b02901000)
   - .zip or .tar.gz only
   - 20% of the final score will be taken off for wrong format

Homework 2

- Your can submit either

  HW 2-1 (HMM Training and Testing)

  HW 2-2 (Speech Analysis)
- You can also submit both
- The higher grade of the two will count as your final score for HW2
- **Deadline: To be discussed**

10% of the final score will be taken off for each day of late submission