AUTOMATIC SPEECH RECOGNITION

Lecture 6

Sounds in Languages II
Spectrogram Reading

Outline

- Spectrogram
- Acoustic properties of
  - vowels
  - consonants
    - fricatives
    - stop consonants
    - affricates
    - nasal consonants
  - semi-vowels
- Spectrogram Reading

Stop Consonants

voiced labial stop → /b/ in bus, เบา

voiceless unaspirated labial stop → /p/ in spin, ปืน

voiceless aspirated labial stop → /p^h/ in pen, พาน

voiced alveolar stop → /d/ in den, เดา

voiceless unaspirated alveolar stop → /t/ in star, ตอน

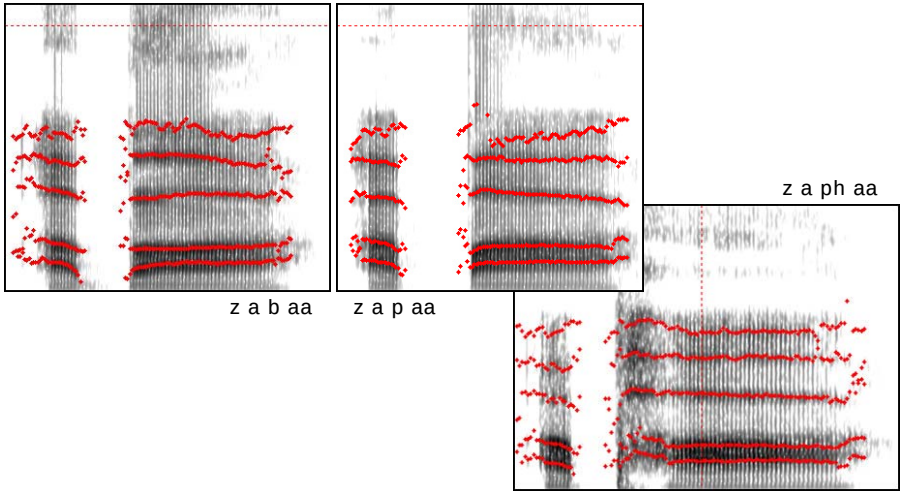
voiceless aspirated alveolar stop → /p^h/ in pen, พาน

voiced velar stop → /g/ in gun, เกา

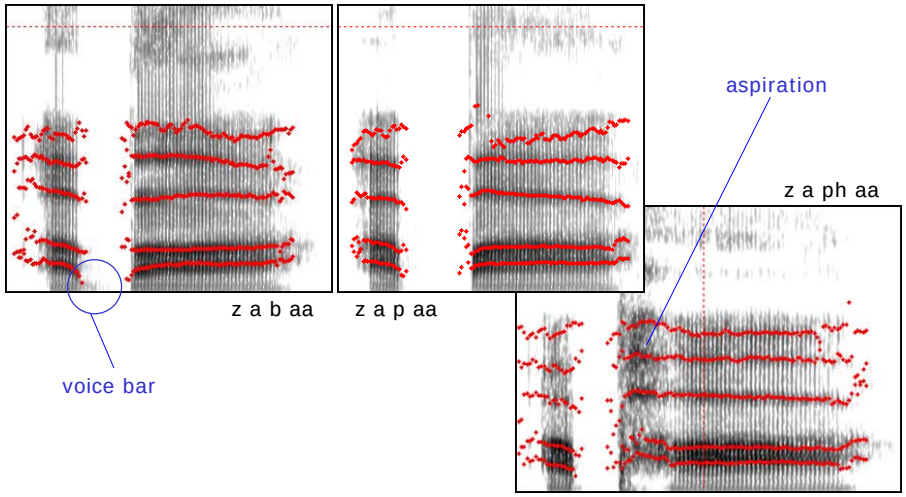
voiceless unaspirated velar stop → /k/ in scar

voiceless aspirated velar stop → /k^h/ in keep, คาน

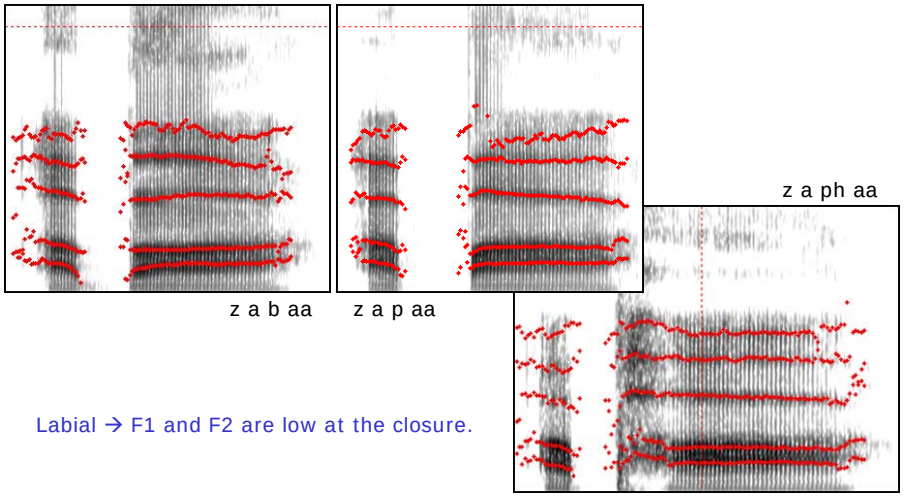
Labial Stop Consonants



Labial Stop Consonants

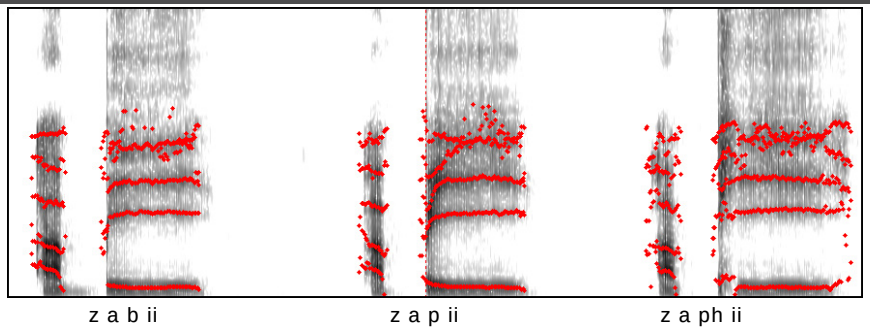


Labial Stop Consonants



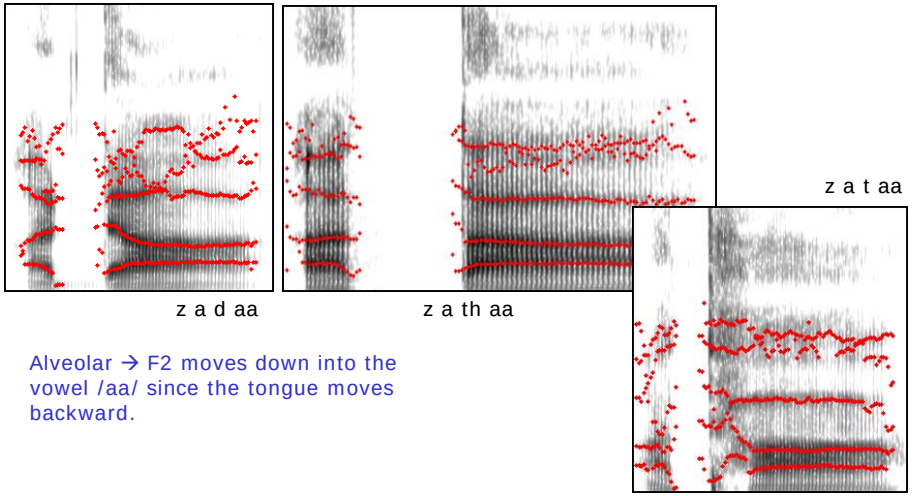
Labial → F1 and F2 are low at the closure.

Labial Stop Consonants



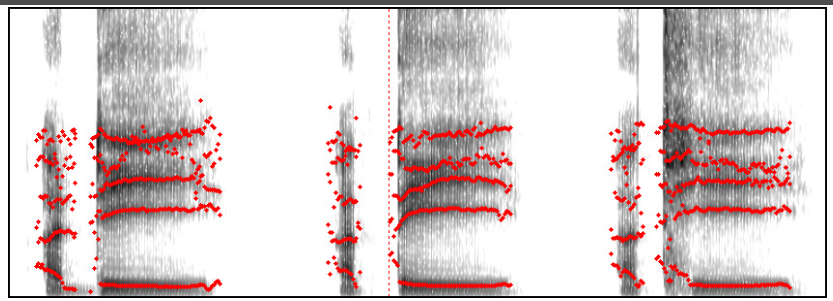
Labial → F1 and F2 are low at the closure.

Alveolar Stop Consonants



Alveolar → F2 moves down into the vowel /aa/ since the tongue moves backward.

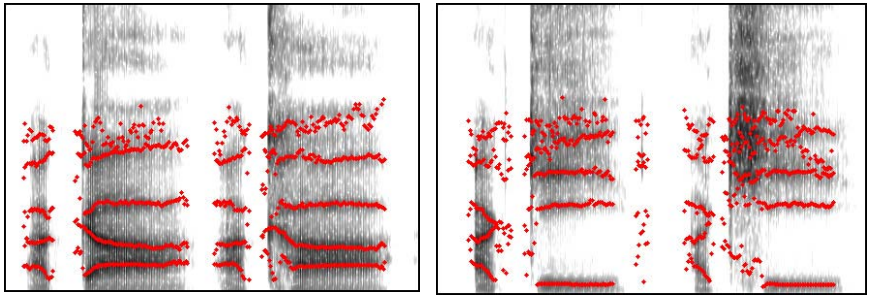
Alveolar Stop Consonants



z a d ii z a t ii z a th ii

Alveolar → F2 moves up into the vowel /ii/ (a little) since the tongue moves a little bit forward from the alveolar stop to the front vowel /ii/.

Velar Stop Consonants



z a g aa z a k aa z a g ii z a k ii

Velar → F2 and F3 tends to move towards each other at the closure.

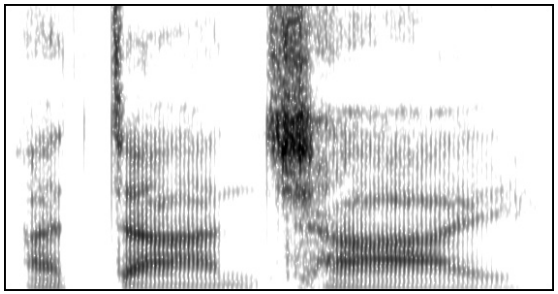
Affricates

- make complete closure like stop consonants
- release the closure and generate the turbulence noise like fricatives

voiceless palatoalveolar affricate →/ch/ in church, ชาน

voiced palatoalveolar affricate → /c/ in judge , จิก

Affricates



z a c aa n^ ch aii

Nasal Consonants

- form a complete closure at some point along the oral cavity
- velopharyngeal port is open during the closure
- no pressure increase behind the constriction
- side branch introduce a pole/zero pair in the spectrum
- decrease in signal amplitude due to loss in the nasal cavity

Nasal Consonants

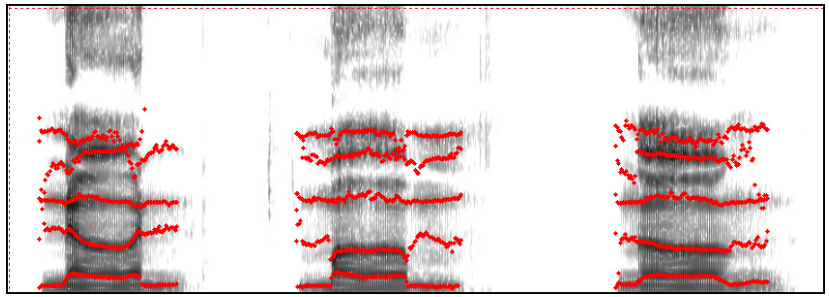
- Three places of articulation like stops

labial nasal → /m/ in man, มาก

alveolar nasal → /n/ in not, นาน

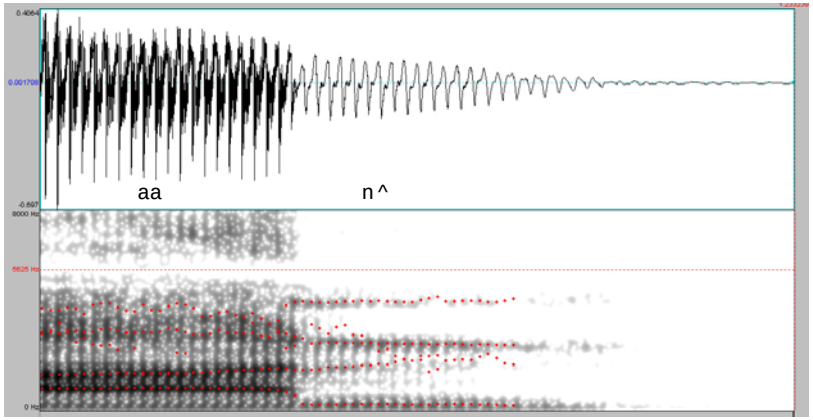
velar nasal → /ng/ in sing, เงา

Nasal Consonants



n qq n^ m qq m^ ng qq ng^

Nasal Consonants



Consonant Summary

Place of Articulation

Manner of Articulation	Labial	Dental	Alveolar	Palatal	Velar
Stop	p b		t d		k g
Fricative	f v	T D	s z	S Z	
Nasal	m		n		ng

voiceless voiced

Classes of Sound

- Vowels
- Consonants
- Semi-vowels

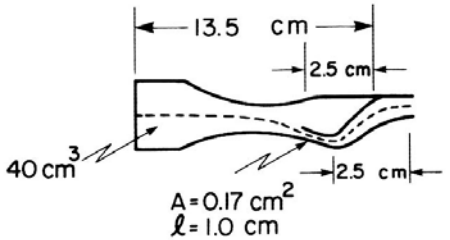
Semi-vowels

- Constriction in the vocal tract
- Slower articulatory motion than other consonants
- Extreme articulation of some vowels
- i.e:
  - laterals (eg. /l/ ลาง, LAN)
  - trills (eg. /r/ ราง)
  - retroflex (English) (eg. /r*/ ring)
  - glide (eg. /j/ ยาว, year)
  - aspirant (eg. /h/ หอน, hat)

} Liquids

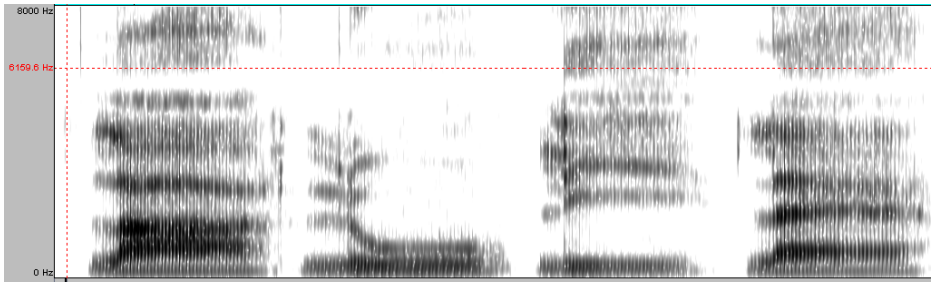
Laterals

- constriction is produced with the tongue blade in contact with alveolar ridge in the midline
- the lateral edges of the tongue do not come in contact with the hard palate → forming side branches



lateral → /l/ in lay, ลาน

Laterals



l aa l uu l ii l xx

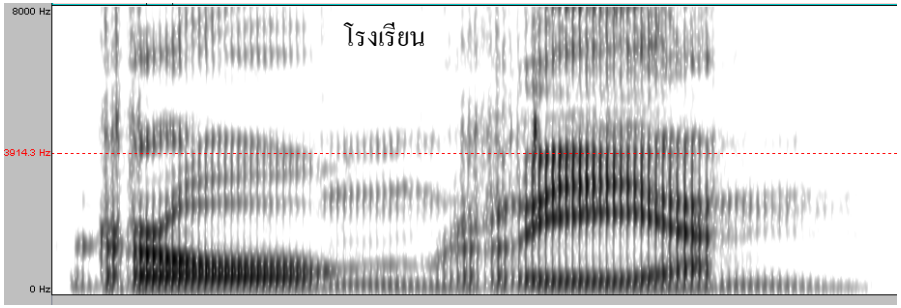
Low F1, Mid to Low F2
Abrupt change in amplitude

Trills

- the tongue blade touches the alveolar ridge and immediately moves away
- move back-and-forth repeatedly (in Thai /r/, if pronounced carefully)

Trill → /r/ in ร<u>า</u>ย

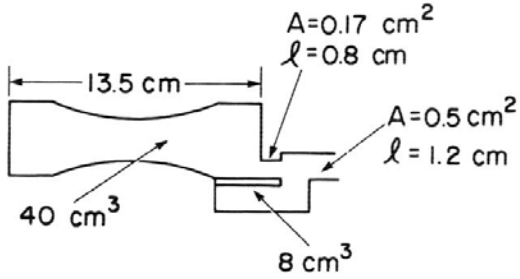
Trills



On-off energy
Very low F3

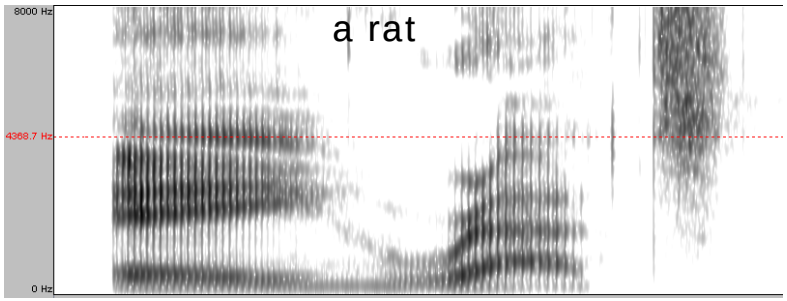
Retroflex

- tongue curled back against the palate



Retroflex consonant → /r*/ in <u>r</u>un

Retroflex



Very low F3

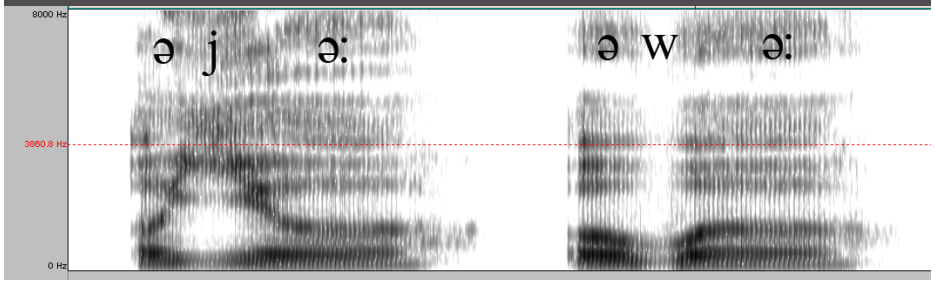
Glides

- extreme version of some vowel
- extreme of /i/→/j/
- extreme of /u/→/w/

Glide /j/ → year, ยาว

Glide /w/ → wear, วาว

Glides



Very low F1
Very high F2

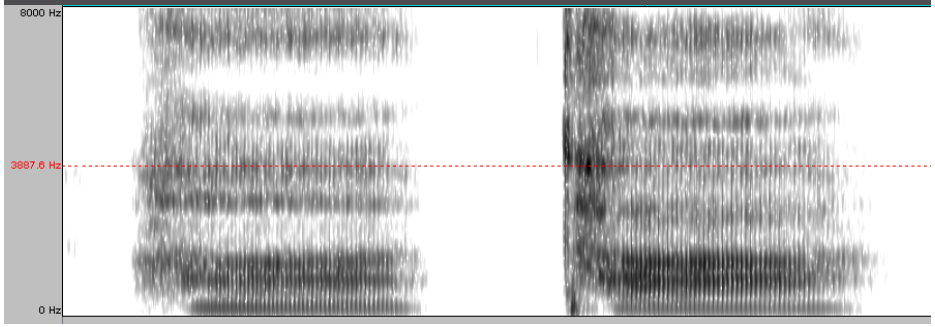
F1, F2 low

Aspirant

- air flow through ‘spread’ glottis
- rapid airflow generates turbulence noise at the glottis
- Glottal fricative

aspiration → /h/ in home, หาม

Aspirant

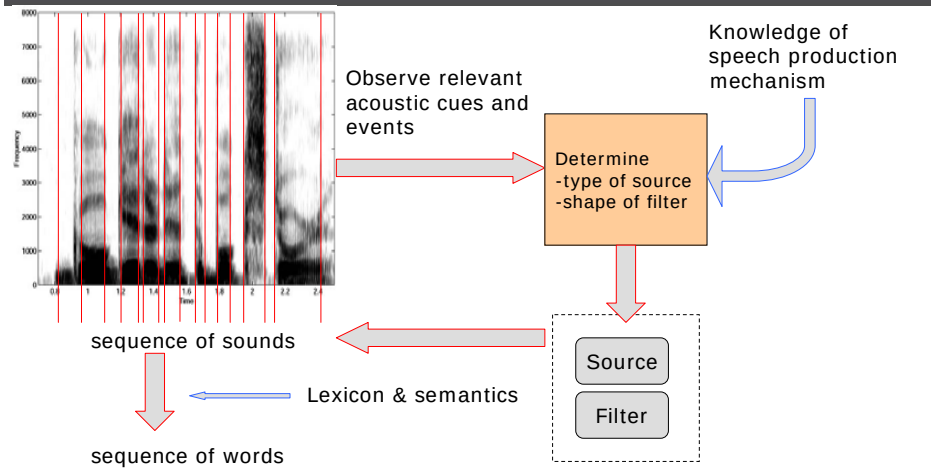


h aa th aa

Spectrogram Reading

- use knowledge of speech production to determine the underlying word sequence corresponding to the speech signal by visually looking at the spectrogram

Spectrogram Reading



Guideline for Spectrogram Reading

1) Determine acoustic boundaries
2) Determine classes of sound

signal → segments

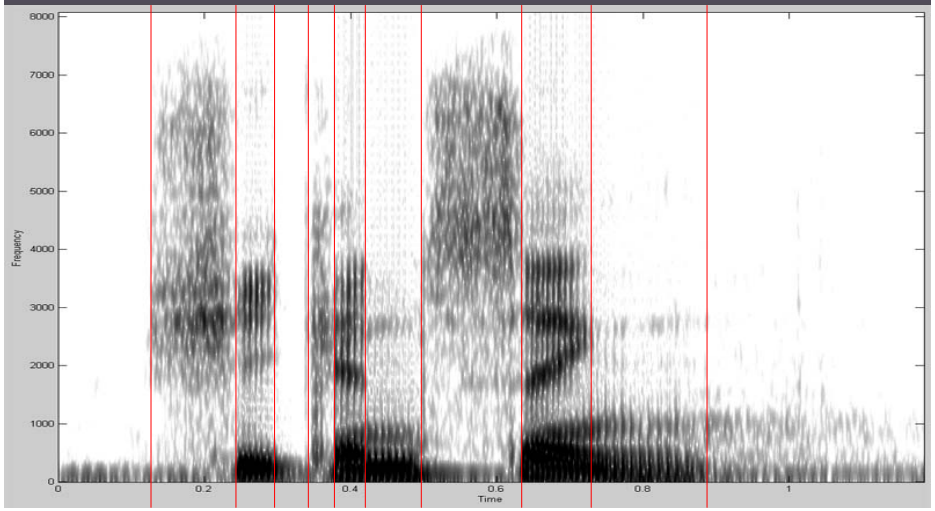
3) Determine more specific manner / place / voicing

segments → sounds (phonemes)

4) Use semantics to verify what are found in earlier steps and determine the final sentence

sounds (phonemes) → sentence (or words)

Step 1 - Determine Acoustic Boundaries

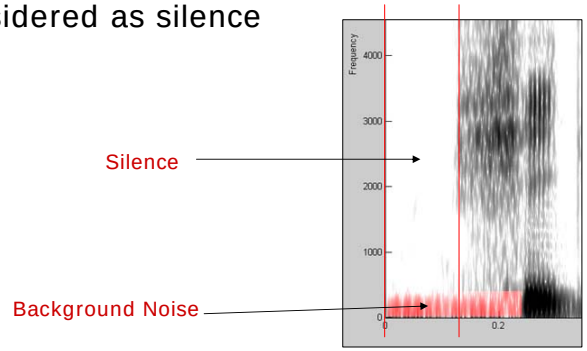


Step 2 - Determine Classes of Sound

- Silence
- Vowels
- Consonants
  - Stop consonant
  - Fricatives
  - Affricates
  - Nasal consonants
- Semi-vowels

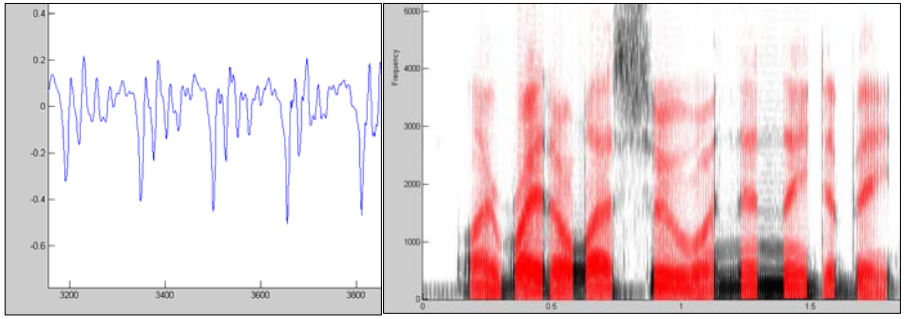
Silence

- no energy (except for some low-frequency background noise)
- closure of stop consonants or affricates are not considered as silence



Acoustic Cues for Vowels

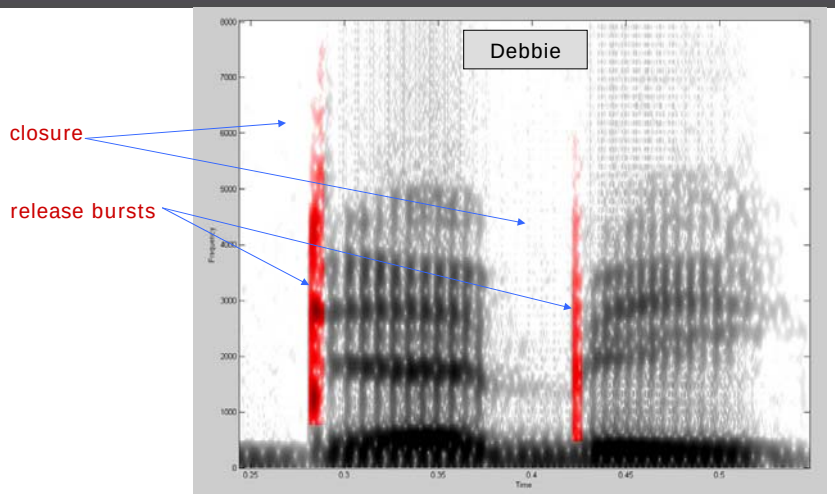
- high energy / maximal airflow
- clear formant structure
- periodic signal (easily seen in time domain)



Acoustic Cues for Stop consonants

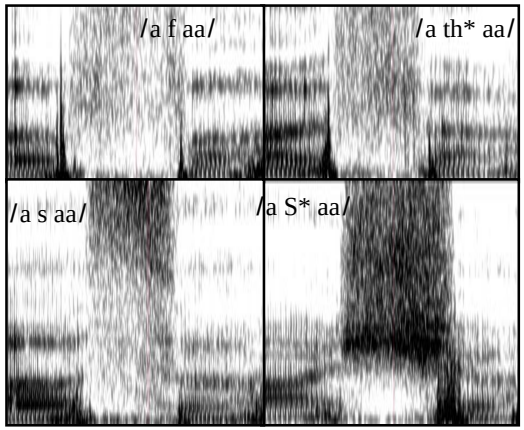
- no energy in the closure, except for voice bar
- after release, there might be release burst
- if 'spread' glottis, there is aspiration noise

Stop Consonants



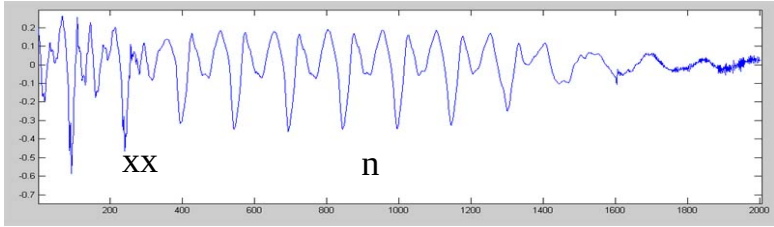
Acoustic Cues for Fricatives

- shaped noise
- Affricates
= stop + fricative

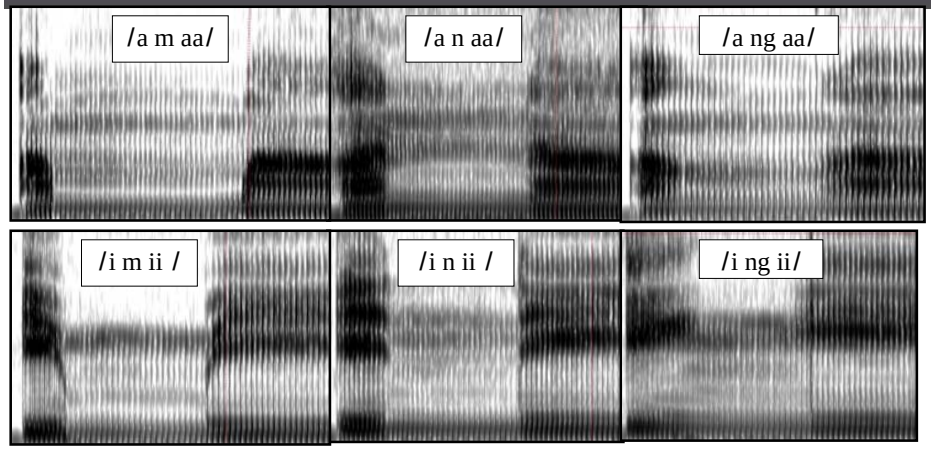


Acoustic Cues for Nasal Consonants

- loss of mid-high freq. energy due to loss in the nasal cavity
- nasalized vowel
  - damping
  - F1 bandwidth increases

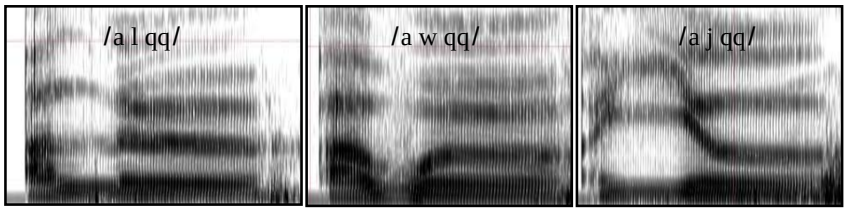


nasals

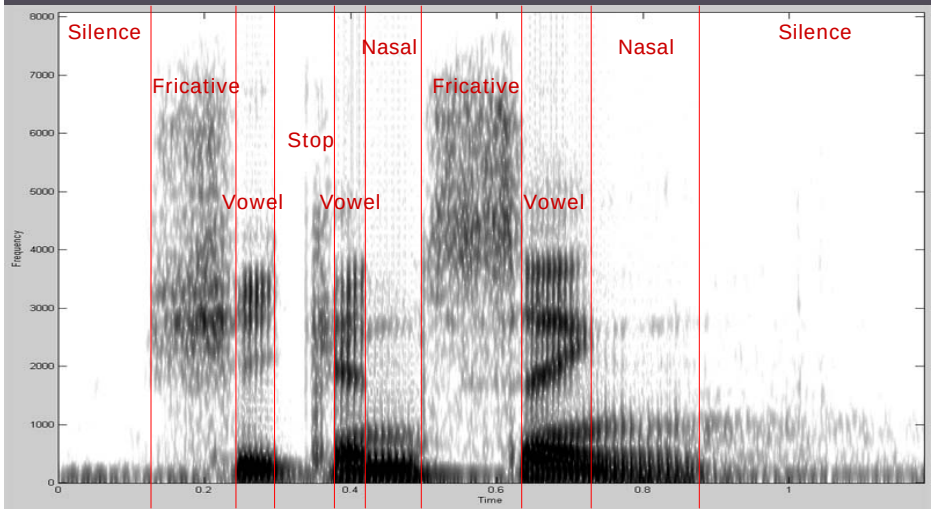


Acoustic Cues for Semi-vowels

- formant movement is more rapid than vowels but not as abrupt as consonants
- might involve abrupt change in the amplitude of some formants



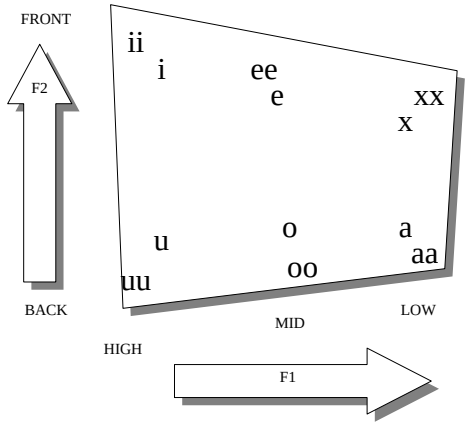
Step 2 - Determine Classes of Sound



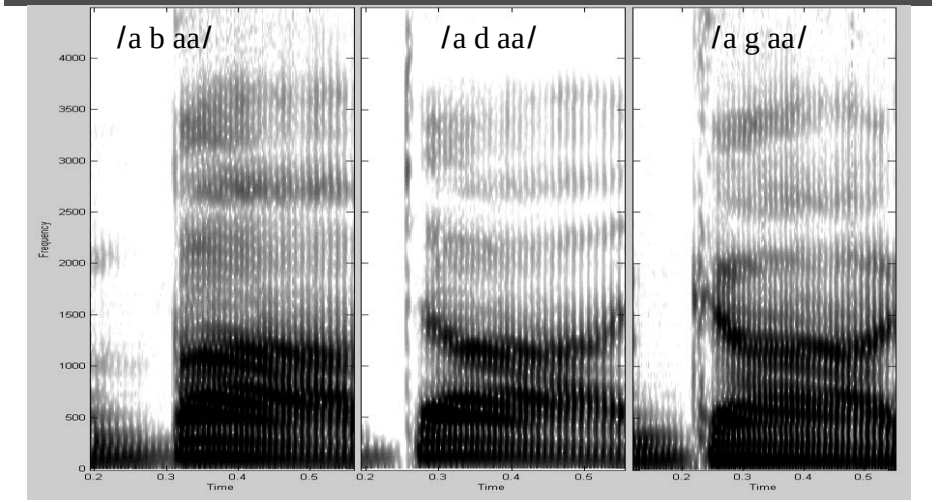
Step 3 - Determine Manner / Place / Voicing

- Vowel → F1 and F2 → front/back, high/mid/low
- Stop
  - → burst shape & formant movement → place
  - → voice bar & VOT → voicing
- Fricative
  - → noise shape & formant movement → place
  - → voice bar & duration → voicing
- nasal → formant movement → place
- semi-vowel
  - → movement & amplitude of F1,F2,F3 → type

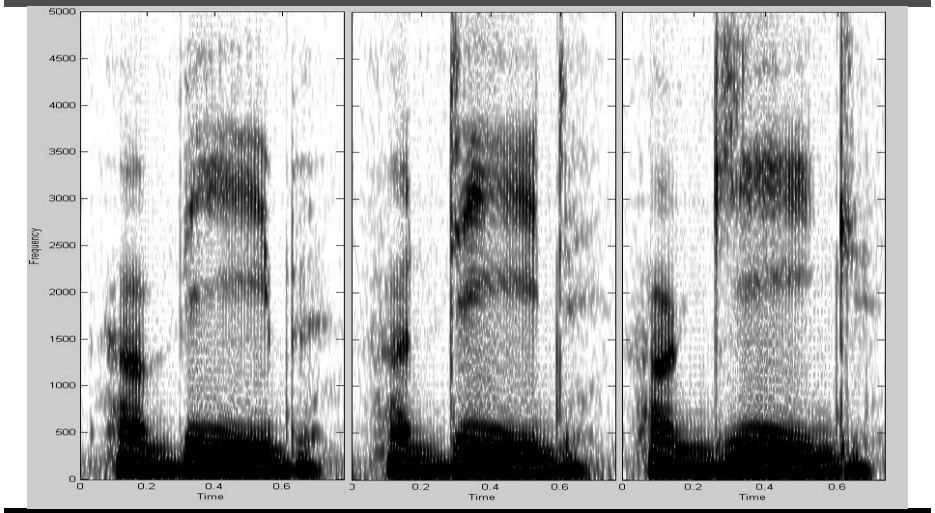
Vowel Frontness & Height



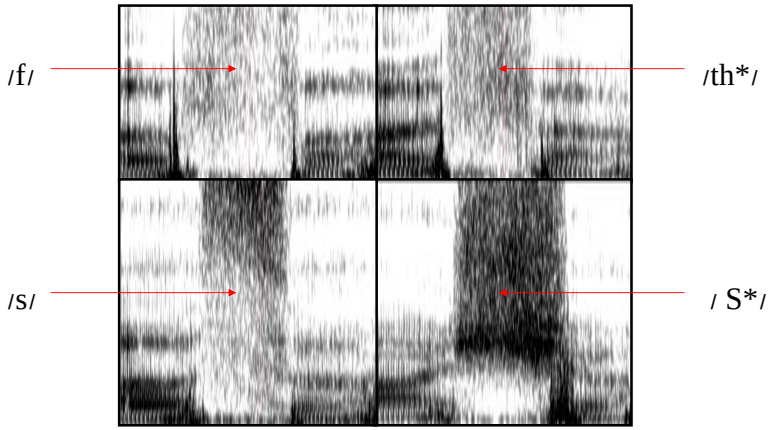
Stop Place of Articulation



Stop Place of Articulation



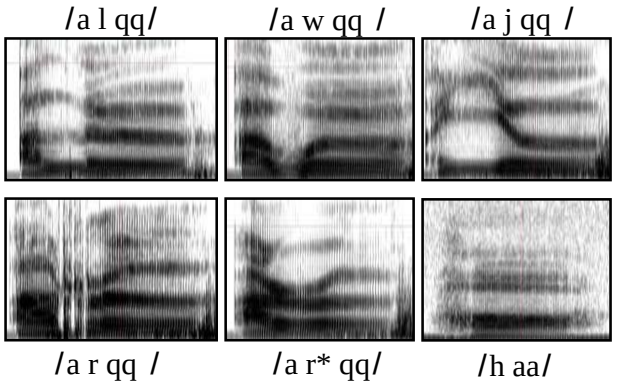
Fricative Place of Articulation



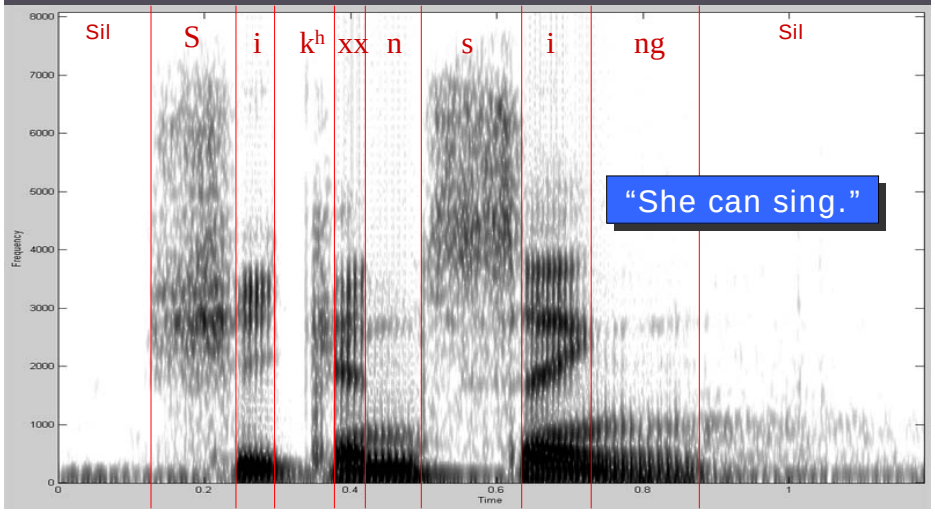
Nasal Place of Articulation

- look at the movement of the formants into/out of adjacent vowels
- same technique as determining stop consonant place of articulation from formant movement

Semi-vowels



Step 3 & 4



Spontaneous vs. Citation