Nature of a Speech Signal

1- gross level : |m|;|a|;|s|;|t|;|t|;

(constant silence but |s| ≠ |t|, it will find like "masks")

Relative amplitudes

Amplitude, periodicity, pitch, noise, articulation pattern

2- Source & System level

Nasal stop, vocal fold vibration (time varying excitation)

Time varying vocal tract system

- Nasal resonances
- Vocal tract resonances

|t| → burst + resonances

|p| → no resonances

|a| => variation in the resonances at initial, central & end regions due to the influence of adjacent units

Supra segmental

Cursive script & speech are analogous

Constraints at different levels

- production level (we can not produce the speech at 5 Hz of 60 Hz movements of articulators)
- Legal sound units of the language
- Supra segmental knowledge constraints (can be very in limited way)
- Language constraints in finding sequence of sounds, words, sentences, etc --

=> All these constraints provides robustness in speech communication

Some important points

- Trying to produce speech by newly born by child
- Speech vs audio, image & video
- Speech has unique features as human enclosed with both production & perception
- Most of the speech systems need to have the knowledge of both production & perception
- A person is deaf (no perception capability) & can child may not produce speech (i.e. production & perception are interrelated)
- Ability of humans can get the meaning of different utterances from the power of prosody
- Cold countries — Languages will be consonant based
  hot countries — Languages will be vowel based
- Raising a voice VS. changing the volume

Knowledge Sources in Speech

Text transcription

Speech signal

Labeling details & sequence of sound units

Duration patterns

Energy contour

Pitch contour < Local / global

Spectrogram

Speech information at different levels

Segmental level (10– 30 ms)

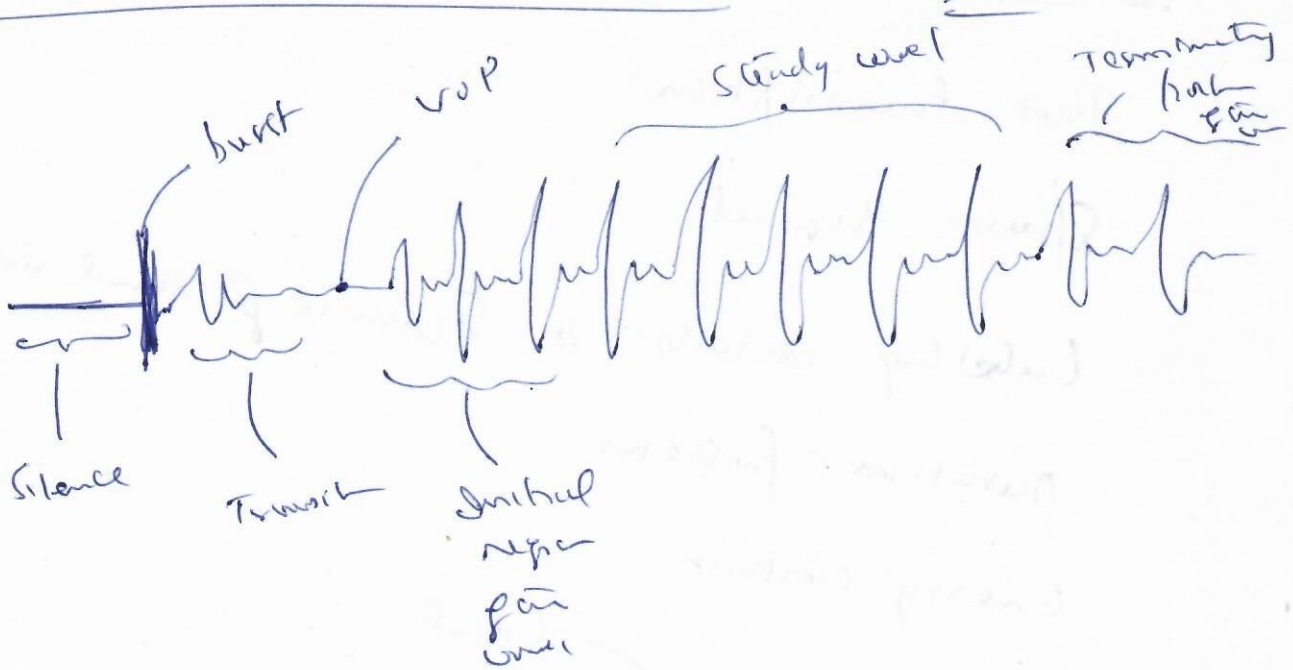
- periodicity & pitch, voiced & unvoiced
- Spectral envelope
- Formant structure

Supra-segmental level (100 – 300 ms)

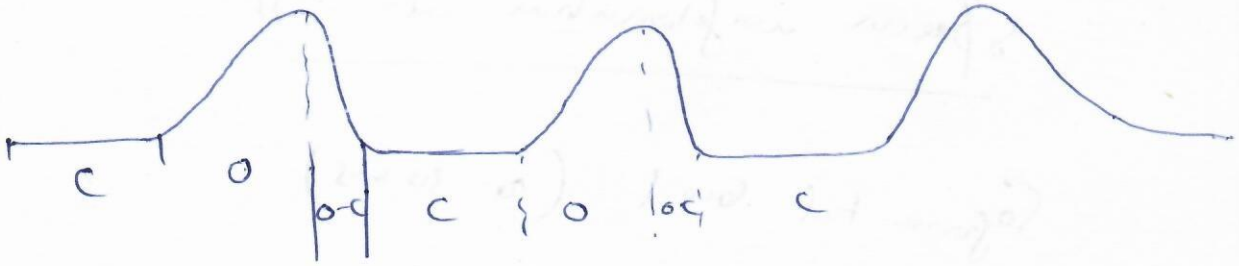
- Duration patterns < different regions within a syllable / duration of different units in the utterance
- co-articulation
- Pitch contour
- Energy contour
- Stress & prominence

prosodic information ⇒ conversational speech in degradation ...

Different regions in the syllable "ka"



Subsegmental level (1 - 3 ms)



Importance of subsegmental information

- coarticulation: influence of adjacent units in articulating the present unit.
- Ability of humans to cope at the ambiguity at different distances from the source of production
- Example: written and subsegmental
  Trying to produce speech by reading them by the