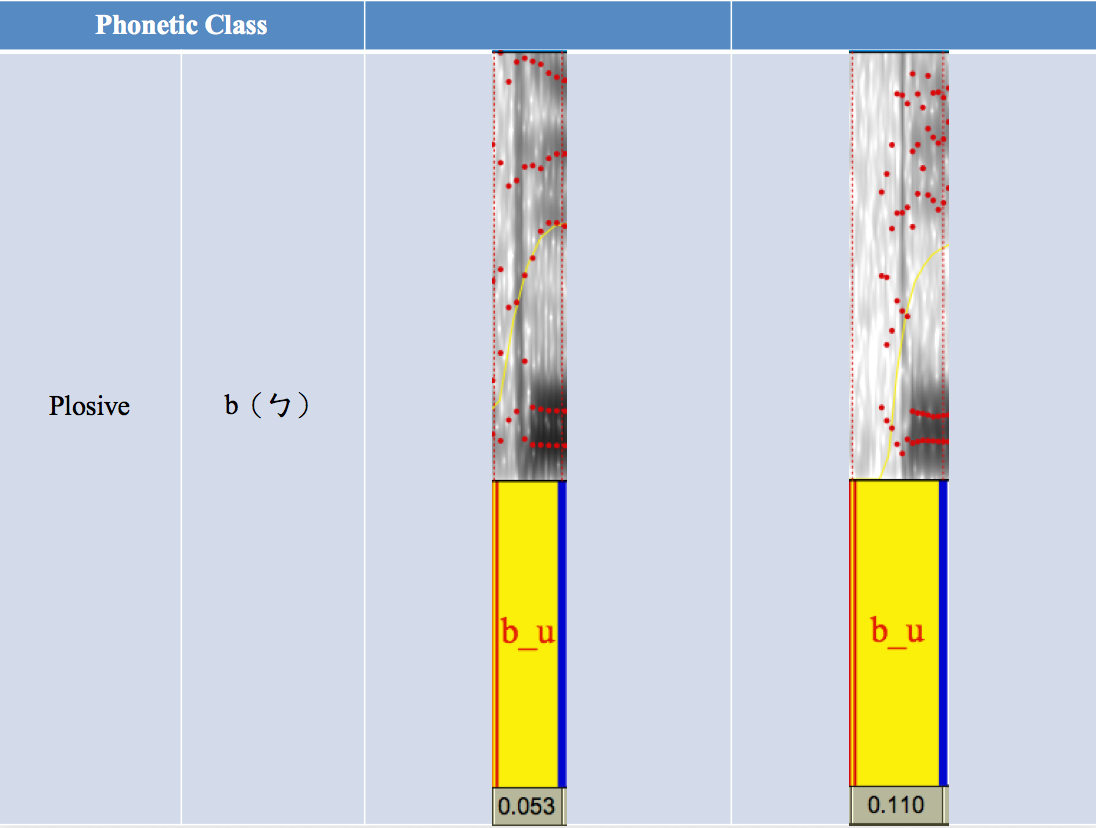
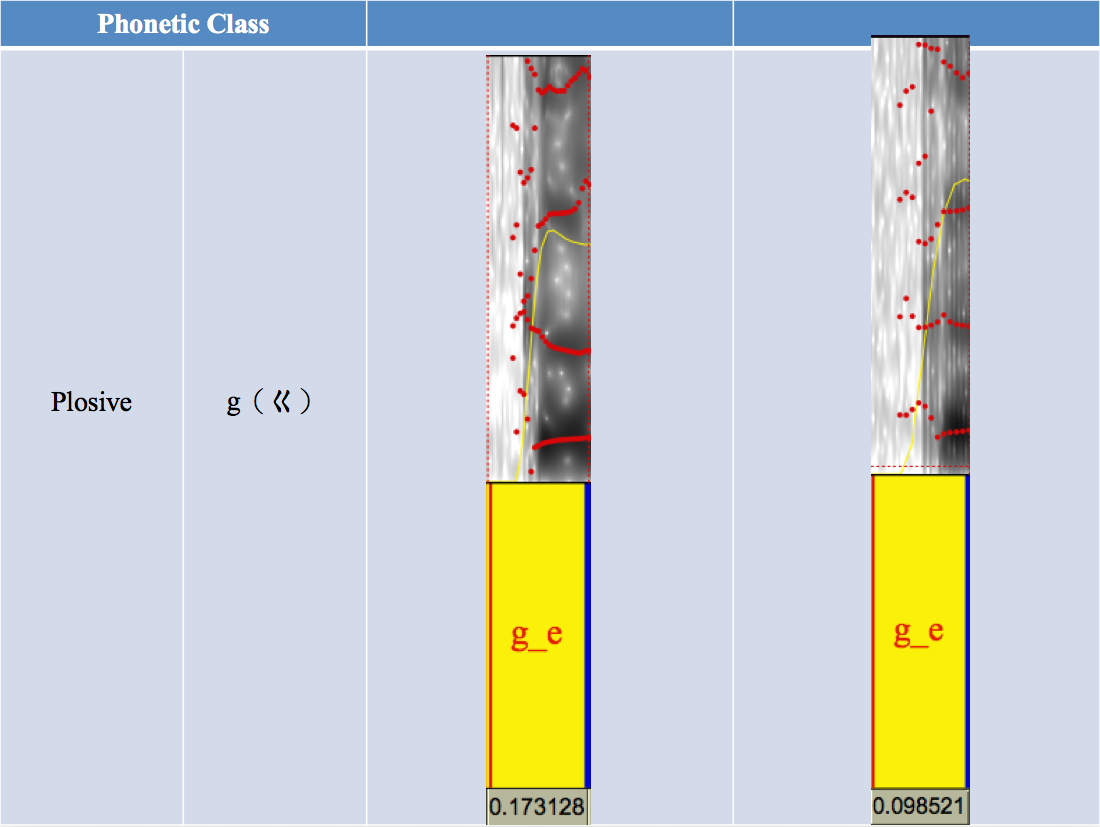
Fundamentals of Speech Signal Processing 2015 Spring

**Homework 2-2 Report**

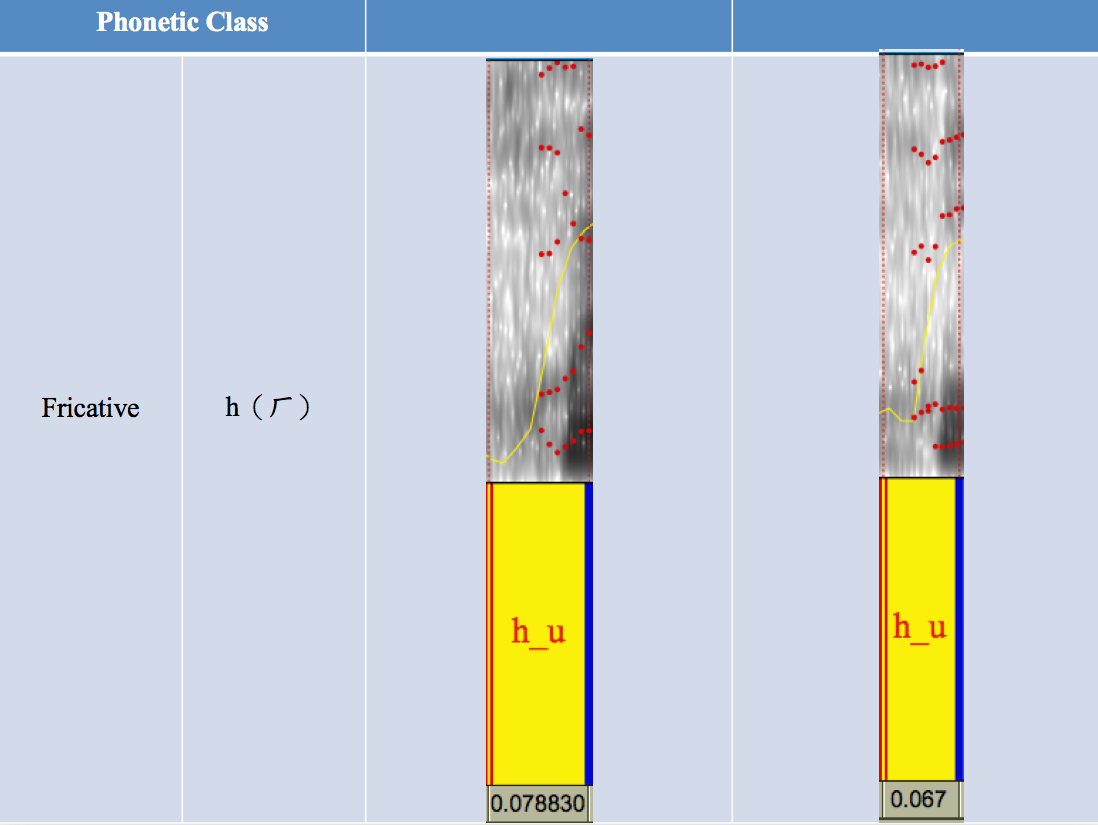
*Part 2*

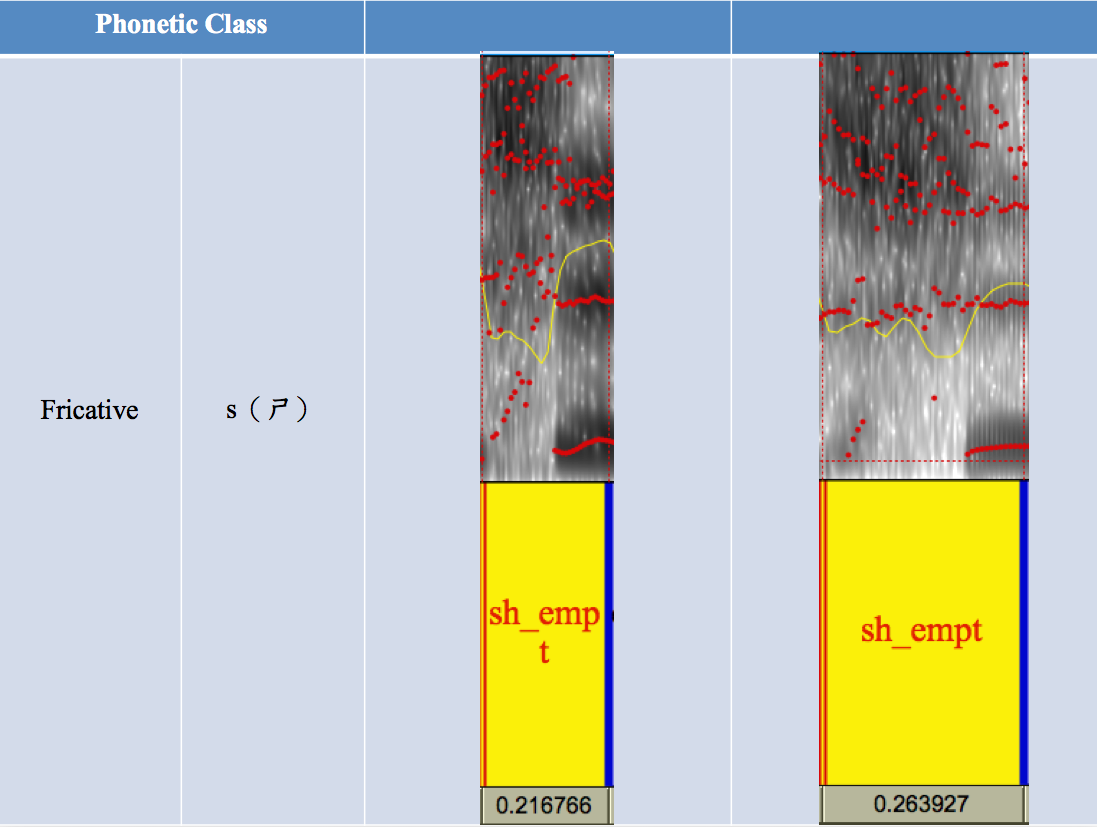
Plosive



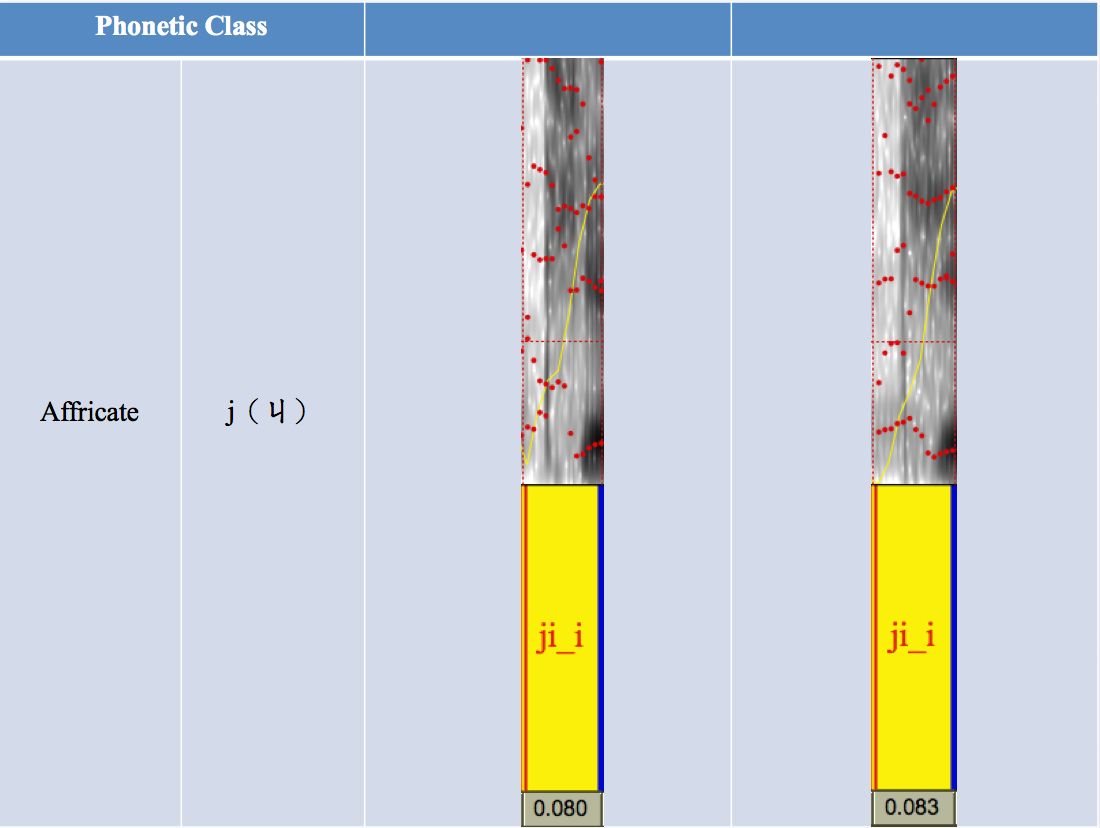


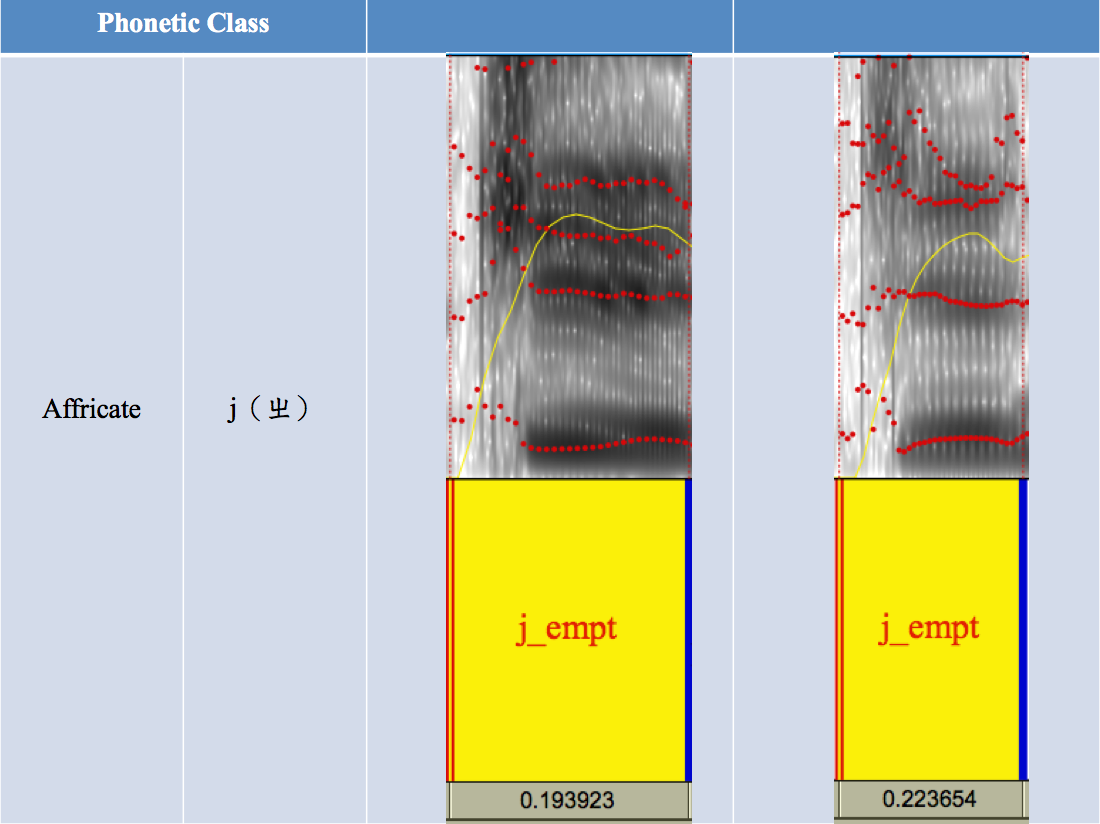
Fricative



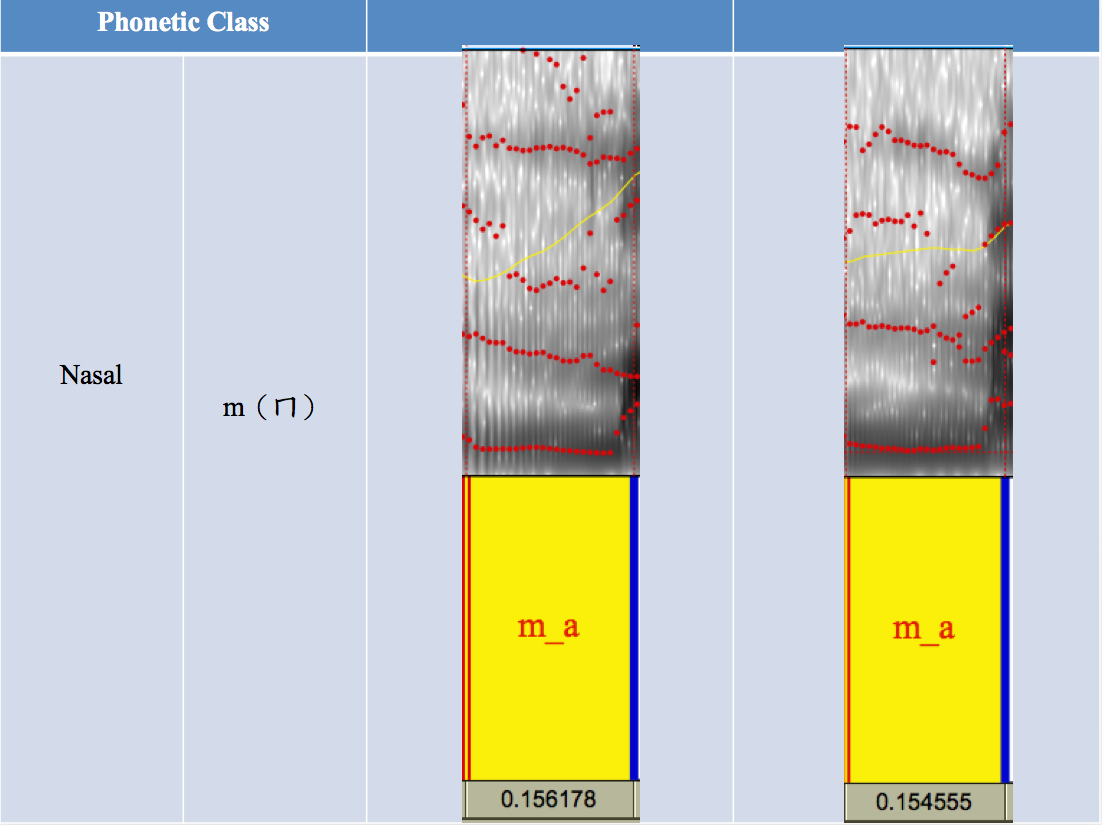


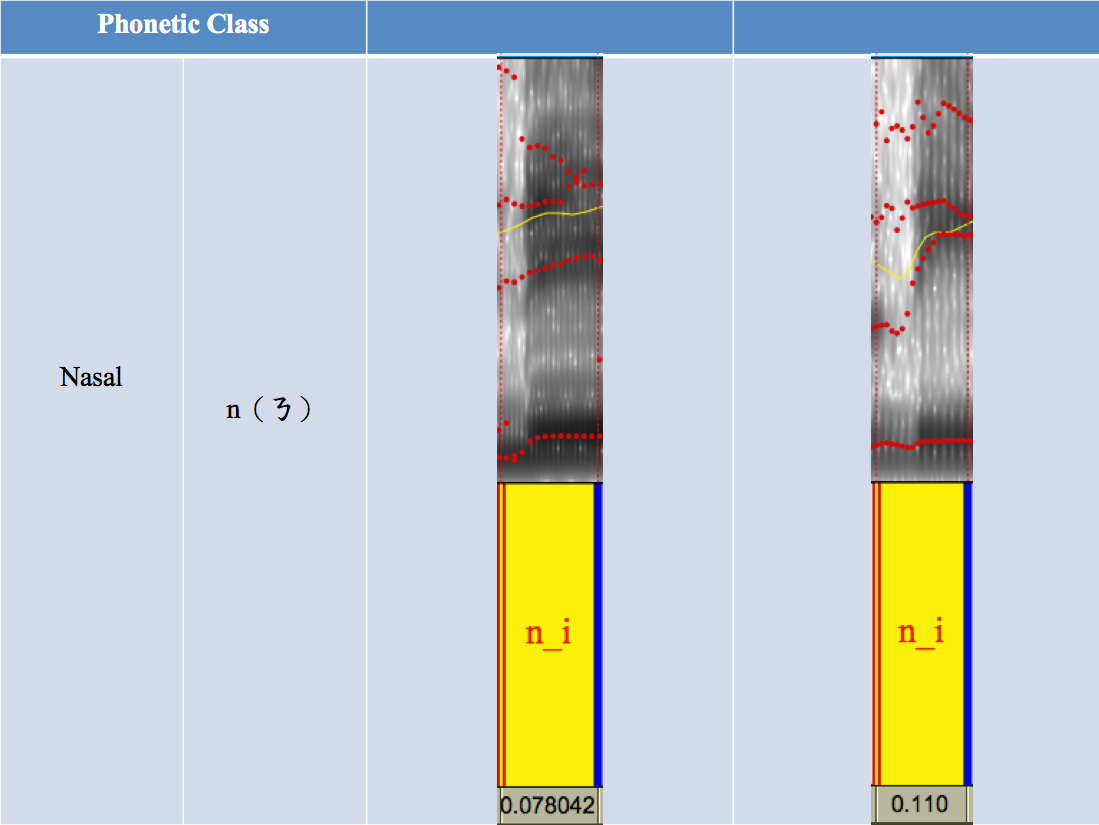
Affricate





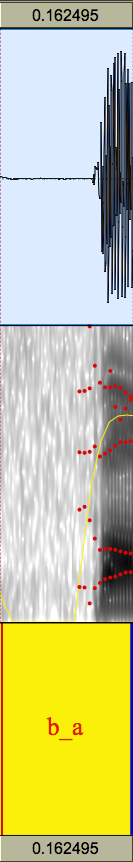
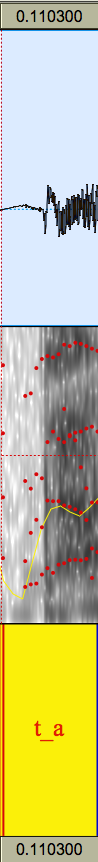
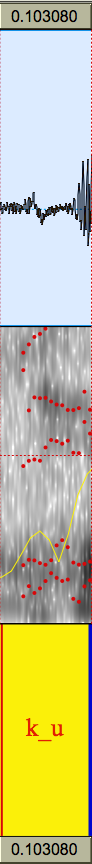
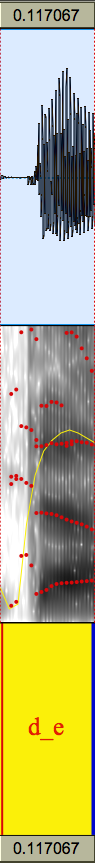
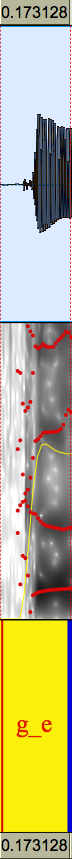
Nasal





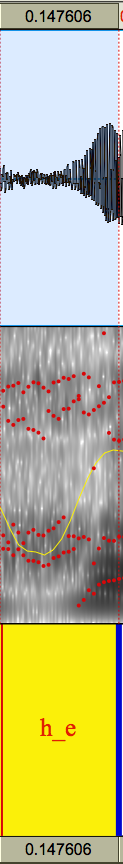
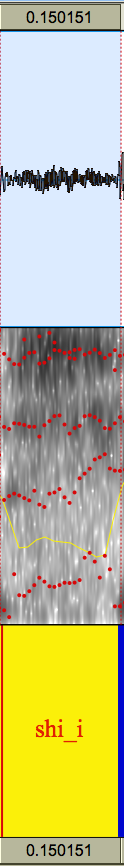
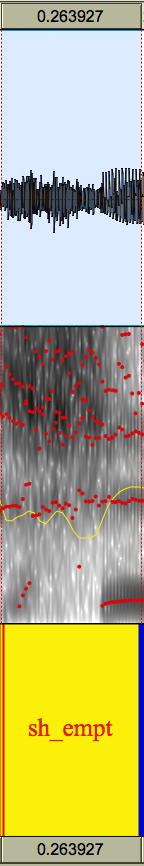
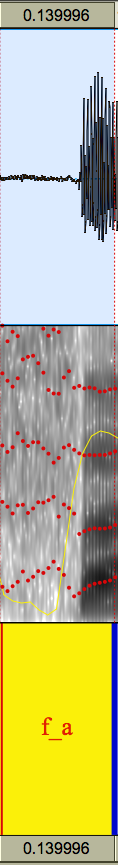
*Part 3*

1. (20%) What are the consistencies of the spectrogram in each phonetic class? (Plosive, Fricative, Affricate, Nasal)
2. Plosive的音具有很明顯的共通性，以下面這幾張spectrogram截圖為例：

 ../Plosive/螢幕快照%202015-05-08%20下午5.00.17.png    ../Plosive/螢幕快照%202015-05-08%20下午5.05.03.png 

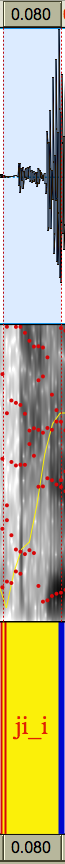
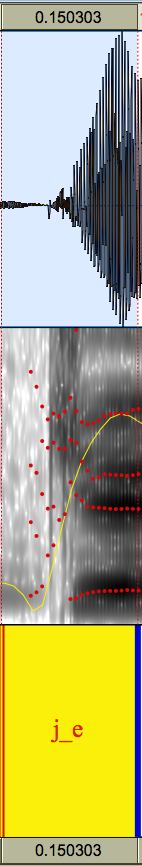
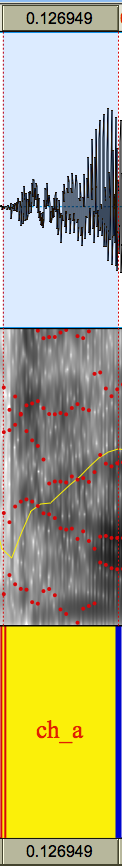
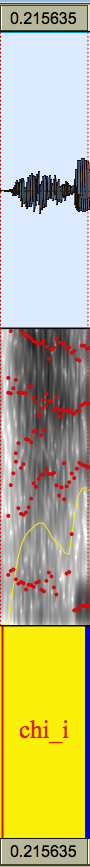
我們可以發現有關聲音訊號強度（黃線）的部分，Plosive都有一開始強度低，然後突然增強的一致性；另外一點值得注意的，是背景的spectrum大多是由白急遽轉黑，不論是訊號強度或是光譜，都呈現極具的變化，或許這就是「爆破音」這個名稱的由來吧。

1. Fricative以下列截圖為例：

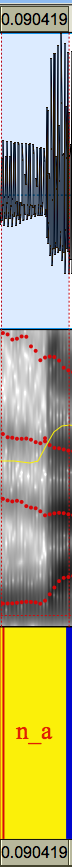
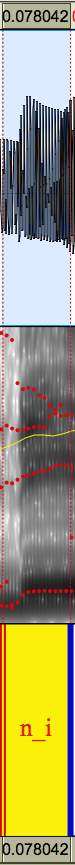
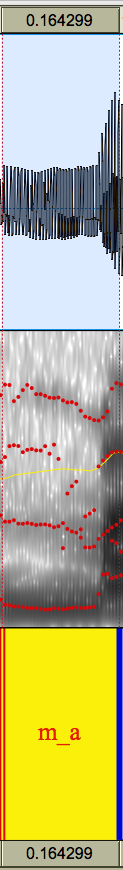
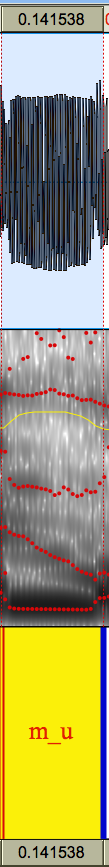
相較於上面的爆破音，擦音的訊號強度（黃線）就比較平順，雖然還是有一定的起伏變化，但是不會一下子由弱突然衝到強；另外spectrum的黑白分佈也比較平均。有趣的事，我發現第四張圖，也就是ㄈ這個音，就我所觀察到的這兩個性質，跟爆破音很類似（急遽的訊號強度變化，與背景光譜的黑白混雜程度）。

1. Affricate以下列幾張截圖為例：

    ../Affricate/螢幕快照%202015-05-08%20下午5.26.10.png

鼻擦音就我所觀察的這幾張圖，比較難發現他們共同的特徵，有些像爆破音，如ji\_i，有些又像擦音，如ch\_a。我猜測鼻擦音之所以命名為鼻擦音，是因為它的性質介於猜因跟鼻音吧！

1. Nasal以ㄋ、ㄇ各兩張截圖為例：

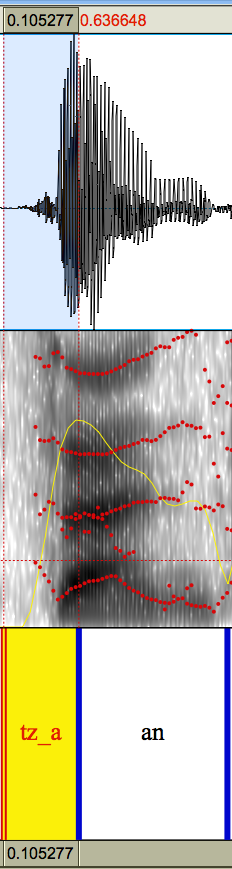
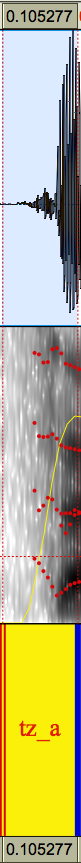
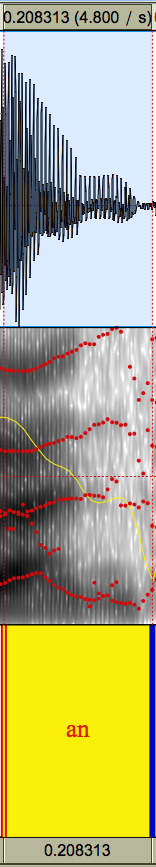
   

鼻音的特徵一目了然，那就是訊號強度相當一致、平，幾乎就是一條直線（當然還是會因為後

面的dependency而有點起伏）。

1. (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 (ex: sh) to model initials?

做完這次的作業，覺得在大多數情況下，initial跟final其實還算好分辨，而且在模糊無法辨別（有時真的很難切開）的時候，其實偷偷看一下訊號的分佈可以幫助判斷，例如：欲將「攢」切開成聲母跟韻母，也就是想切開tz\_a跟an的音，有時候並不是那麼容易判斷，這時偷偷看一下訊號強度（黃線），就可以猜測最高峰是個轉捩點，大膽嘗試切開之後重聽一遍，發現效果很不錯。

🡪

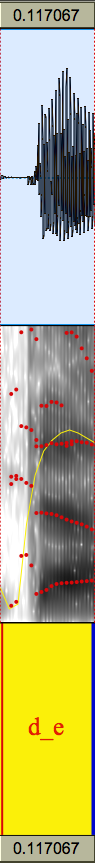
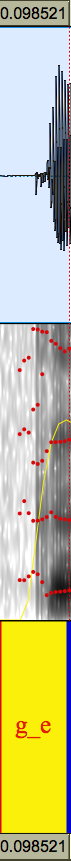
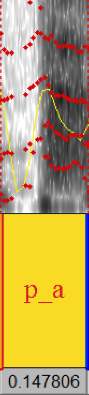
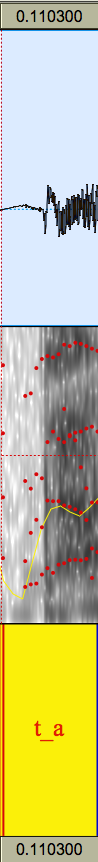
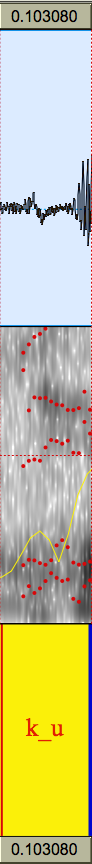
我認為妥善運用RCD也可以幫助判斷切開點，同樣用上圖為例，判斷ㄗ的切開點時，若是只用tz\_empt（without dependency）的資訊會很難抓，用tz\_a就沒這問題！

1. (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)

在發ㄆ音的時候，會先發出一些氣音，而發ㄅ音就是直接出聲。用這個特點來看，會發現ㄉ跟ㄊ也是類似：發ㄊ音時也是先發出一些氣音，而發ㄊ音時就是直接出聲；ㄍ跟ㄎ的比較一樣，ㄎ會先發出氣音，ㄍ則是直接出聲。

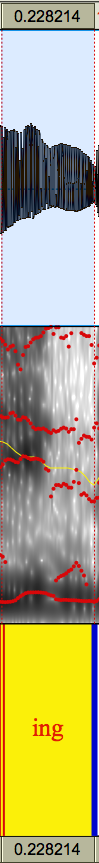
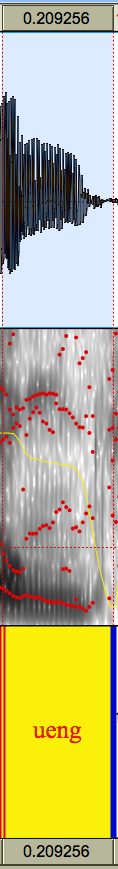
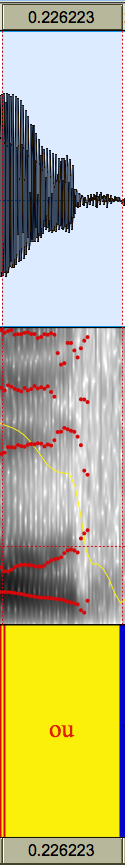
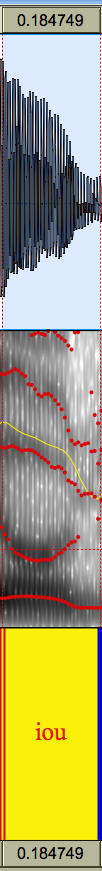
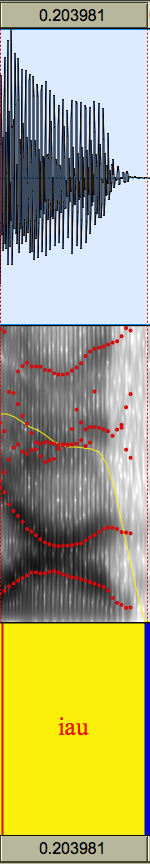
由spectrogram來看：

ㄅ、ㄉ、ㄍ一類（直接發聲）： ㄆ、ㄊ、ㄎ一類（先發出氣音）：

../Plosive/螢幕快照%202015-05-08%20下午5.00.17.png     

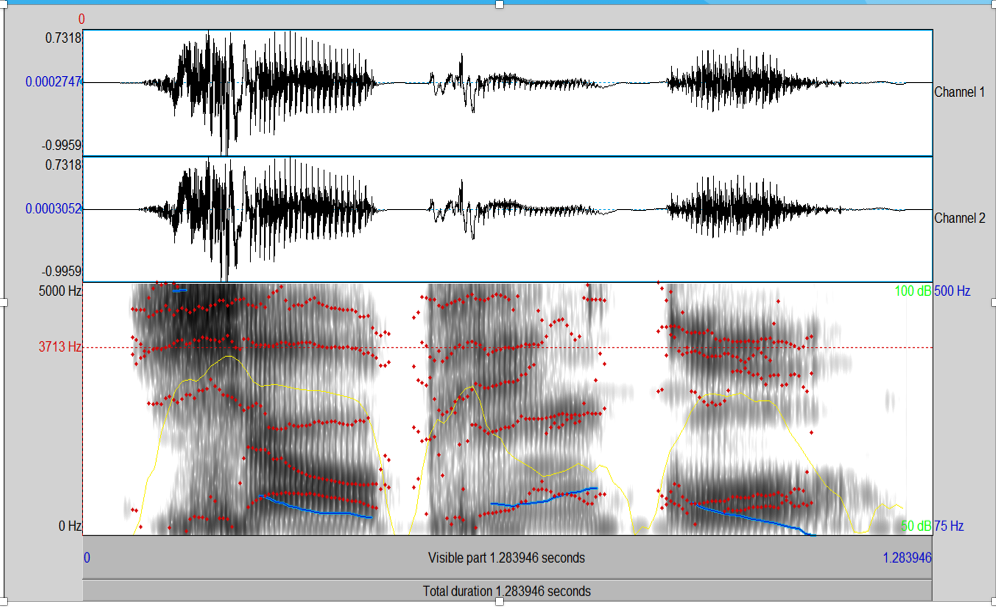
我們可以發現ㄅ、ㄉ、ㄍ這一類的訊號長相比較單純，由弱直接跳到強，有種直接「爆」的感覺；而ㄆ、ㄊ、ㄎ這一類的訊號稍微變化複查一點，雖然趨勢也是由弱到強，但中間稍有曲折，或許這就是因為先發出氣音的影響吧。

1. (10%) Take a look at the spectrogram of finals. Are there any simple rules to discriminate initials from finals provided only spectrogram?

以上五張圖是一些finals的截圖，我們可以大致歸納出一個簡單規則，那就是訊號強度最終會降低！有些finals或許可以維持強度不下降太快一段時間，但是最終還是會劇烈下降。

*Bonus*



第一個字的initial，因為訊號強度急遽的提升，所以我猜是Plosive。