AUTOMATIC SPEECH RECOGNITION

Lecture 6

Sounds in Languages II
Spectrogram Reading

Last Lecture

• Spectrogram
• Acoustic properties of
  – vowels
  – fricatives
  – stop consonants
  – affricates
  – nasal consonants

This Lecture

• Other classes of sound
  – semi-vowel
• Spectrogram reading

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/ LAN)
    - trills (eg. /r/ รู)
    - retroflex (English) (eg. /r/ ring)
    - glide (eg. /j/ ย)ear
    - aspirant (eg. /h/ ฮ)at

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

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 ลาน, 泰

Laterals

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 รู
Trills

- On-off energy
- Very low F3

Retroflex

- Tongue curled back against the palate

\[A = 0.17 \text{ cm}^2 \]
\[l = 0.8 \text{ cm} \]
\[A = 0.5 \text{ cm}^2 \]
\[l = 1.2 \text{ cm} \]

Retroflex consonant → /r*\ in run

Retroflex

- A rat

Glides

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

Glide /j/ → year, ย
Glide /w/ → wear, ว
Aspirant

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

aspiration → /h/ in home, ห าม

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

- Observe relevant acoustic cues and events
- Determine type of source and shape of filter
- Knowledge of speech production mechanism

Guideline for Spectrogram Reading

1) Determine acoustic boundaries
2) Determine classes of sound
3) Determine more specific manner / place / voicing
4) Use semantics to verify what are found in earlier steps and determine the final sentence

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
- closure
- release bursts
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

Acoustic Cues for Nasals

- 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

- Silence
- Fricative
- Nasal
- Stop
- Vowel

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

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

/\textit{f}/  /\textit{th}/  /\textit{s}/  /\textit{S}/

/\textit{a l qq}/  /\textit{a w qq}/  /\textit{a j qq}/  /\textit{a r qq}/  /\textit{a r^* qq}/  /\textit{h aa}/
Step 3 & 4

“She can sing.”

Spontaneous vs. Citation