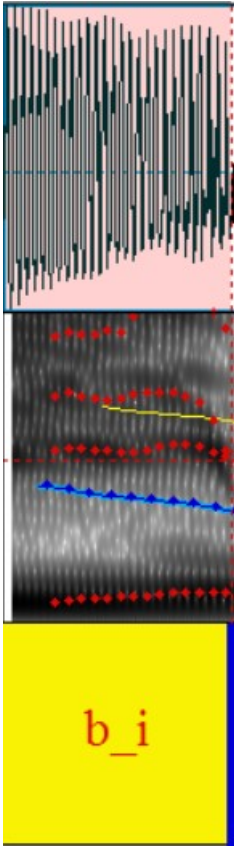
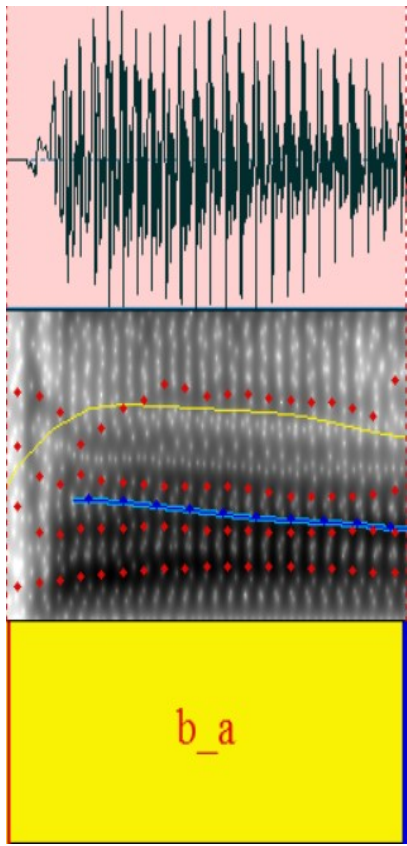
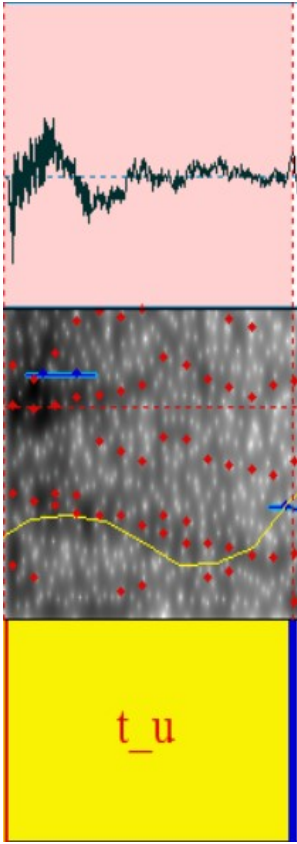
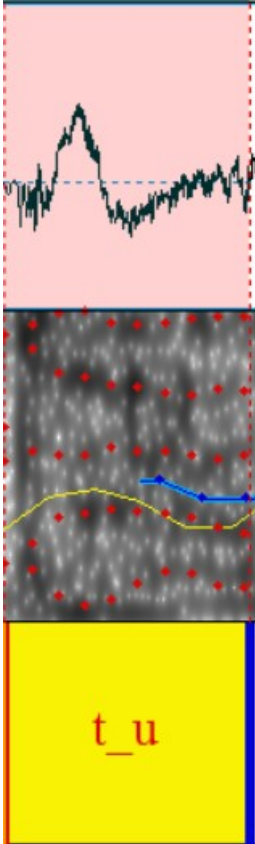
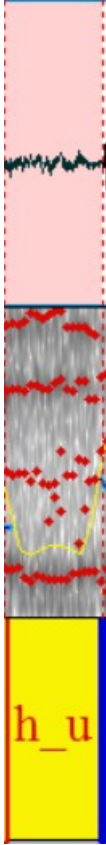
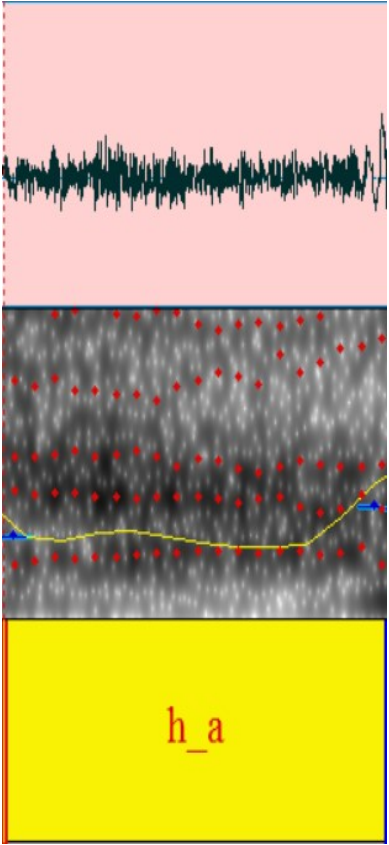
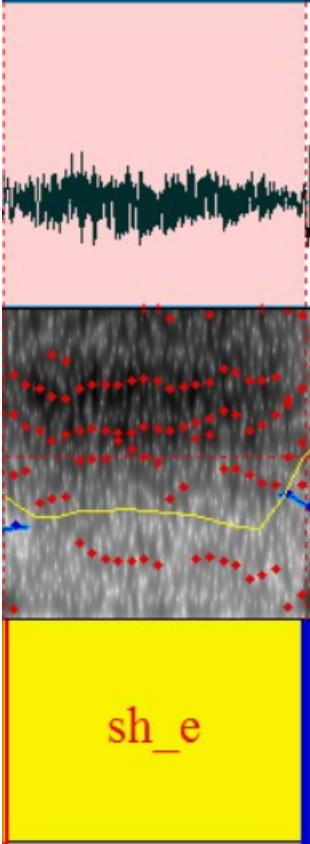
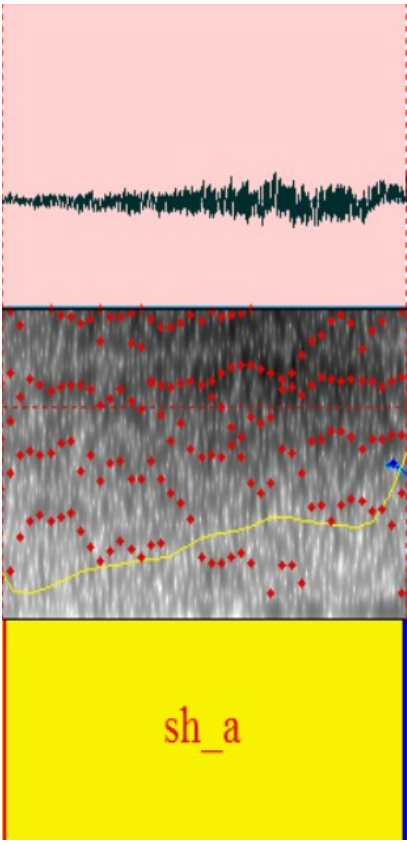
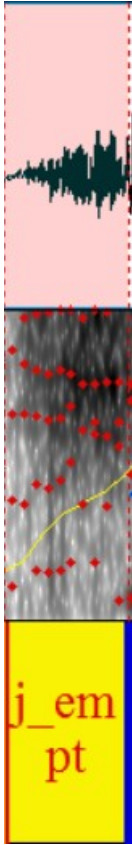
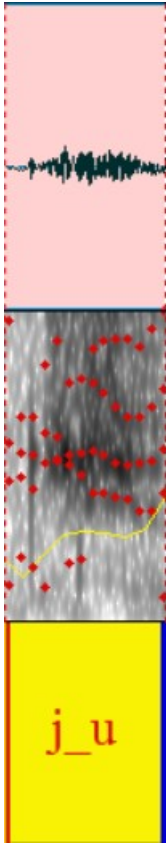


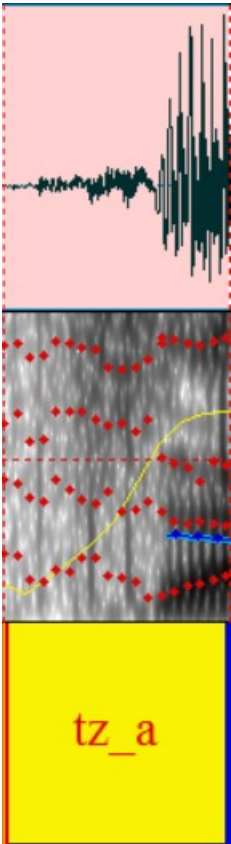
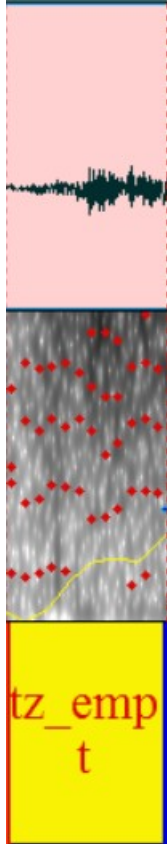
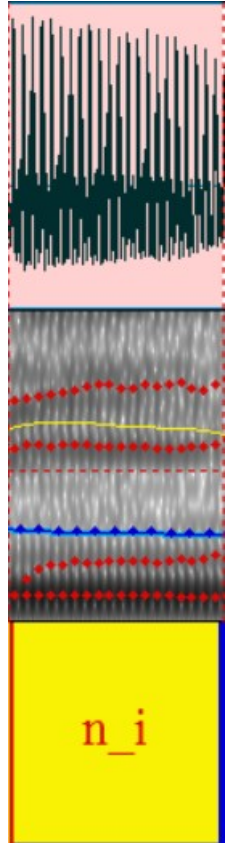
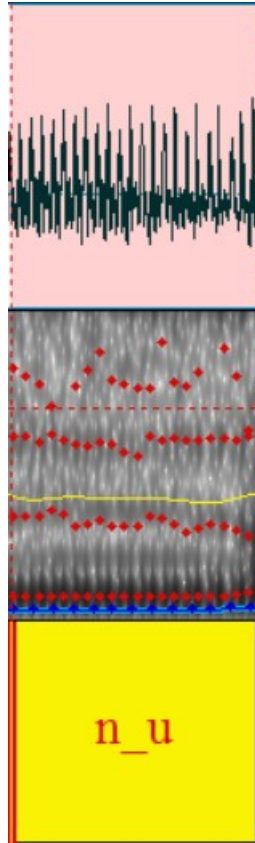
DSP homework 2-2 report

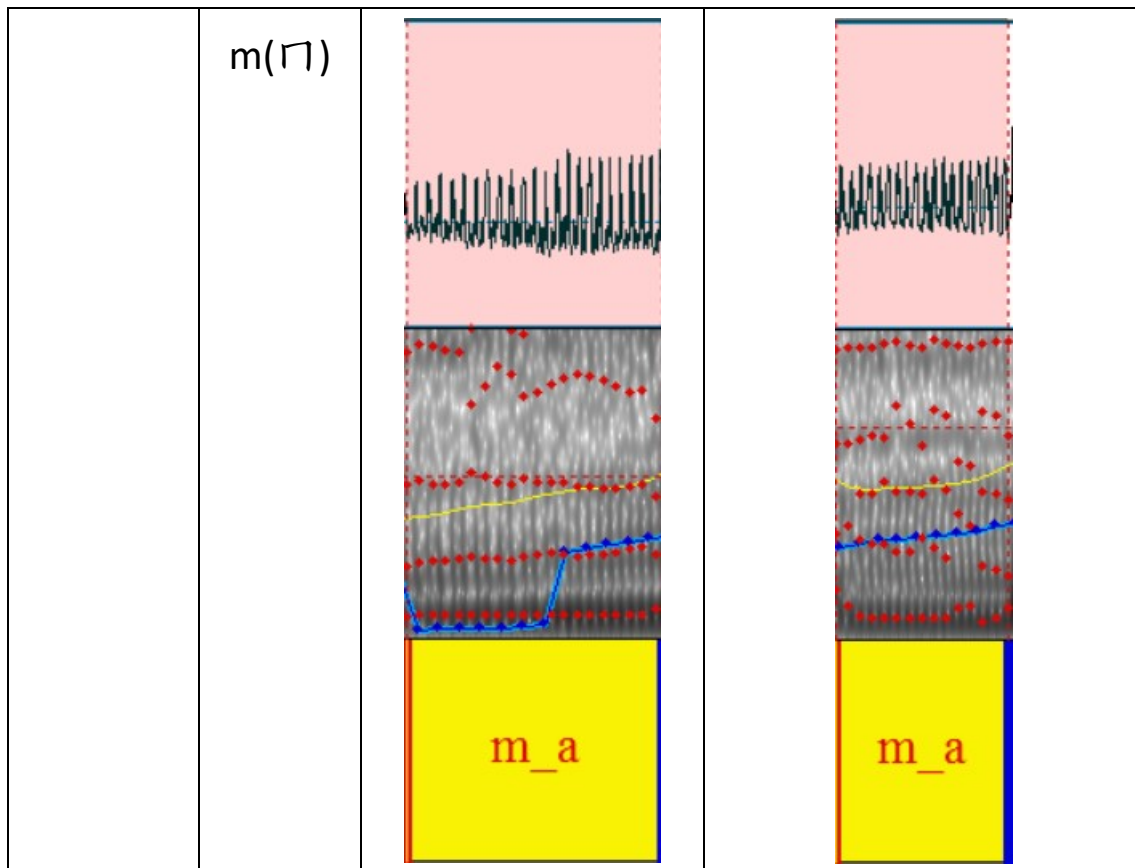
B04902028 資工三 洪浩翔

Phonetic class		figure	
plosive	b(ㄅ)	 <p>The figure for 'b_i' consists of three vertically stacked panels. The top panel is a waveform plot with a pink background, showing a series of high-frequency pulses. The middle panel is a spectrogram with a grey background, displaying horizontal lines of energy. A blue line with diamond markers tracks the formants, and a yellow line with diamond markers tracks the spectral envelope. The bottom panel is a solid yellow rectangle with the label 'b_i' in red text.</p>	 <p>The figure for 'b_a' consists of three vertically stacked panels. The top panel is a waveform plot with a pink background, showing a series of high-frequency pulses. The middle panel is a spectrogram with a grey background, displaying horizontal lines of energy. A blue line with diamond markers tracks the formants, and a yellow line with diamond markers tracks the spectral envelope. The bottom panel is a solid yellow rectangle with the label 'b_a' in red text.</p>

	t(ㄊ)		
fricative	h(ㄏ)		

	sh(尸)		
affricate	j(尸)		

	$\text{tz}(\mathcal{T})$	 <p>tz_a</p>	 <p>tz_emp t</p>
nasal	$\text{n}(\mathcal{Z})$	 <p>n_i</p>	 <p>n_u</p>



1. What are the consistencies of the spectrogram in each phonetic class?

1. **Plosive**: 最前端會出現較混亂無規律的振動，因為發聲時不同於聲帶發聲的方法所導致，如發聲時氣流受阻斷再突然打開，氣流受阻再瞬間釋放也造成在混亂的氣音後會突然出現高峰的現象，在不規律的振動之後接的就是聲帶發聲所形成的規律震動。

2. **Fricative**: 頻譜幾乎整段都是較無規律的，因為發聲時器官收縮造成狹小的通道讓氣流通過摩擦發聲，不同於 **plosive** 的是，**fricative** 因為是讓氣流通過而非突然釋放，所以可以延續混亂的振動。

3. **Affricate**: 頻譜基本上也較無規律，這是因為此類發聲方法是融合 **plosive** 閉塞氣體再釋放以及 **fricative** 用狹小通道發聲的特性，所以也會出現混亂的震動，此外，也會有像 **plosive** 突然出現高峰震動的狀況。

4. **Nasal**: 相較於上三者，**nasal** 的震動較規律，因為 **nasal** 為口腔收縮讓氣流改通過鼻腔，所造成不規律震度會較少。此外，**nasal** 的能量衰減也會較快。

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?

1. 常常是無法明顯分辨的，因為在正常狀況下念一個字時，initial 跟 final 很容易混在一起，尤其是當念得較快的時候。舉例來說，當念“寧”時，基本

上都會 **ning** 一次念過去，而不會像 **ni-ing** 有明顯分開，如此 **ㄣ** 的音就不容易判斷出來。

2. 同上，因為有時候不容易區分 **initial** 跟 **final**，這導致實作時 **initial** 的音聽起來會混進後面的音，也就是後面的音會影響 **initial** 之後的走向，因此讓 **initial** 加上後綴作為表示混在 **initial** 之中後續的音，因為將影響納入本身判斷條件，可以較不加的方式容易判斷而不受影響，可以正確判斷後續音的走向，因此可以增加正確率。

3. What are the differences when pronouncing ㄣ & ㄣ? How can you tell the differences in spectrogram for ㄣ & ㄣ? (You may also want to compare ㄣ & ㄣ, ㄣ & ㄣ respectively)

1. ㄣ & ㄣ: 在發ㄣ時，嘴唇先輕輕閉上，在發聲時瞬間打開讓氣流流出。在發ㄣ時，一樣嘴唇先輕輕閉上，但在讓氣流瞬間流出前，嘴唇先稍微收縮蓄力，再打開讓氣流衝出，如此才有別於ㄣ的發音。至於兩者音頻圖的部分，由上可知兩者都會先一段較無規律的震動，然後再瞬間產生規律的較大震動，但是因為ㄣ在發音時會先嘴唇收縮蓄力，所以在釋放嘴唇衝出空氣的時刻，ㄣ所產生的震動應較ㄣ來的相對大，此外，也因為ㄣ要先蓄力，所以發生較大的震動的時刻會較晚。

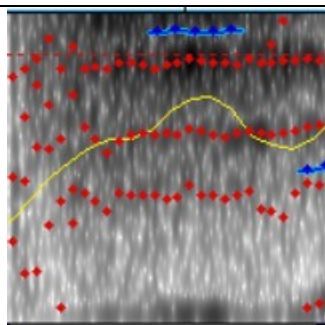
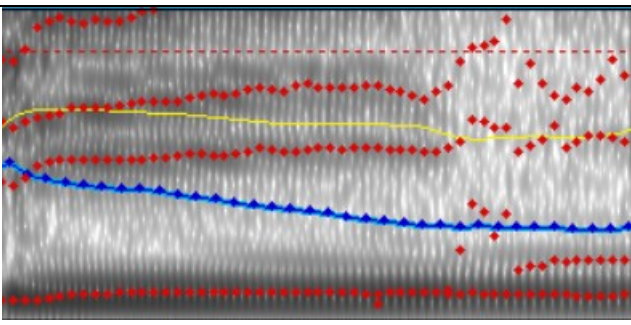
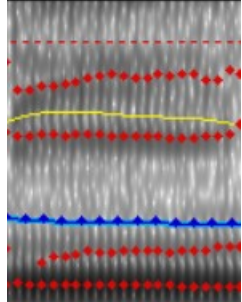
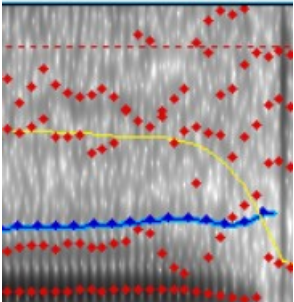
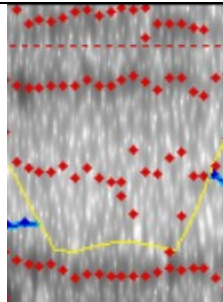
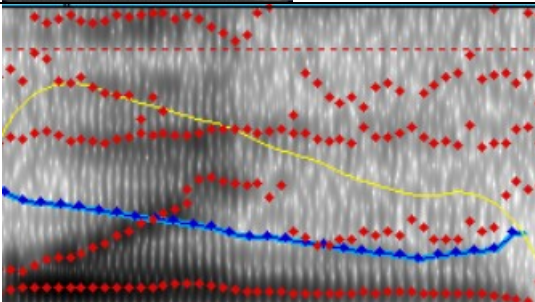
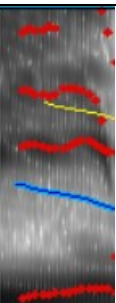
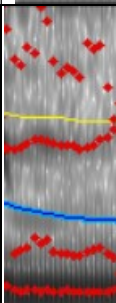
2. ㄣ & ㄣ: 在發ㄣ時，舌頭頂在口腔頂，在發聲時原地輕輕彈一下；在發ㄣ時，舌頭捲曲頂著口腔頂，在發聲時向前彈。如要分辨兩者音頻圖，因為ㄣ在發聲時舌頭捲曲，會比較有蓄力的狀況，所以發聲瞬間相較於ㄣ會有較大的振幅。

3. ㄣ & ㄣ: 在發ㄣ時，舌根跟口腔靠近食道頂擠壓，在發聲時原地輕輕彈一下，瞬間讓氣流通過；在發ㄣ時，舌根一樣擠壓，但如同前兩者一樣收縮，在發聲瞬間往前彈。要分辨兩者音頻圖，也類似上兩個例子一樣，ㄣ會有較大的震動出現。

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

一般狀況來說，要明顯區分 **initial** 跟 **final** 是不容易的，如同在 2-1 所言，常常會有混砸在一起、互相影響的部分，但是透過觀察，還是可以大概指出 **initial** 和 **final** 的位子。若仔細觀察頻譜圖，不難發現 **initial** 在頻譜圖中高頻位置的分布較 **final** 多，而 **final** 則主要峰值出現在低頻，藉由這樣的特色，則大概可以判斷 **initial** 跟 **final** 的位置。

Ex:

Initial	final
	
	
	
	

5. Bonus:

我猜測的電影名稱為“捍衛任務”，原因如下:首先，先收集 2017 年出的電影且為續集者，再來，第二個字最下面兩條 **format** 分開，與“ㄣ”的音非常吻合，因此先猜測有ㄣ在第二個字的名稱，經過過濾，得到的可能剩兩者:“捍衛任務”以及“神鬼奇航”，最後，原圖最後兩個字的 **pitch** 應該是 4 聲，與“任務”兩字相符，因此猜測為捍衛任務。