

Knowledge Sources in Speech

The slide features a white background with the title 'Knowledge Sources in Speech' in a blue sans-serif font. At the bottom, there are decorative blue elements: a horizontal bar on the left with a gradient and a small icon of a pen nib, and a large, abstract blue shape on the right that resembles a stylized wave or a speech bubble tail.

Knowledge Sources in Speech

- **Message:** Thought to be conveyed
- **Speaker:** Who the speaker is?
- **Language:** Language in which speech is produced
- **Naturalness:** Pleasing quality of speech
- **Intonation:** Increasing/decreasing pitch
- **Duration:** Variations in the duration patterns
- **Stress:** Uttering with special attention (words or syllables)
- **Accent:** Dialect region of person
- **Emotions (mood):** State of mind
- **Health:** Condition of speech production organs
- **Device and Channel:** Microphone and medium in which speech is collected

Illustration of knowledge sources in speech

Text : maataa aur pitaa kaa aadar karnaa chaahie

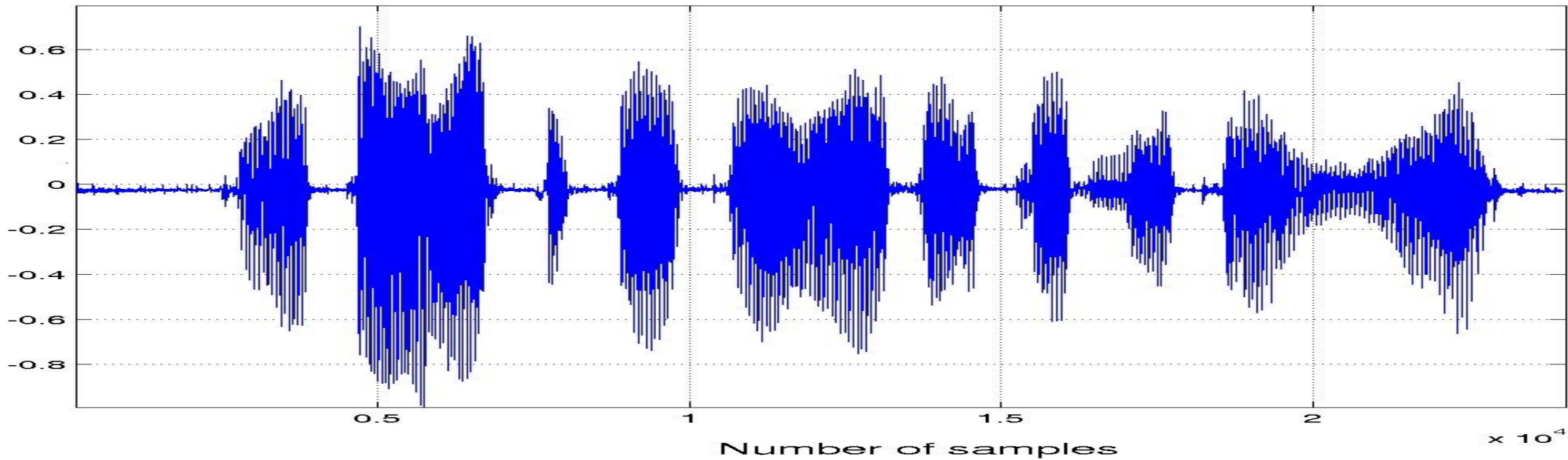
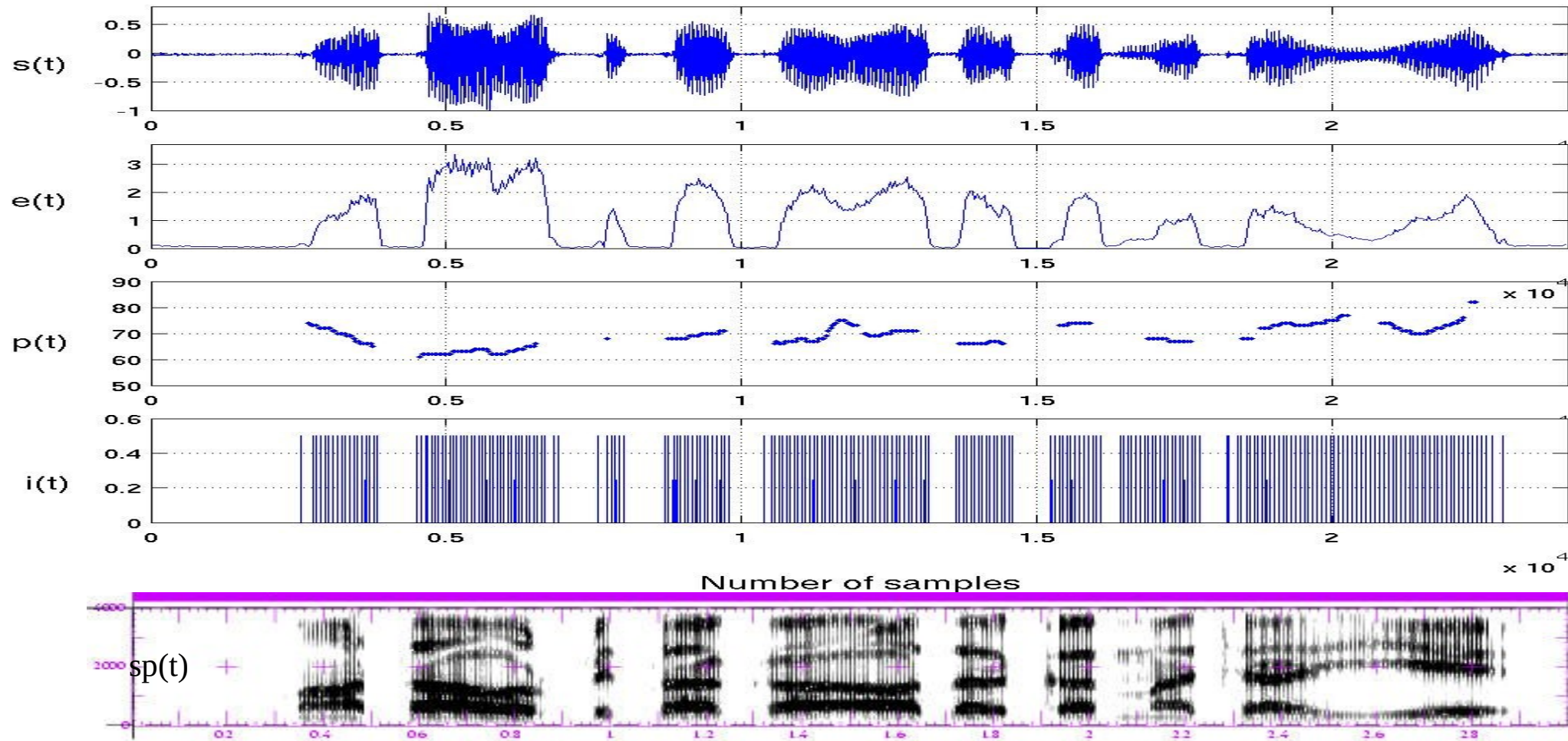
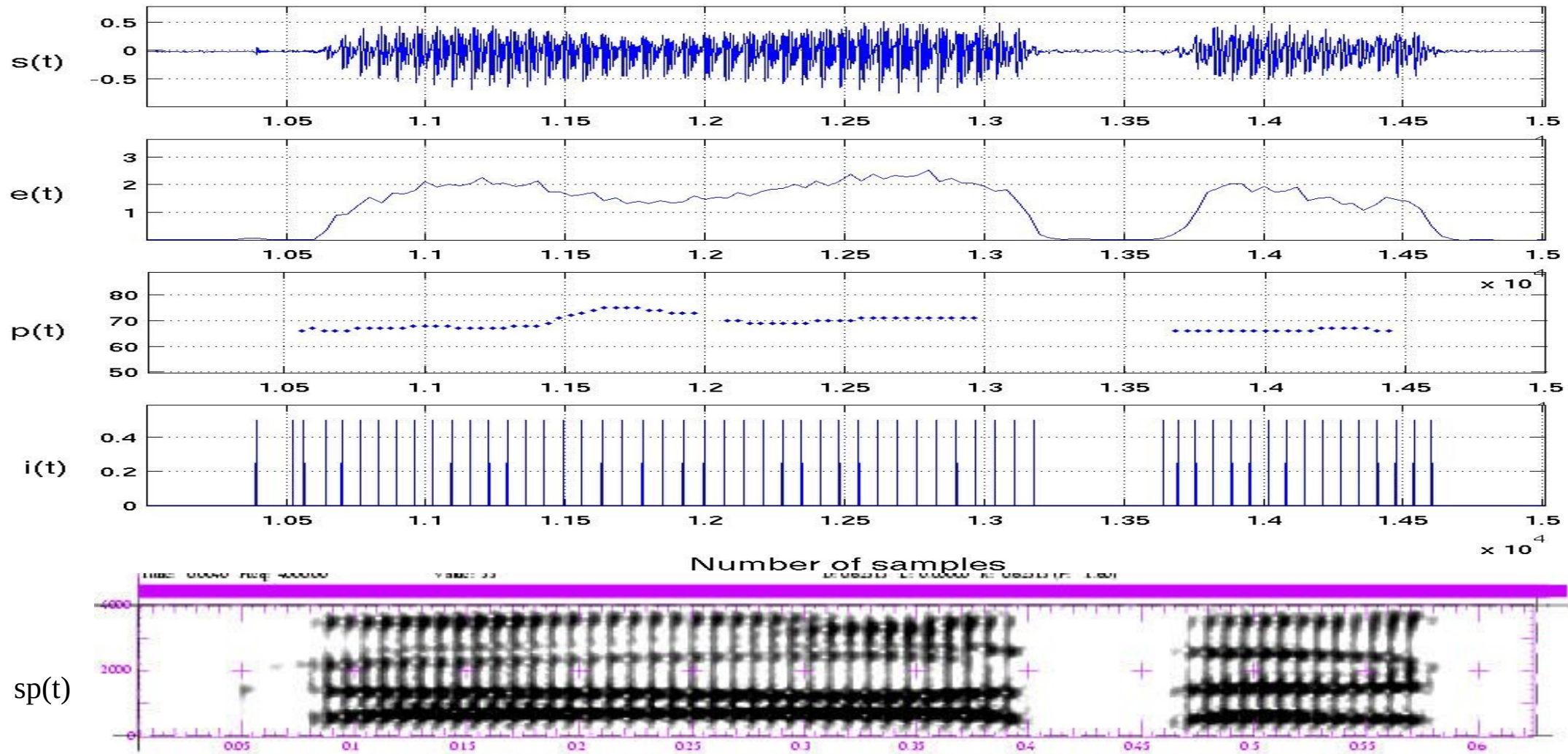


Illustration of knowledge sources in speech (Contd..)



$s(t)$: speech waveform, $e(t)$: energy contour, $p(t)$: pitch contour, $i(t)$: instants of significant excitation and $sp(t)$: spectrogram

Illustration of knowledge sources in speech (Contd..)



$s(t)$: speech waveform, $e(t)$: energy contour, $p(t)$: pitch contour, $i(t)$: instants of significant excitation and $sp(t)$: spectrogram

Knowledge sources in speech (analysis point of view)

- **Segmental level (10-30 ms)** : Positioning and movement of articulators, shapes of cavities (oral and nasal), voiced and unvoiced excitation, periodicity (pitch) and formant structure
- **Sub-segmental level (3-5 ms)**: Glottal pulse shape, open and closure regions of glottis, consonant regions (burst and transition regions)
- **Supra-segmental level (>100 ms)** : Prosody (duration and pitch), stress, prominence, melody, syntax and semantics.
- **Coarticulation level** : Constraints due to linguistic context: Influence of the adjacent units in the articulation of the present unit.

Knowledge sources in speech (analysis point of view)

Information in speech at different levels

1. Segmental level (10 - 30 ms)
 - periodicity (pitch) ✓
 - voiced / unvoiced
 - Formant structure
 - Shape of the vocal tract
2. Supra segmental level (100 - 300 ms)
> 100 ms
Prosodic level

Knowledge sources in speech (analysis point of view)

2. Subsegmental level (Prosodic level)

— Durations patterns — gross level
— finer level

— Intonation (Pitch contour)

— Energy contour

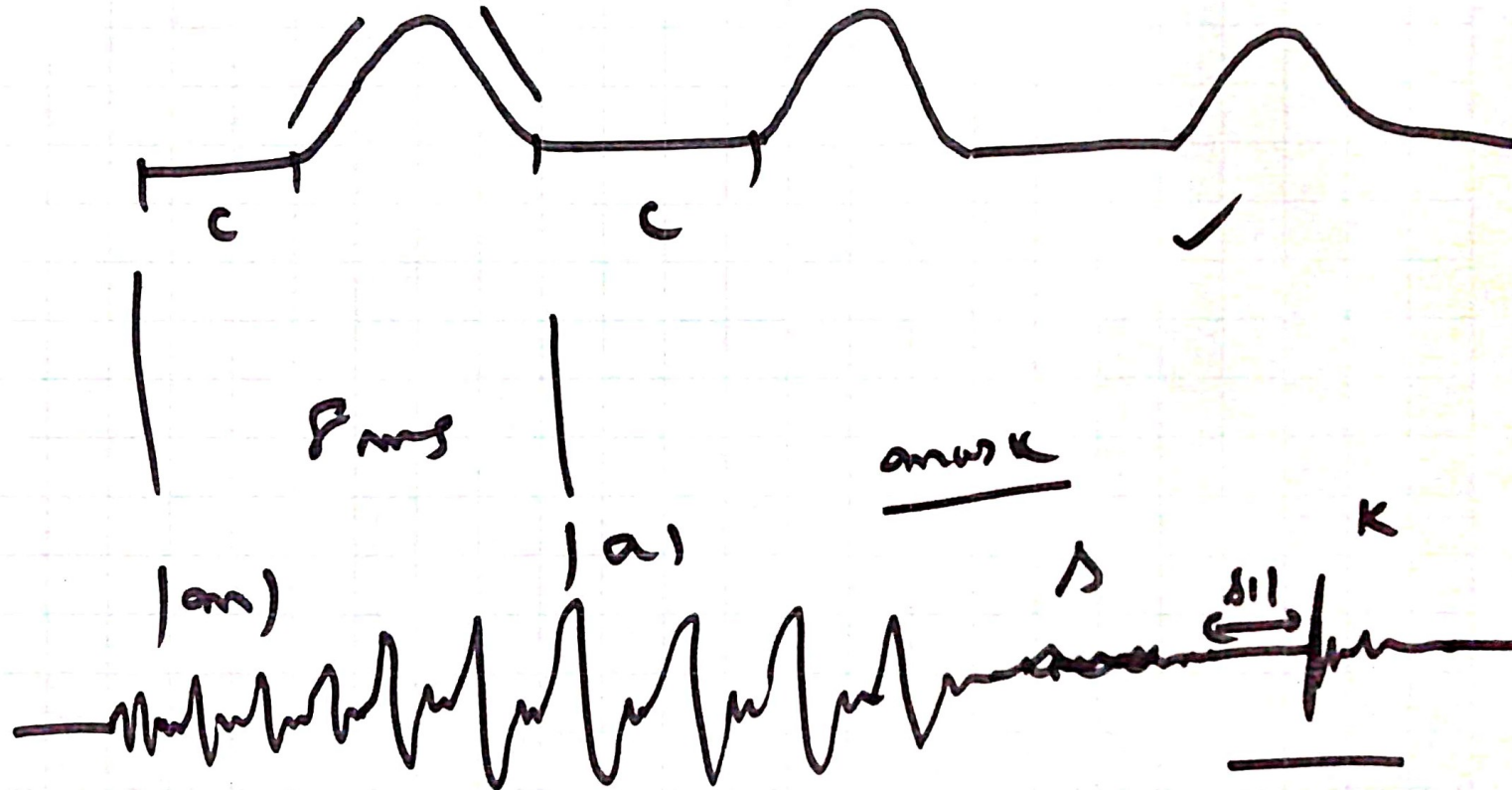
— Co-articulation

— Stress & prominence

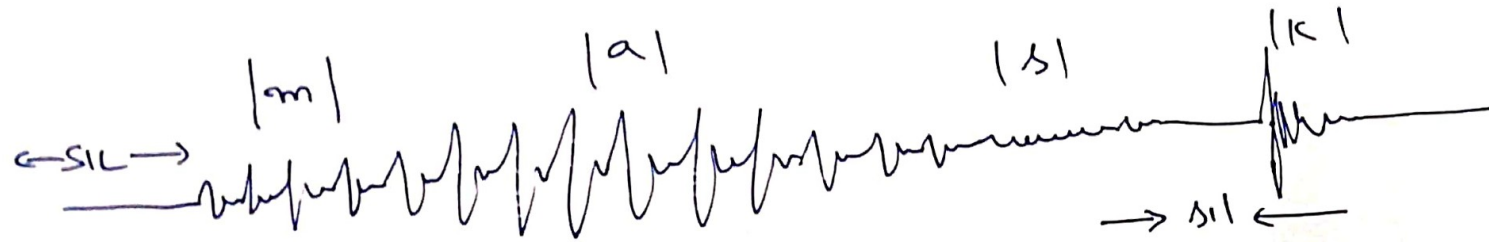
Knowledge sources in speech (analysis point of view)

Subsegmental level (3-5ms)

I.I.T. KGP



Nature of a Speech Signal



Gross Level: $|sil|$; $|m|$; $|a|$; $|s|$; $|sil|$; $|k|$

Relative amplitudes, Periodicity, Noise, Burst, duration patterns

Source & System Level

Nasal stops, Vocal fold vibration

Time-varying vocal tract system: Nasal & VT resonances

$|k|$ → Silence + Burst + Resonances

$|s|$ → Silence + Burst

$|a|$ → Variation of Resonances at initial, middle & End
Co-articulation