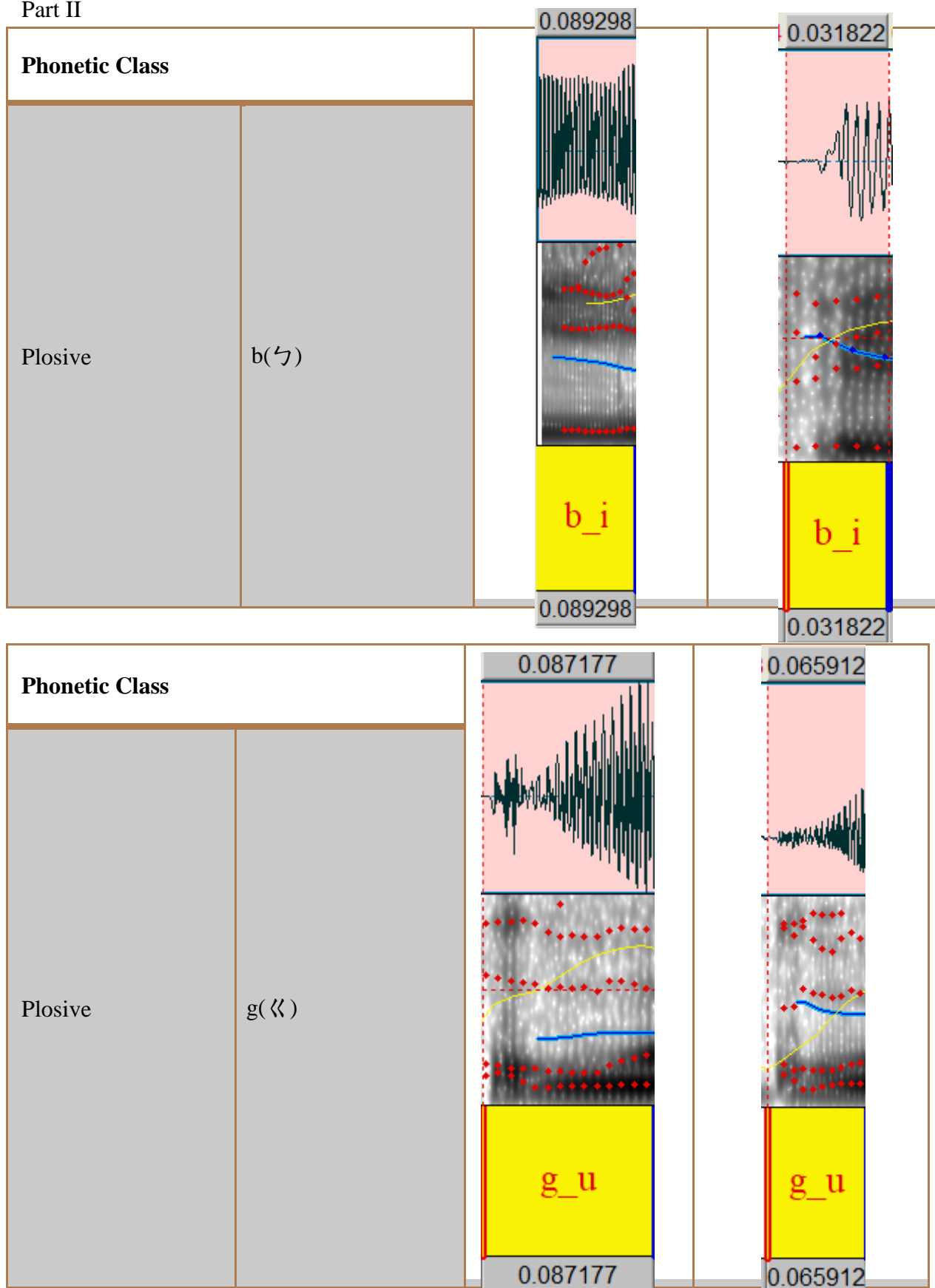
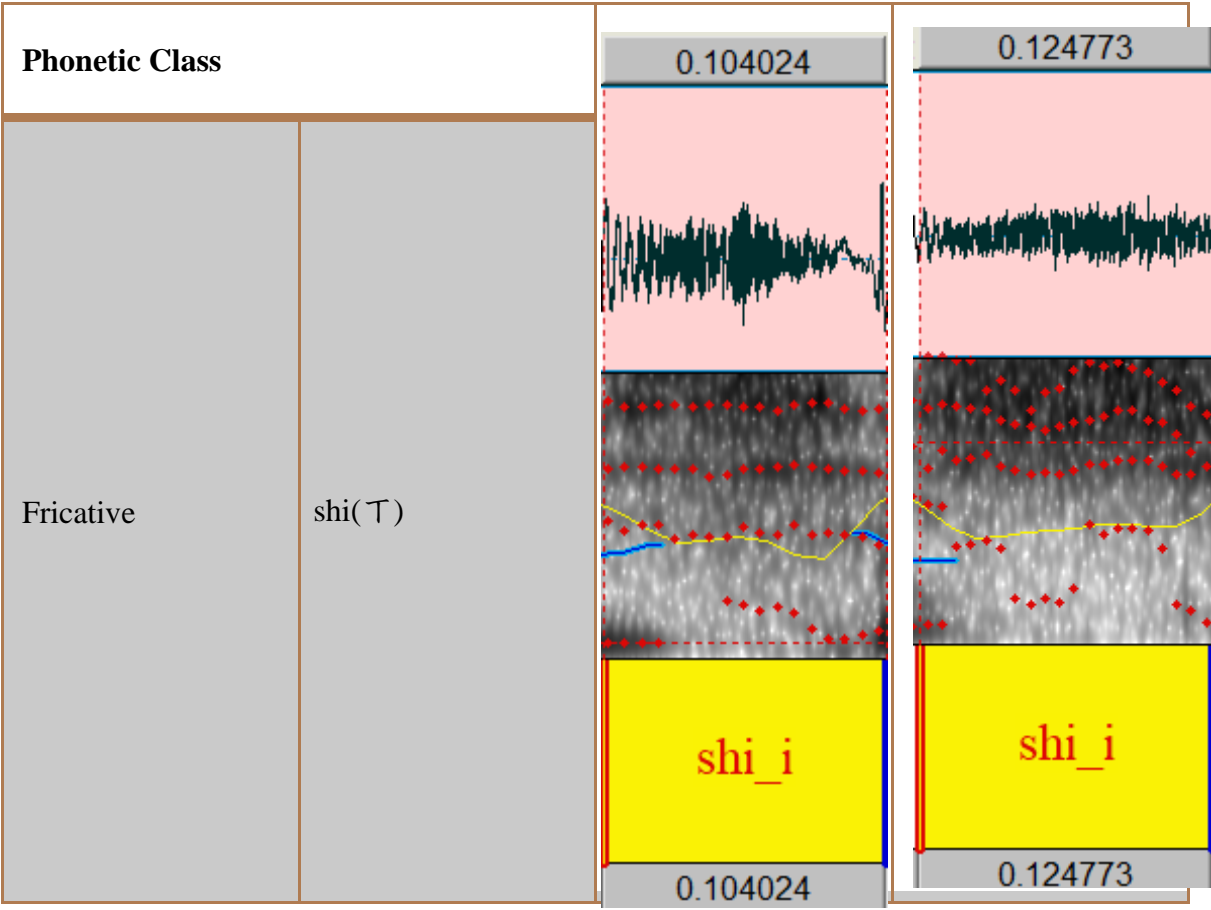
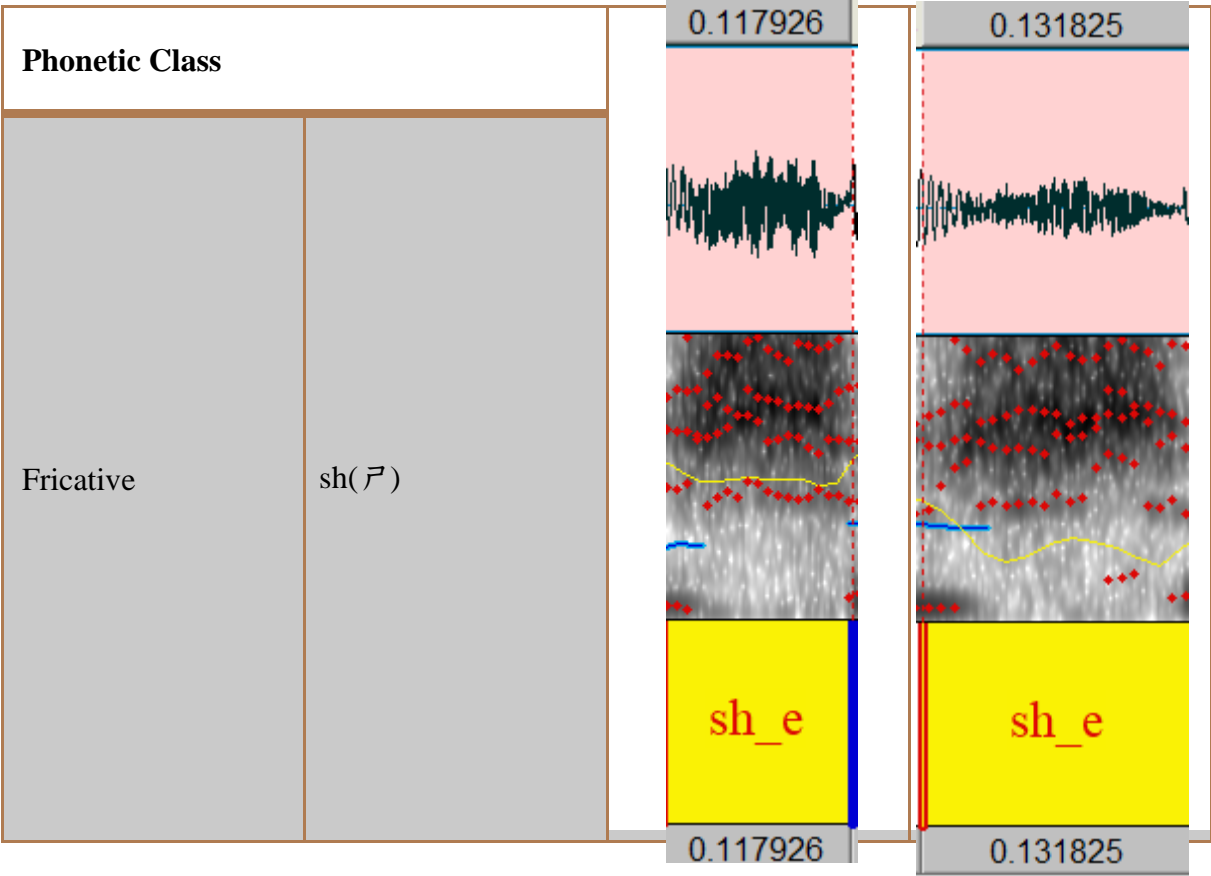


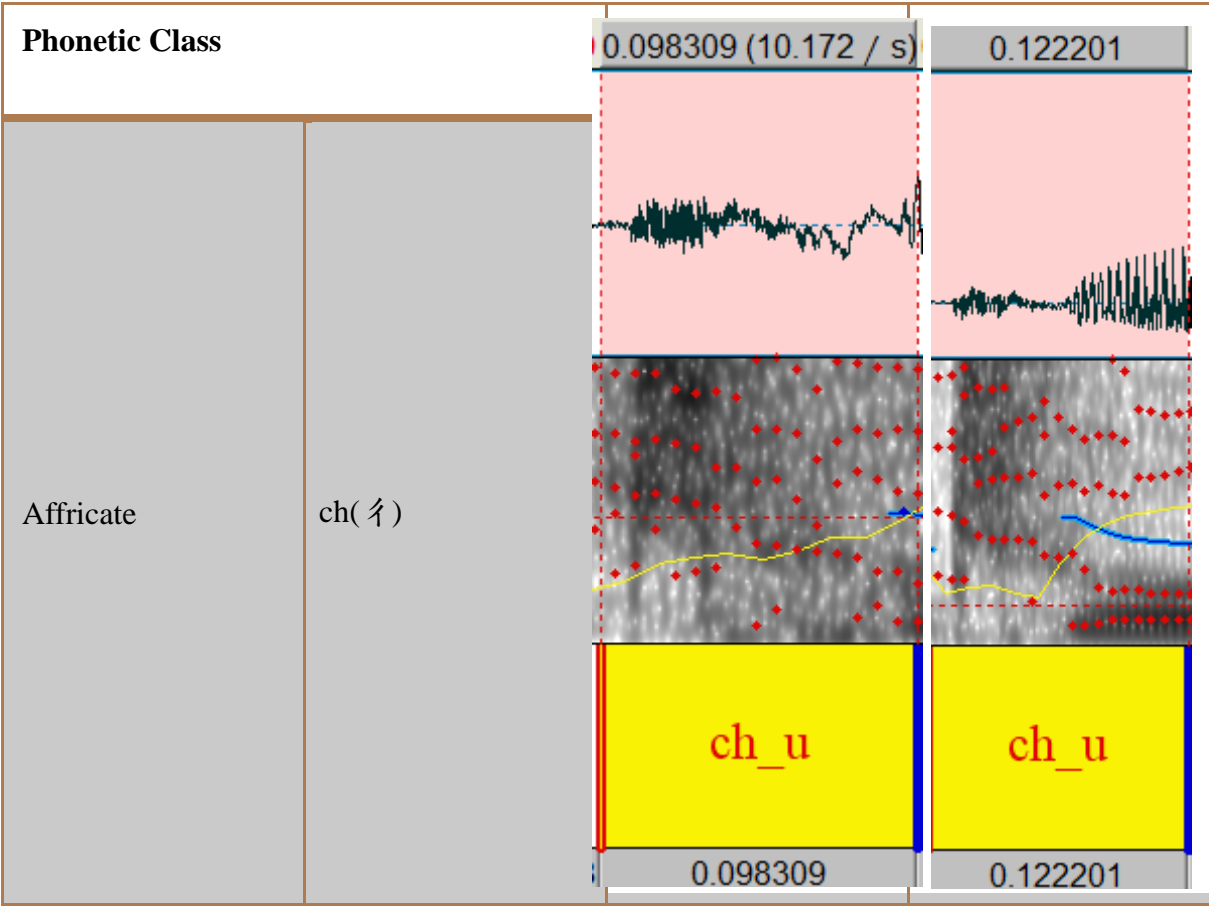
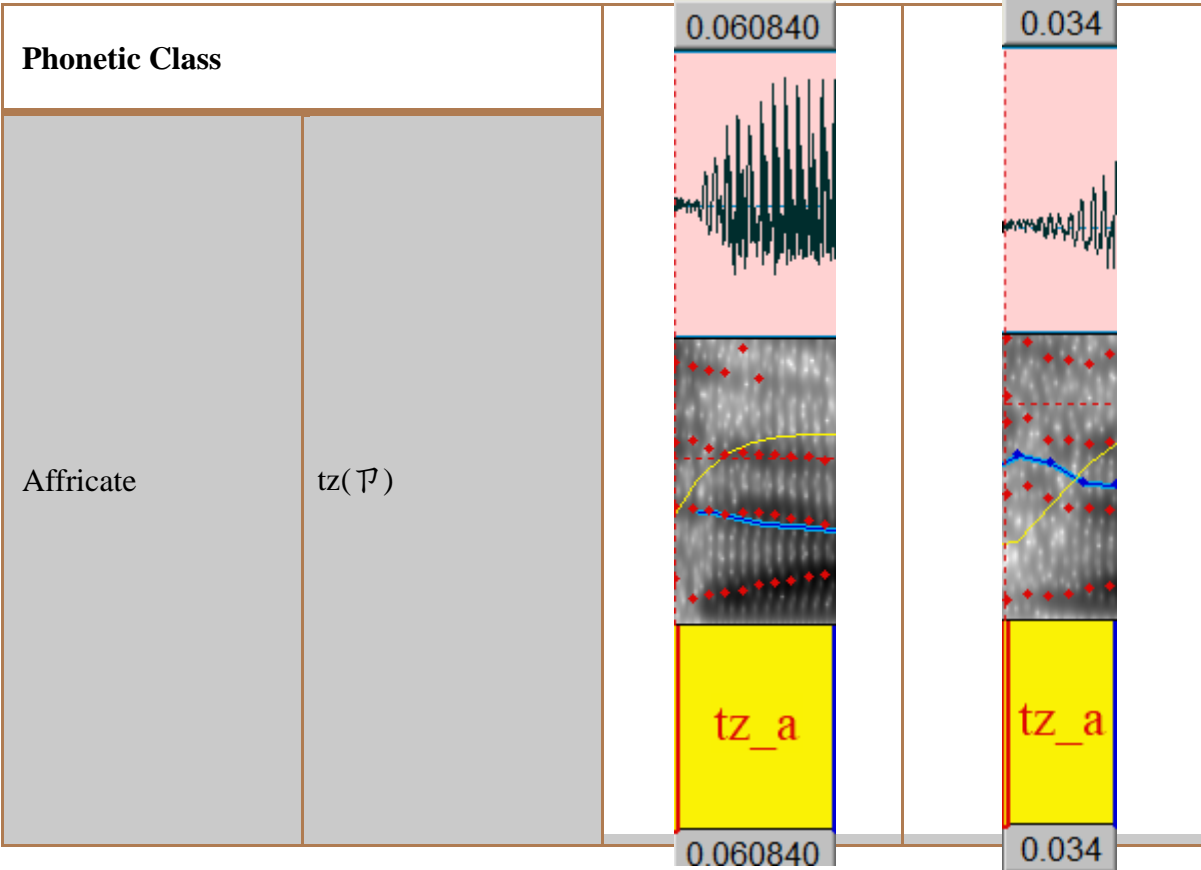
# Digital Speech Processing HW2-2 report

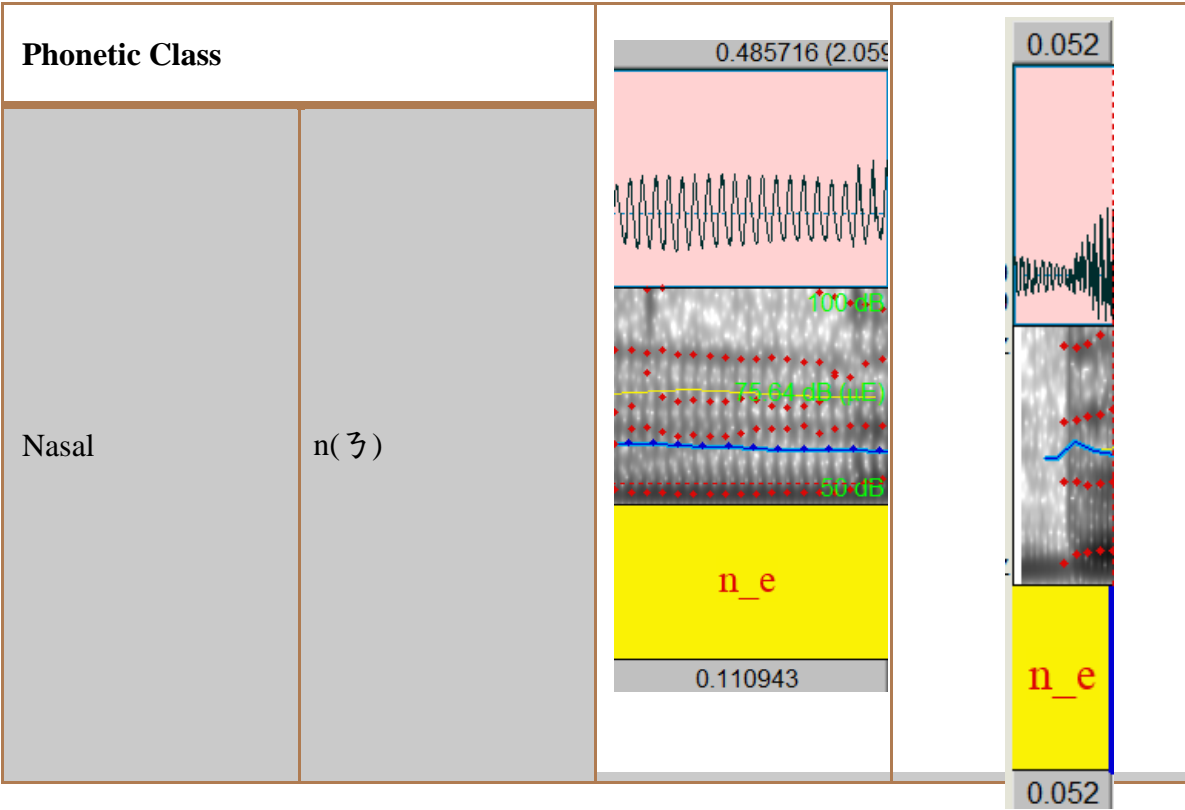
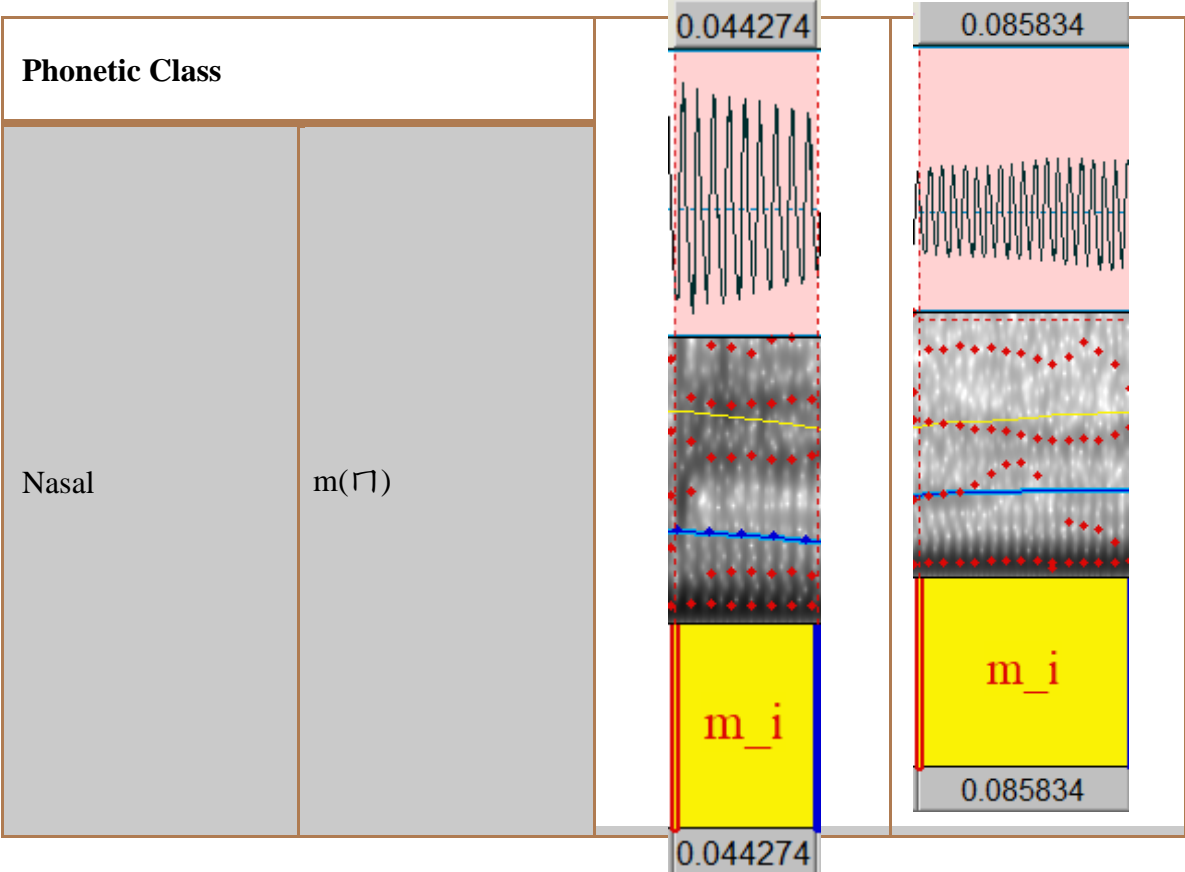
R06942048 電信碩一 林彥伯

## Part II





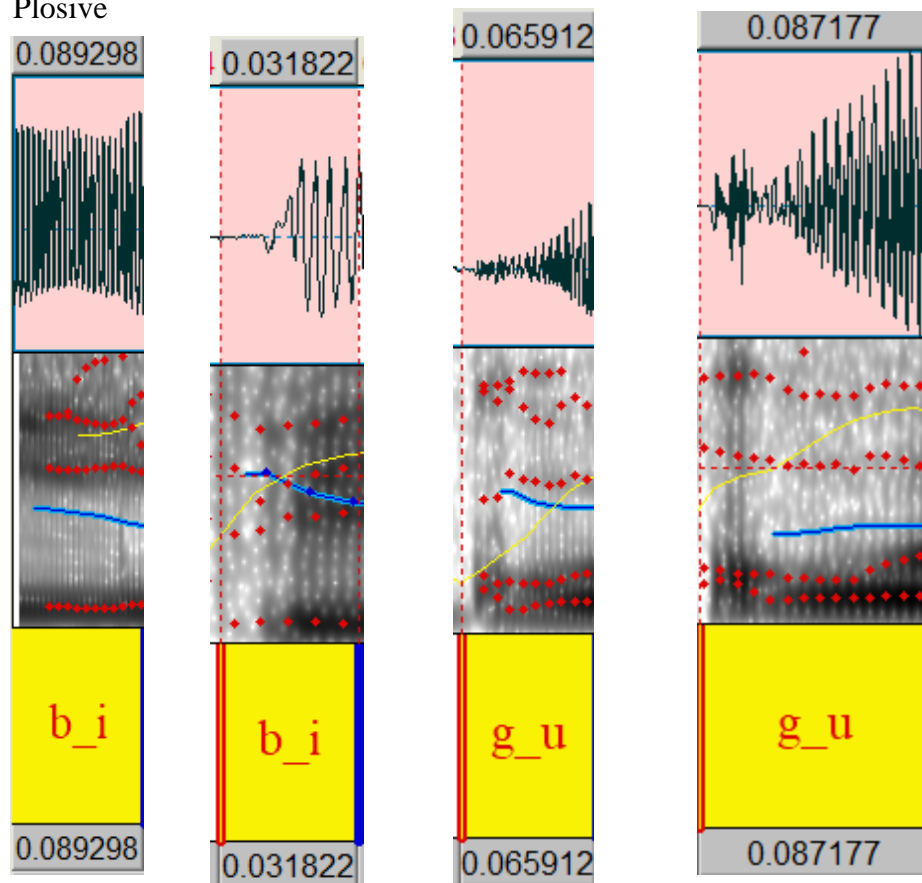




### Part III

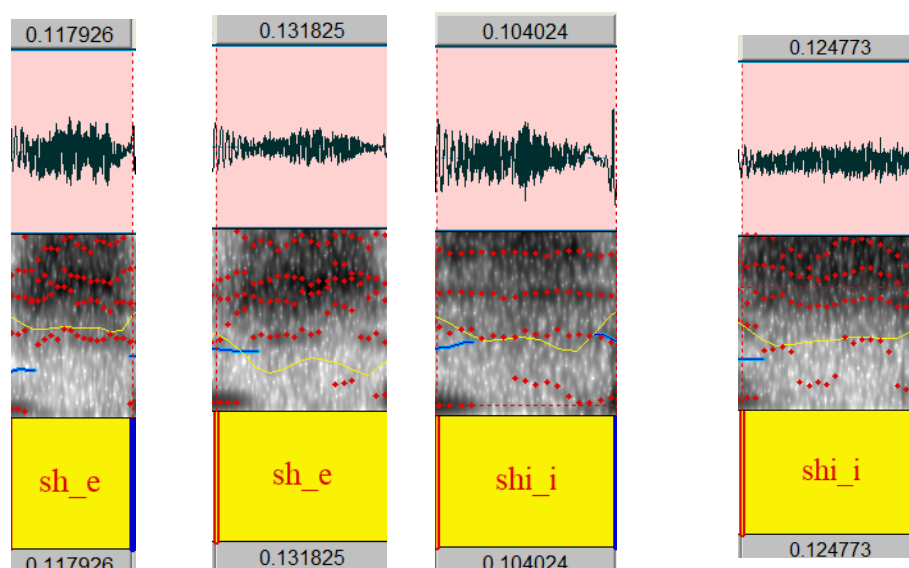
- What are the consistencies of the spectrogram in each phonetic class? (Plosive, Fricative, Affricate, Nasal)?

#### A. Plosive



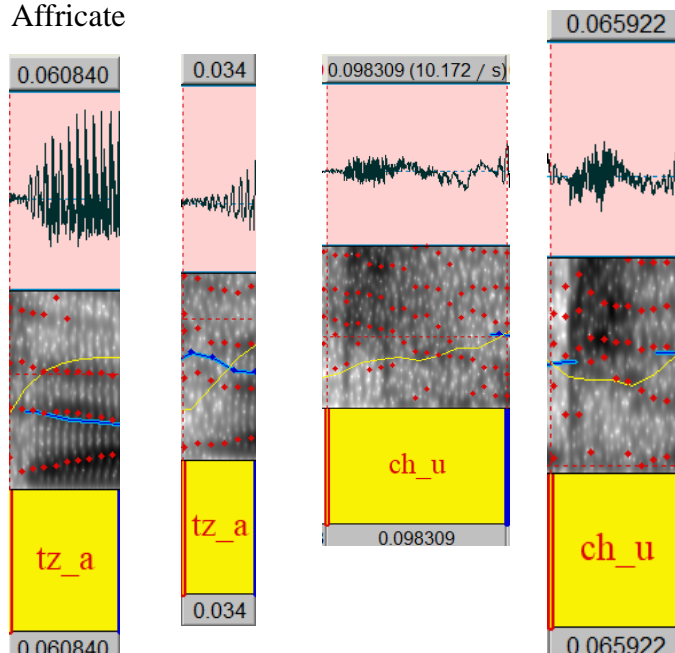
由此圖可以發現，Plosive 在黃色的聲音訊號強度會忽然往上拉，在 time domain 的狀況可以發現有種漏斗形狀加大的感覺。

#### B. Fricative



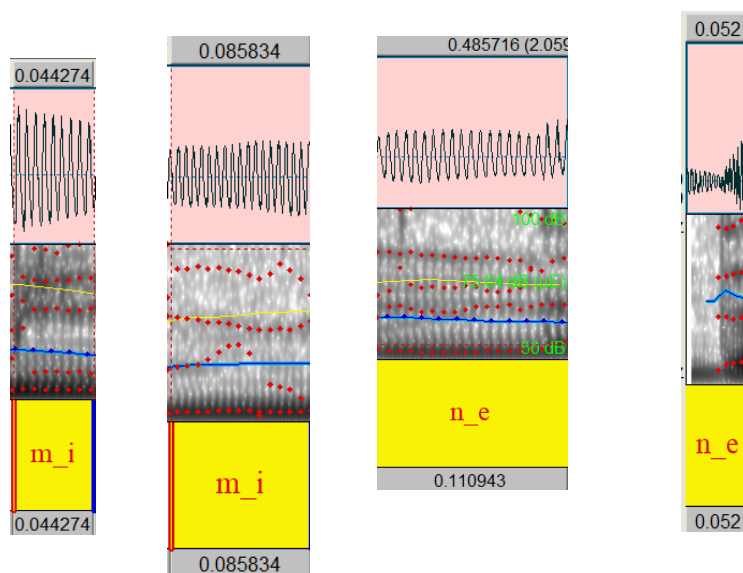
由這些圖可以發現到 Fricative 的黃色能量部分較為平穩，圖片中的聲音為尸  
 厶，這兩種聲音較為接近，所以在 frequency domain 上的能量分布非常相近(都在  
 頻率比較高的地方)

### C. Affricate



由這些圖可以發現到 Affricate，黃色的能量線與 Plosive 有相似的性質，可能是  
 發音的方式與 Plosive 接近再混合舌頭往前牙齒的聲音(可能空氣與舌頭跟牙齒摩  
 擦的感覺吧)

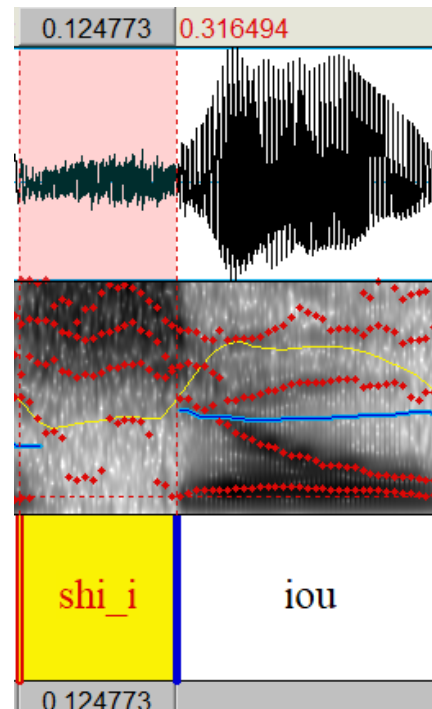
### D. Nasal



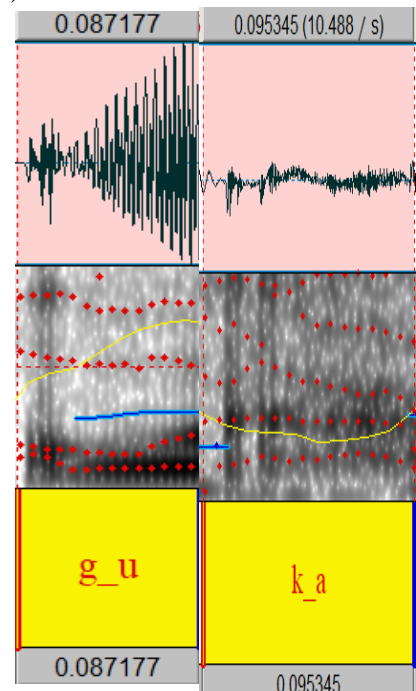
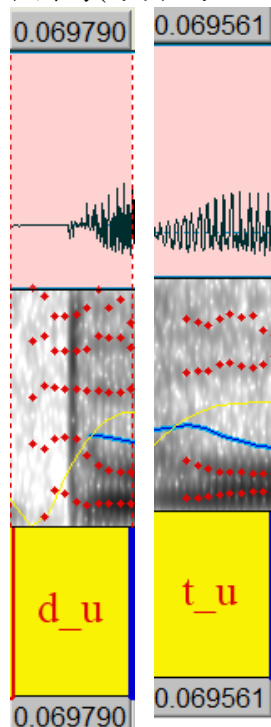
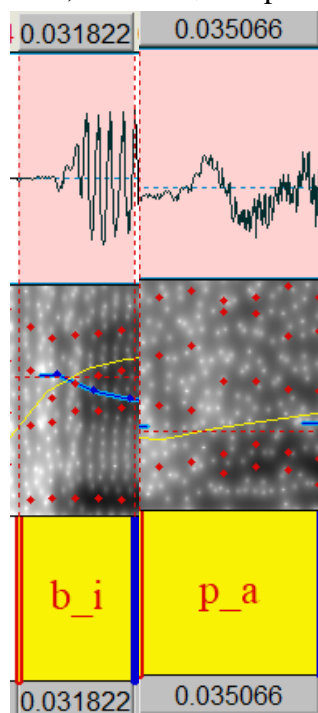
由這些圖可以發現到 Nasal 的所有分布都非常平穩，幾乎是一條直線。

2. 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?

基本上算好分，若分不清楚可以透過 time domain 或 frequency domain 來判斷，像一些 ㄌ、ㄐ 的聲音，在 frequency 的部分會只有高頻(如下圖)。某些 RCD 的組成會使聲母分布在只有高頻或低頻部分。



3. 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 ㄅ & ㄆ, ㄅ & ㄆ respectively)? 下圖為(ㄅㄆ ㄅㄆ ㄅㄆ ㄅㄆ)

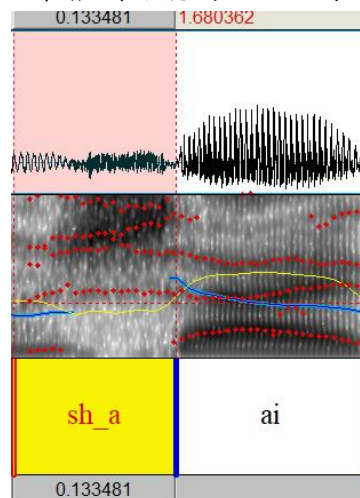




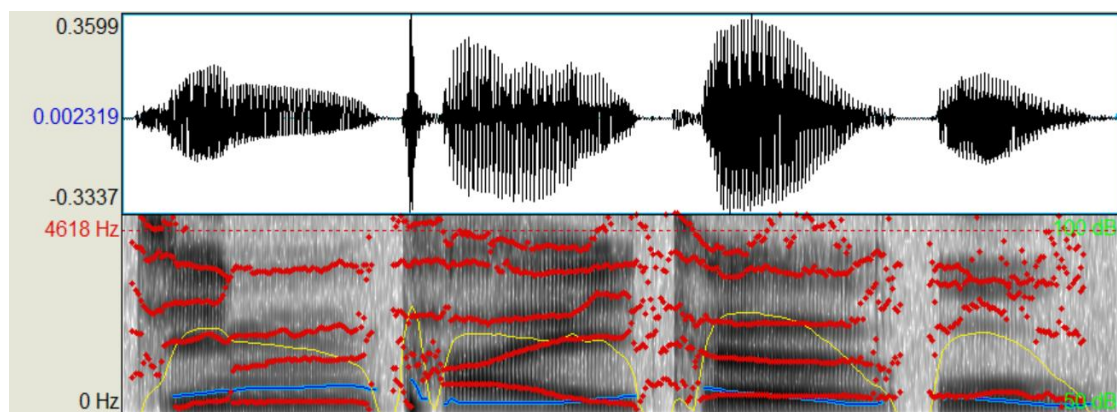
由上圖觀察，雖然能量線都是急速上升，但在觀察 time domain 的訊號時，可以發現到ㄅㄆㄇ有種能量被壓住爆開的感覺(典型漏斗的形狀)。ㄅㄆㄇ雖然也有爆破的感覺，但沒有能量被壓住爆開的感覺，而是直接爆開(由發音去感覺)，在 time domain 的狀況下就不會觀察到漏斗狀的形狀了。

4. Take a look at the spectrogram of finals. Is there any simple rules to discriminate initials from finals provided only spectrogram?

如下圖，只有 spectrogram 的話，可以透過電腦來辨認 finals(人眼應該可以辨識，不過困難度較高)，然後 initial 跟 final 又有一定連貫性，所以可以藉由辨認出的 final 來推測可能的 initial 再用推算出機率最高的 initial。



**Bonus:**



第一個音可能為ㄌ ㄍ ㄎ，因為 frequency domain 的高頻較多，  
第三 四個音 應該是ㄜ 聲，藍線都下降 且為爆破音 (漏斗狀??)

猜是**金牌特務**