Acoustic Phonetics

Analysing a Communication Medium

Dafydd Gibbon

Guangzhou, Autumn 2019

Conclusion

After studying this unit ...

… you should know the basic semiotic and physical foundations on which rhythms and melodies of speech are based
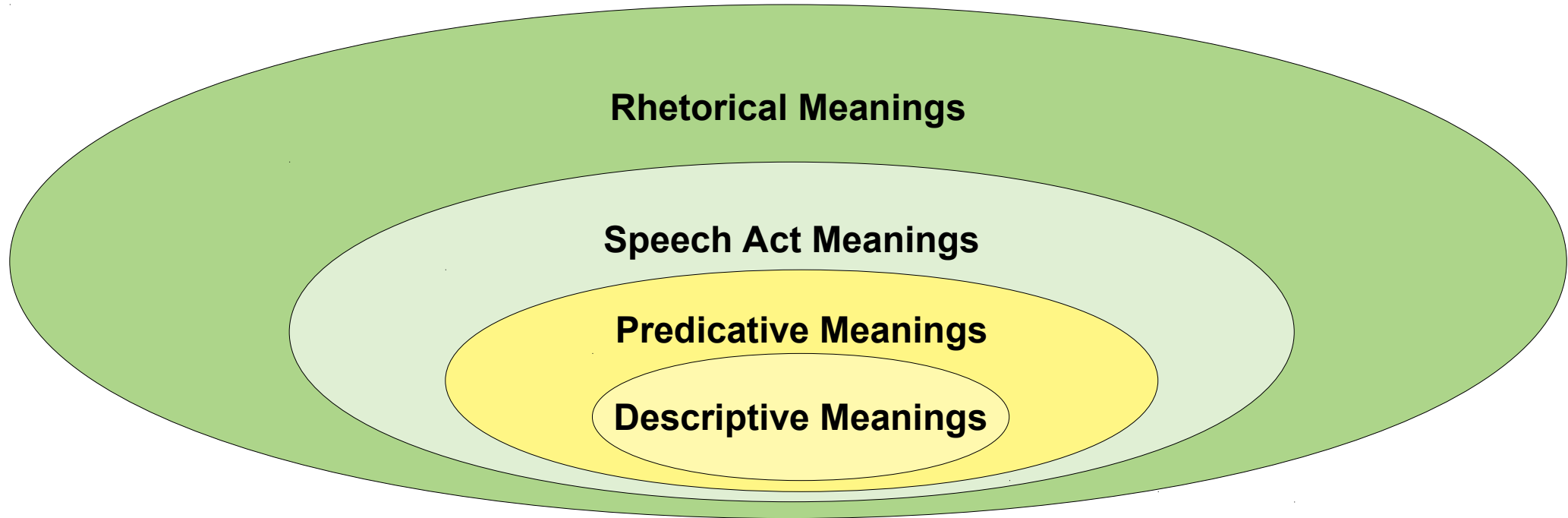
Communication Basics

- Semiotics: the study of signs and their meanings

  - Index: signifies something with a particular time, place or cause

  - Icon: signifies something which is similar to itself

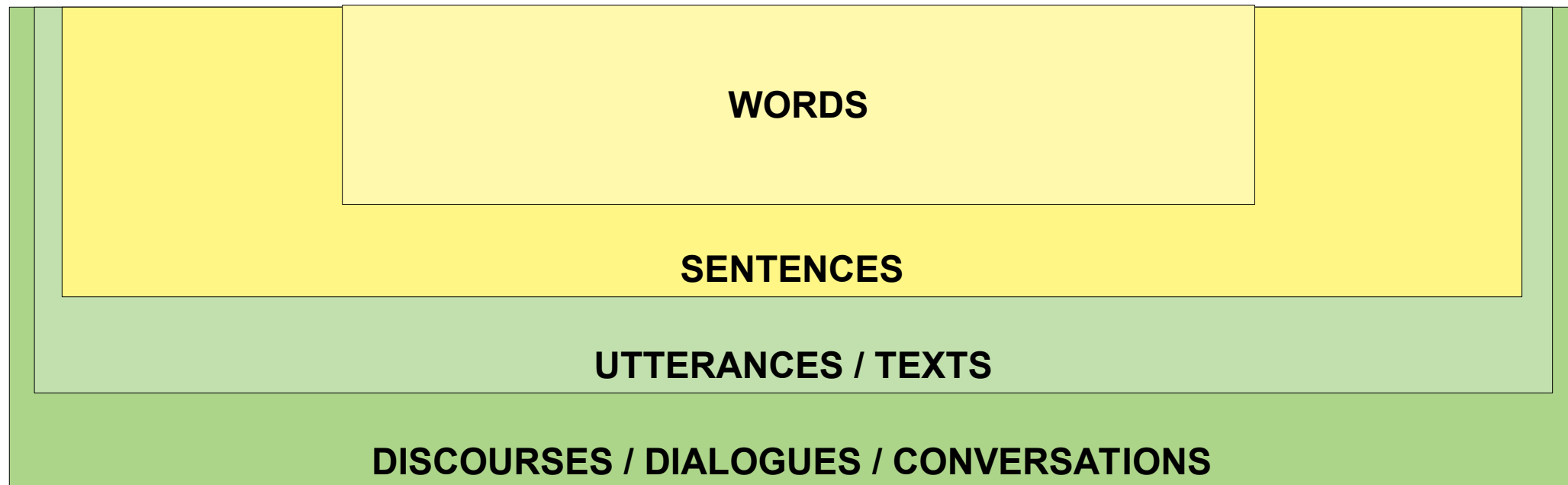 - Symbol: signifies something which is independent of its appearance, place, time or causal influence
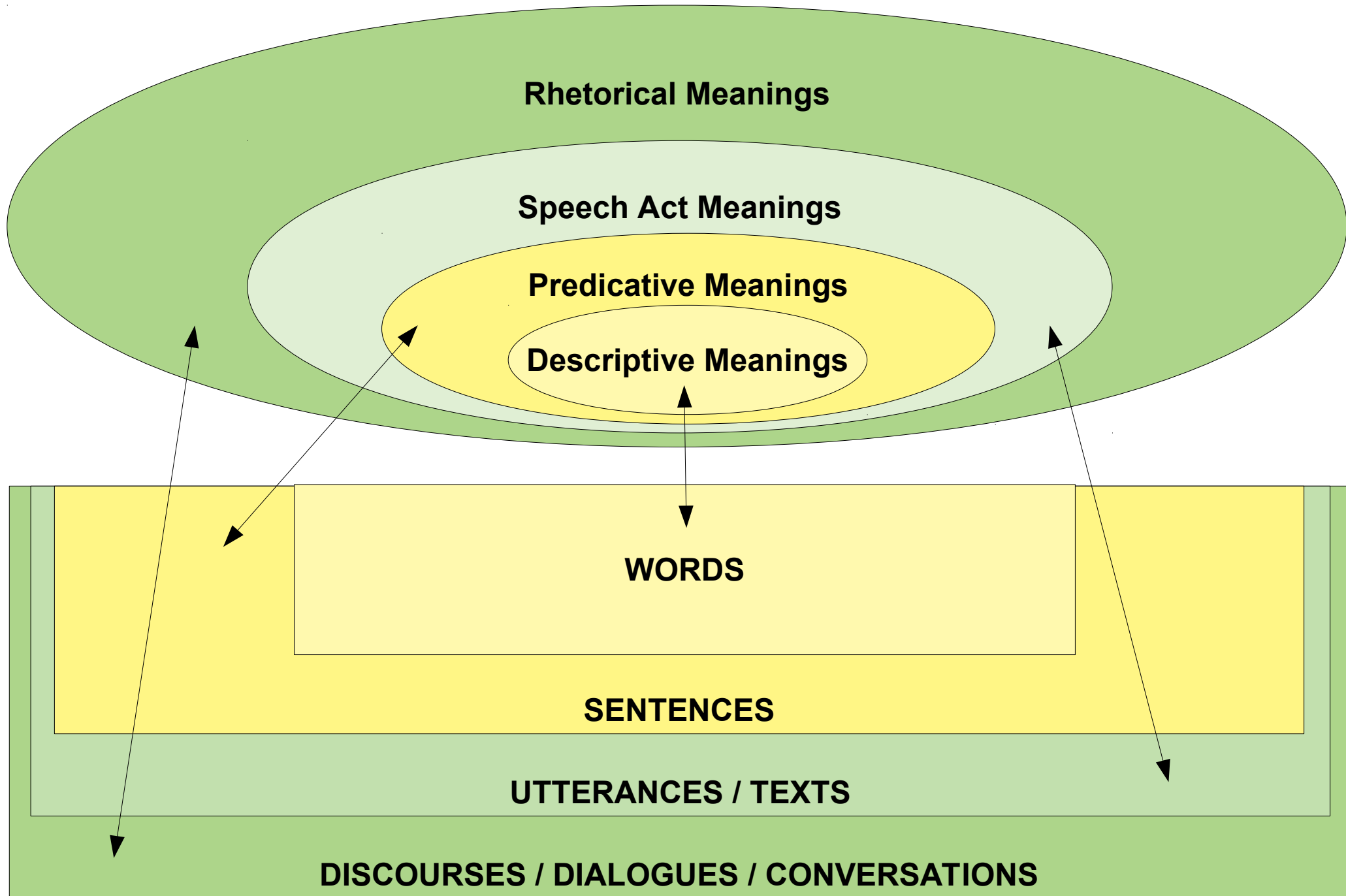
Traffic signs:
Index
Icon or
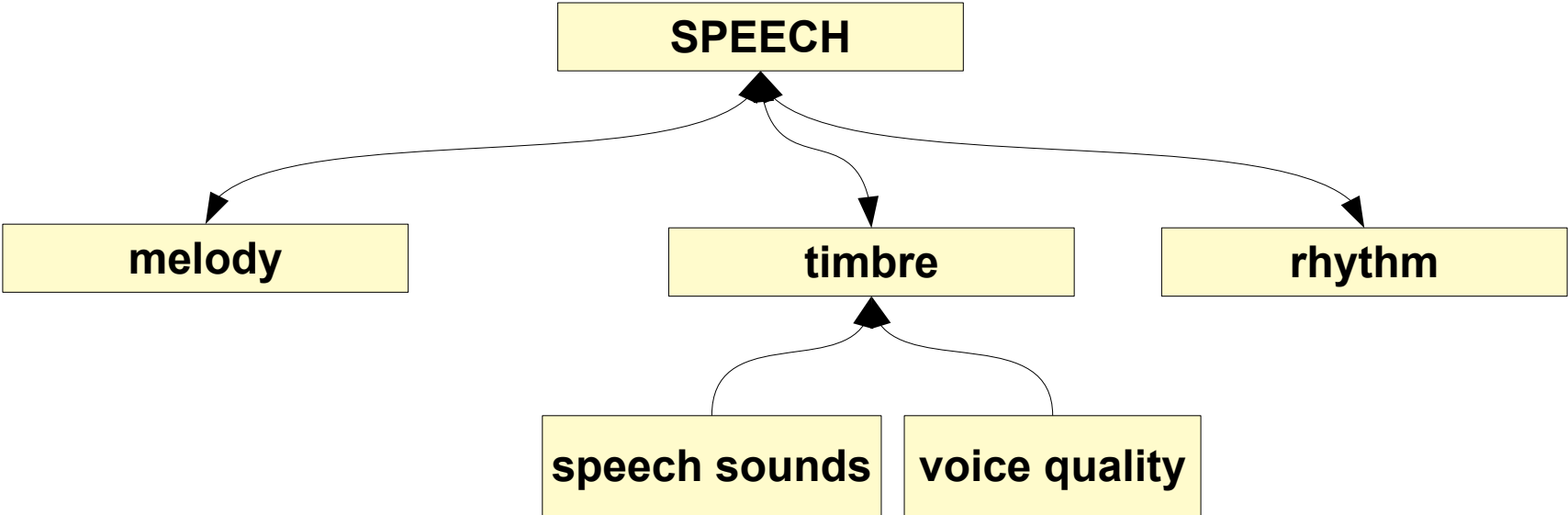Symbol?

We have meanings to communicate / understand

Rhetorical Meanings

Speech Act Meanings

Predicative Meanings

Descriptive Meanings

We have symbols for communication / understanding

WORDS

SENTENCES

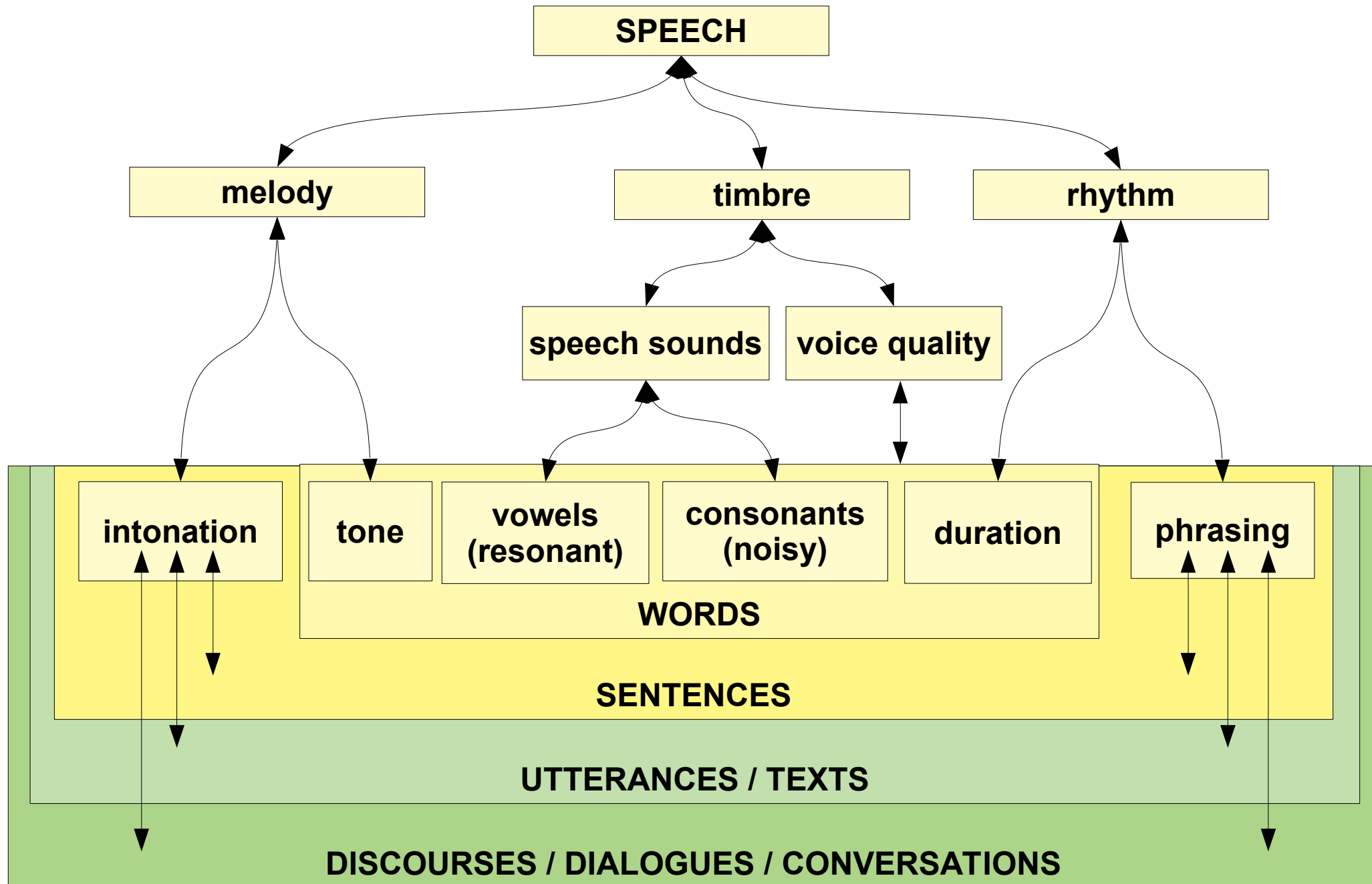UTTERANCES / TEXTS

DISCOURSES / DIALOGUES / CONVERSATIONS
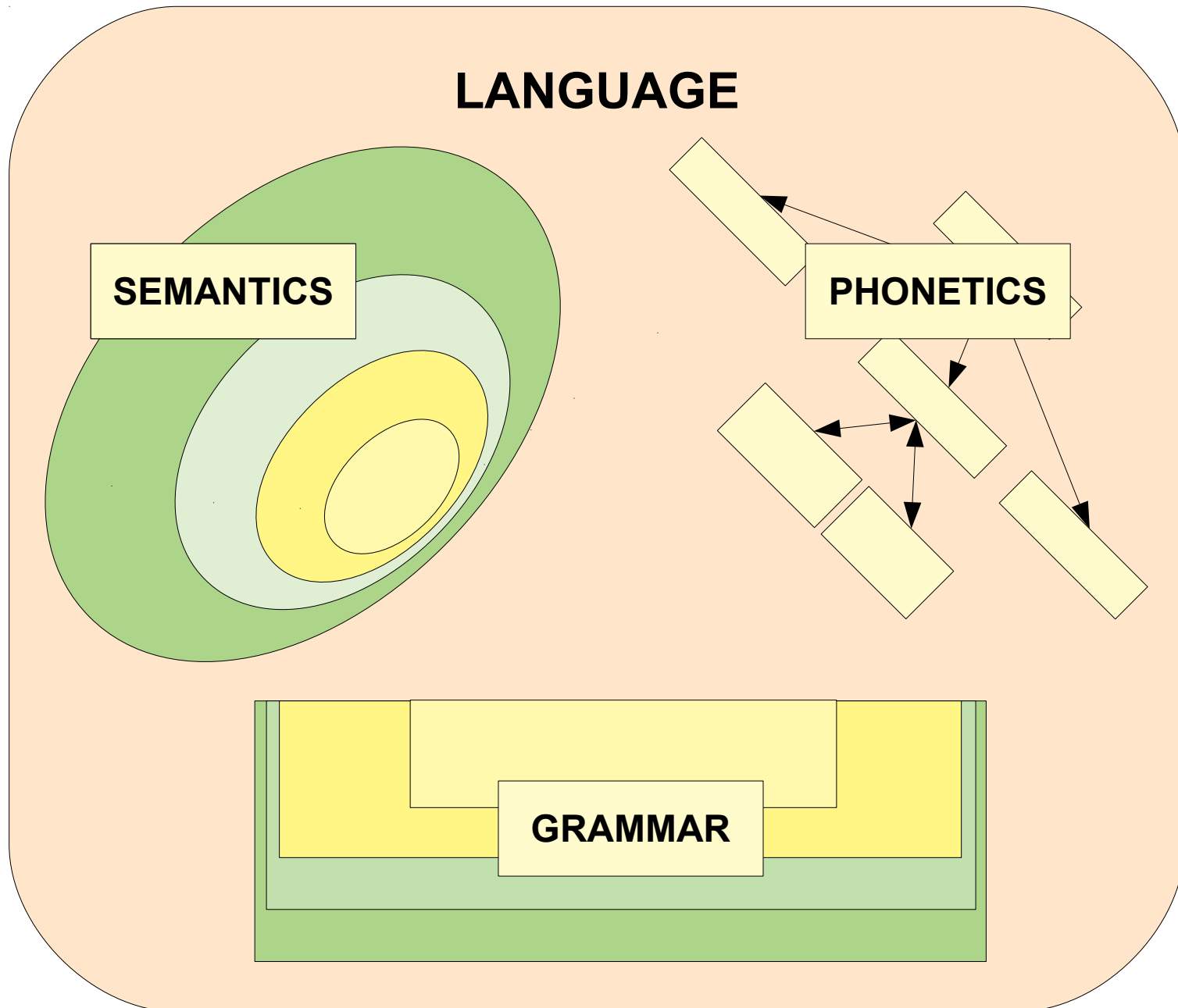
We have symbols for communication / understanding

We have sounds to encode / decode symbols

We have sounds to encode / decode symbols

We have symbols for communication / understanding

LANGUAGE

SEMANTICS

PHONETICS

GRAMMAR

We have channels and media to communicate through

OTHERS

CHANNEL

MEDIUM

YOU

ME

- - - THE WORLD - - -
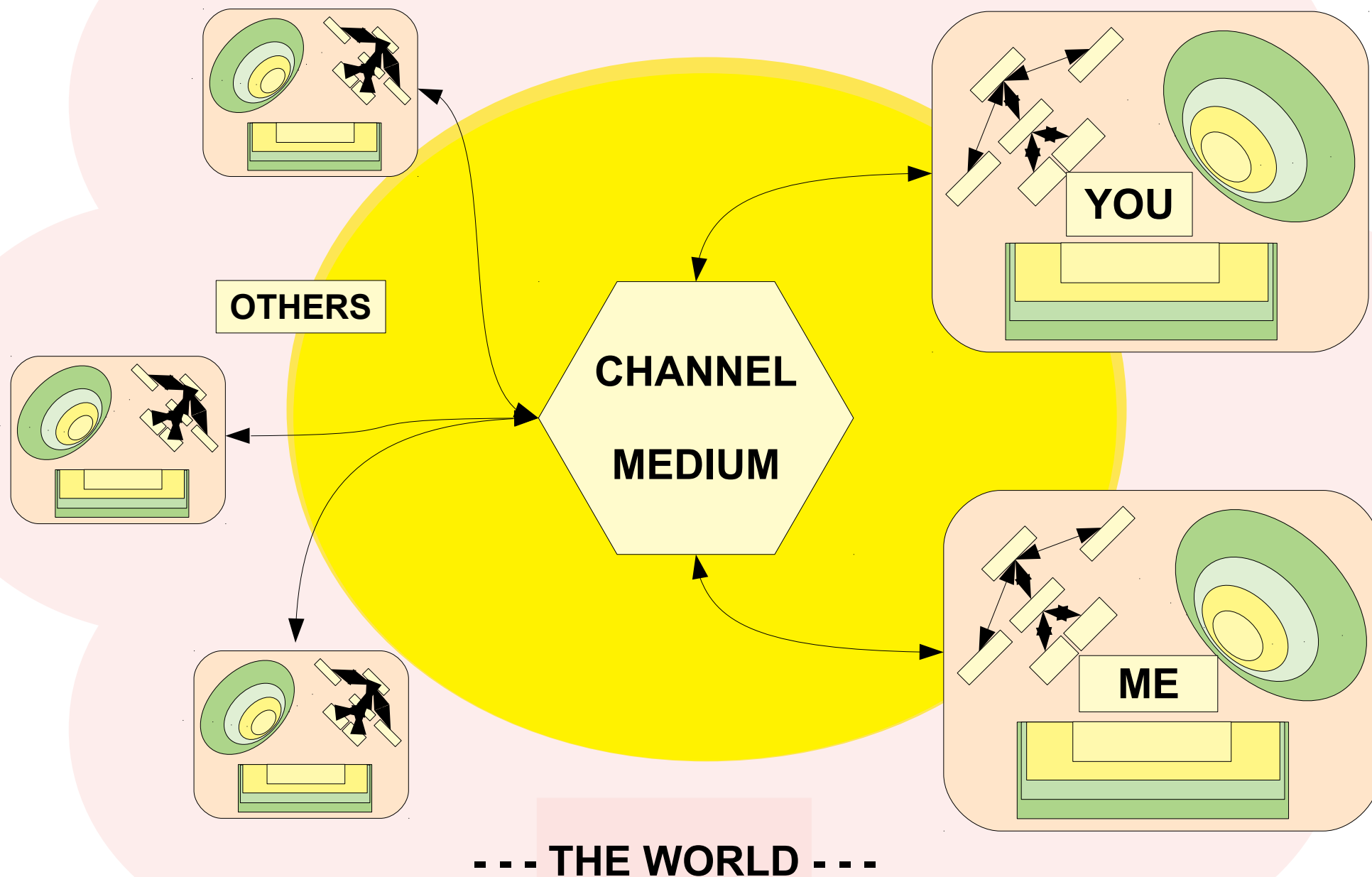
PHONETICS

D. Gibbon, Acoustic Phonetics

The Phonetic Cycle

- ## The Articulatory Domain (Articulatory Phonetics)
    - The IPA (A = Alphabet / Association)
    - The Source-Filter Model of Speech Production

- ## The Acoustic Domain (Acoustic Phonetics)
    - The Time Domain: the Speech Wave-Form
    - The Frequency Domain: simple & complex signals
        - Fourier Analysis: the Spectrum
        - Pitch extraction
    - Analog-to-Digital (A/D) Conversion

- ## The Auditory Domain (Auditory Phonetics)
    - Anatomy and Physiology of the Ear

The Methods of Phonetics

- ## Empirical Methods
    - Direct observation ("impressionistic"), usually based on articulatory phonetic criteria
    - Measurement
        - position and movement of articulatory organs
        - the structure of speech signals
        - mechanisms of the ear and perception in hearing

- ## Formal Methods
    - Statistical evaluation of direct observation and measurements
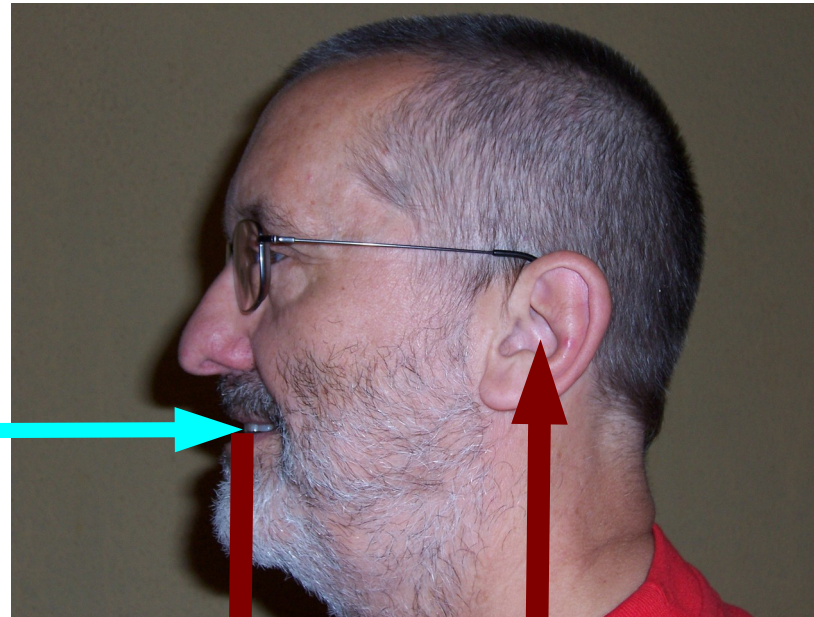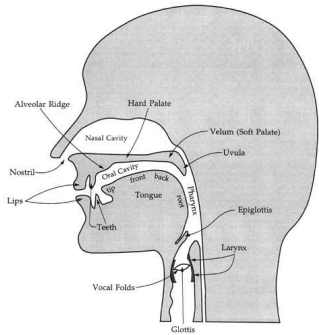   - Creation of formal models of production, transmission and perception

The Domains of Phonetics: the Phonetic Cycle

A tiger and a mouse were walking in a field

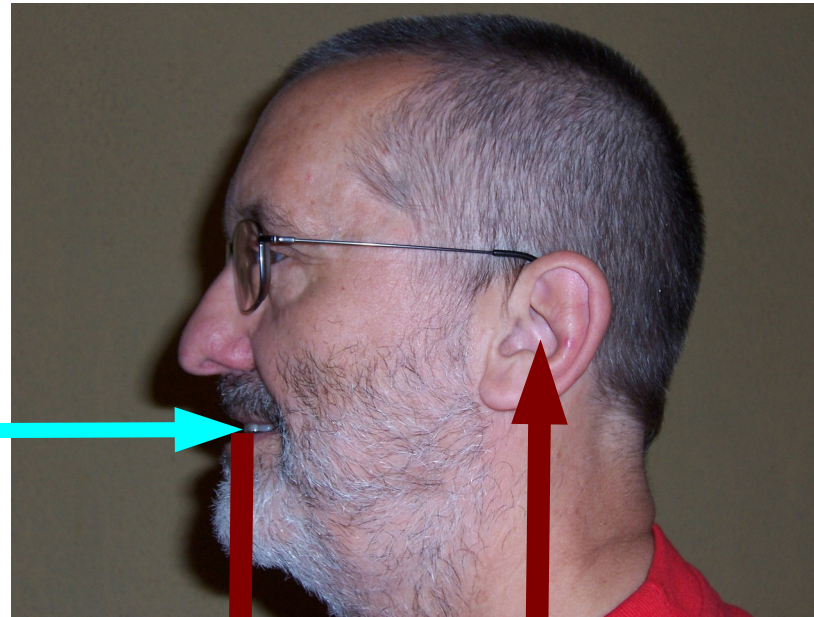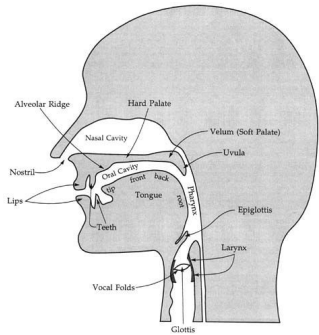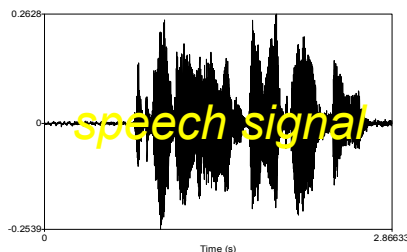# The Domains of Phonetics: the Phonetic Cycle

Articulatory Phonetics
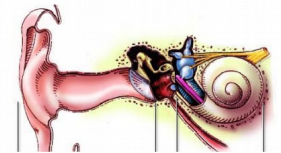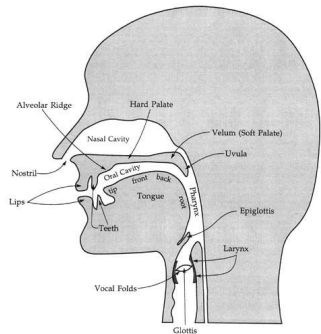
A tiger and a mouse were walking in a field...

The Domains of Phonetics: the Phonetic Cycle

Articulatory Phonetics

A tiger and a mouse were walking in a field...

speech signal

Acoustic Phonetics

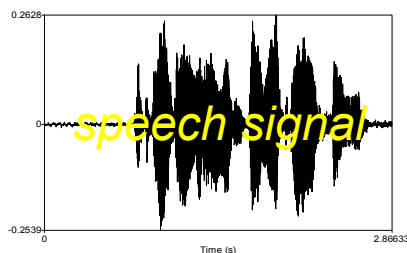D. Gibbon, Acoustic Phonetics

The Domains of Phonetics: the Phonetic Cycle

Articulatory Phonetics

Auditory Phonetics

Acoustic Phonetics

A tiger and a mouse were walking in a field...
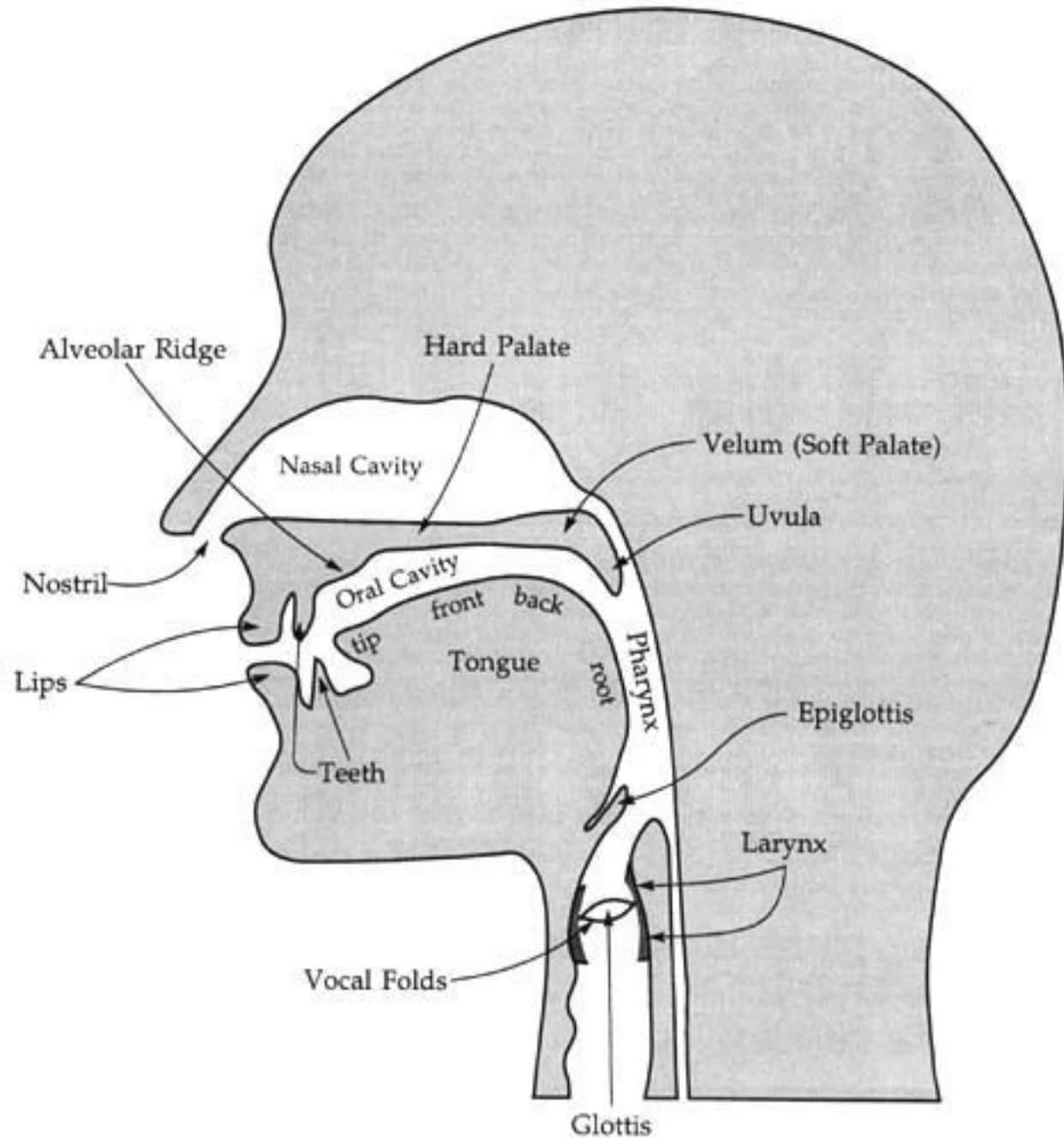
speech signal

Summary – Phonetic Domains

- Define each of the following:
    - articulatory phonetics?
    - acoustic phonetics?
    - auditory phonetics?
- Which parts of the head are they associated with?
- What is the "phonetic cycle"?

Articulatory Phonetics
(Speech Production)

The articulatory domain

- Domain of speech production
- Articulatory organs are relatively easily observable
- Domain of reference for phonetic categories of the IPA
- Investigated via
  - corpus creation
  - experiment paradigm

The IPA (A = Alphabet / Association)

- **IPA**: 1897
  - Latest revision: 2005
  - <u>articulatory</u> categories
- **transcription** of
  - the **phonemes**
  - of **all languages of the world**
- **phoneme:**
  - a vowel or consonant which distinguishes words in at least one language

THE INTERNATIONAL PHONETIC ALPHABET (revised to 1993)

The IPA (A = Alphabet / Association)

- **transcription** of the **phonemes** of **all languages of the world**

- **phoneme:** a vowel or consonant which distinguishes words in at least one language
  - /kɪt/ /kæt/ /kɒt/ /kʌt/ /kuːt/
  - /kɜːt/ /kaːt/ /kɔːt/ /kəʊt/ /kɪːt/

- **vowels** are the main factors in **rhythm** and **melody**

The <u>Source-Filter Model</u> of Speech Production

- A "model" is a simplified representation of relevant features of reality (it also adds its own artefacts)
- In the **Source-Filter Model** of speech production, the sound is generated by the **SOURCE** and modified by the **FILTER**
- Two types of **source**:
    - Larynx: melody (tone, intonation)
    - Narrowing / closing of the mouth (noisy consonants)
- Three types of **filter**:
    - the PHARYNGEAL CAVITY (throat)
    - the ORAL CAVITY (mouth)
    - the NASAL CAVITY (nose)

The Source-Filter Model of Speech Production

The Source-Filter Model of Speech Production

D. Gibbon, Acoustic Phonetics
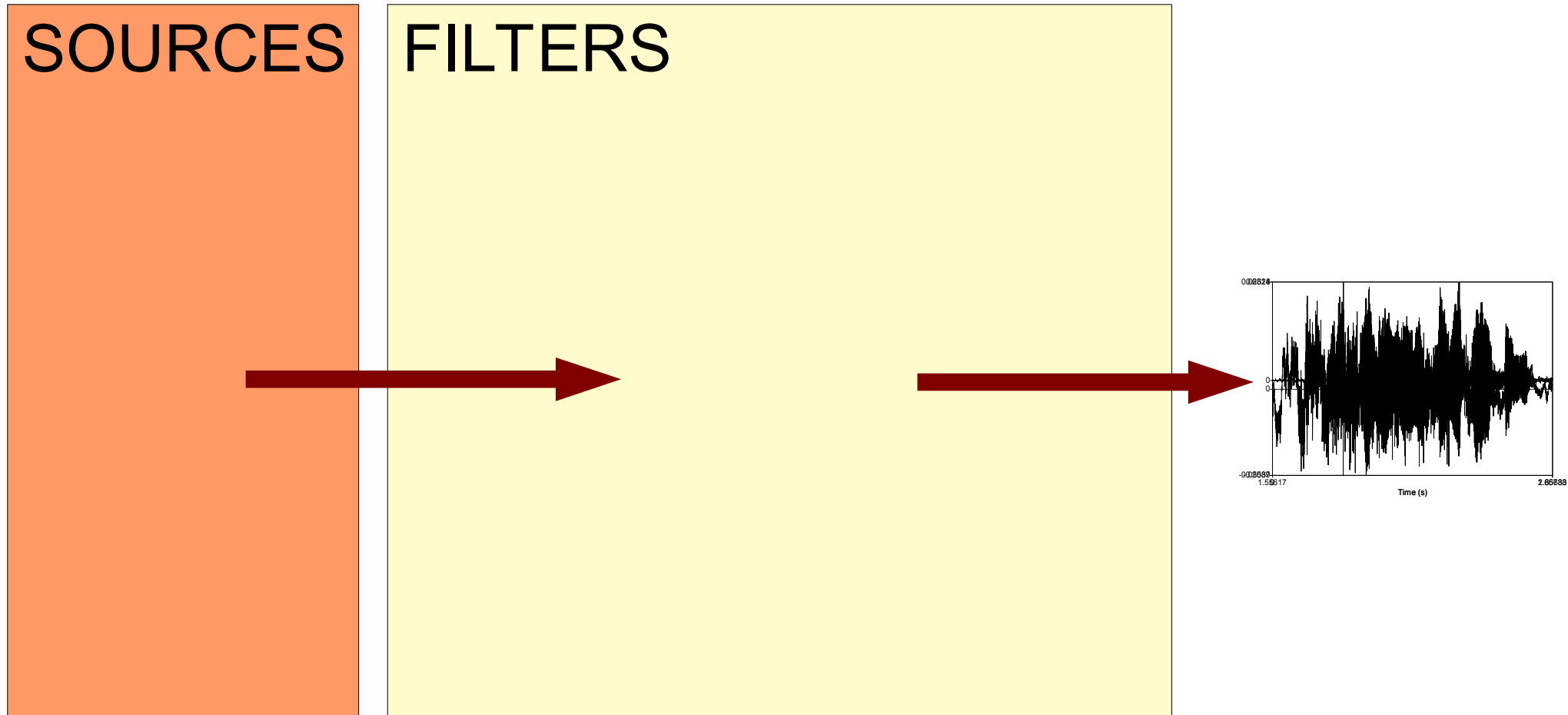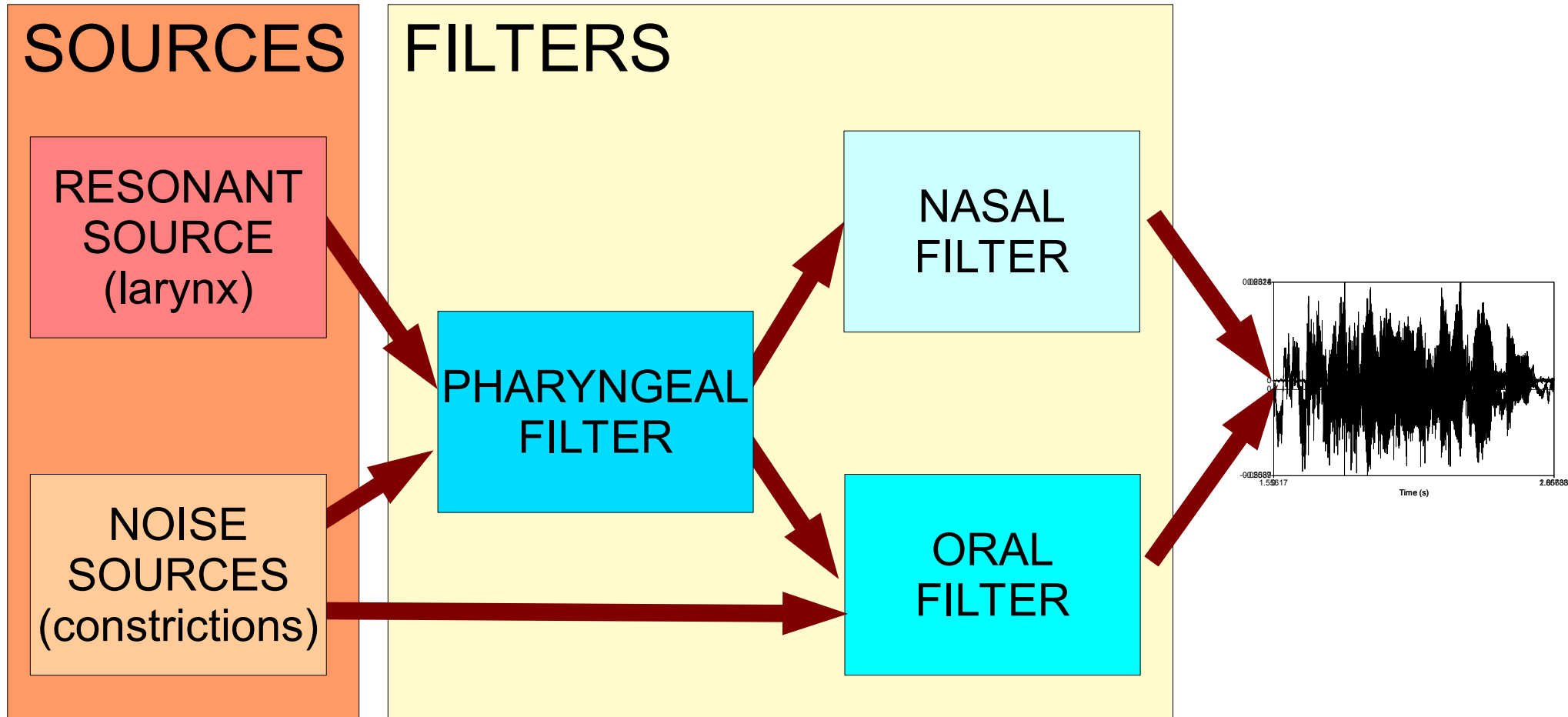
Summary: Articulatory Phonetics

- Which are the main articulators involved in
  - vowel production?
  - consonant production?
  - tone production?
- Produce these consonants, followed by the vowel [a]:
  - voiceless bilabial fricative
  - voiced palatal stop
  - voiceless labial-velar stop
  - implosive velar stop
  - velar nasal
- What is the source-filter model?
  - Illustrate this, referring to the difference in sound between speaking in a tiled bathroom and in the open air.

Acoustic Phonetics
(Speech Transmission)

The acoustic domain

- **Acoustic phonetics** is concerned with investigating the **transmission** of speech signals through
    - gases such as air, other substances (e.g. bone, tissue)
    - electronic amplification and storage
- The basic **parameters** of the speech signals are
    - **Amplitude** → **energy, intensity, loudness**
    - **Frequency** → **melody, pitch: tone, intonation**
    - **Time** → **duration, rhythm**
- The **methods** used to analyse speech signals are:
    - observation, transcription, description by a trained phonetician
    - analog-to-digital (A/D) conversion
    - mathematical definitions of filters and transformations

The Speech Wave-Form

- Speech is transmitted through air (and other substances) as a regular wave of pressure changes:
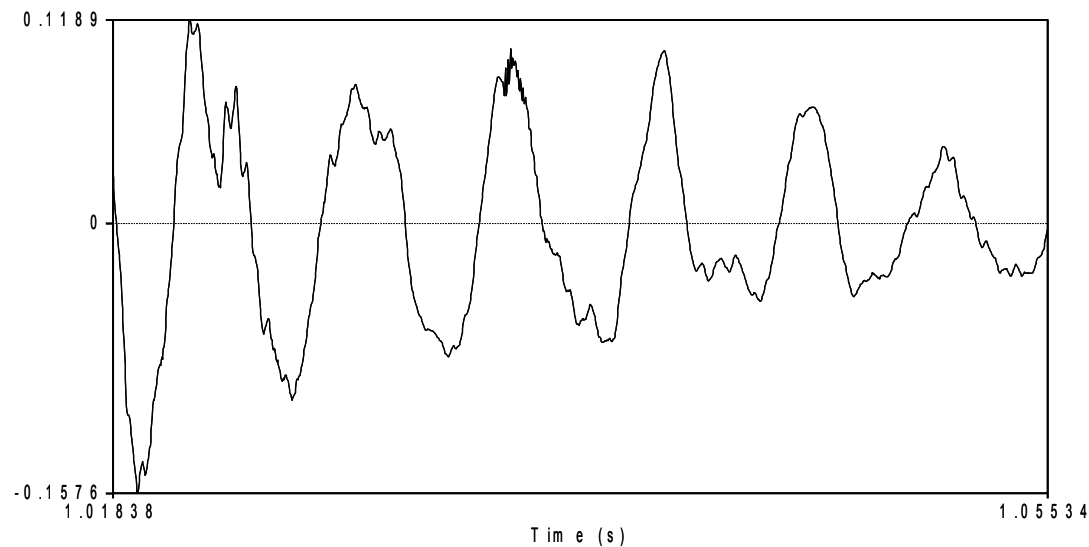
- The changes in air pressure
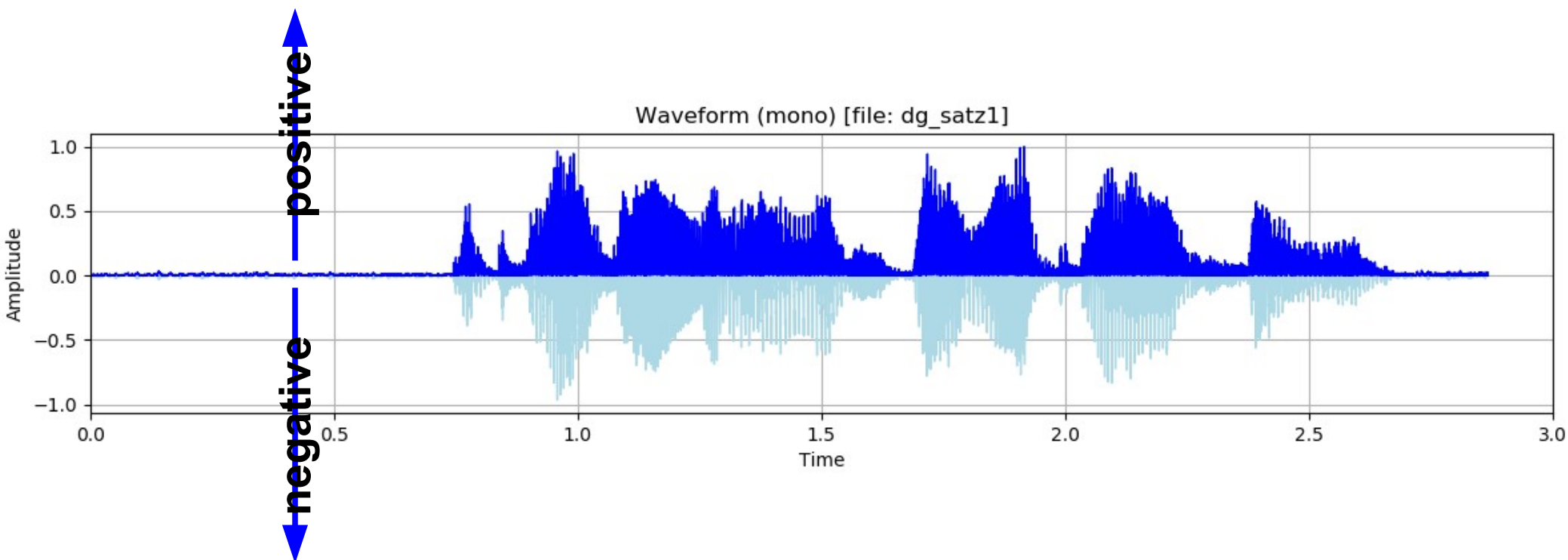 - can be <u>heard</u>
  - can be <u>measured</u> (like the waves on the ocean)
  - the measurements can be <u>visualised</u> and used for calculating statistical <u>models</u> of the structure of speech

Basic Speech Signal Parameters

Waveform (mono) [file: dg_satz1]
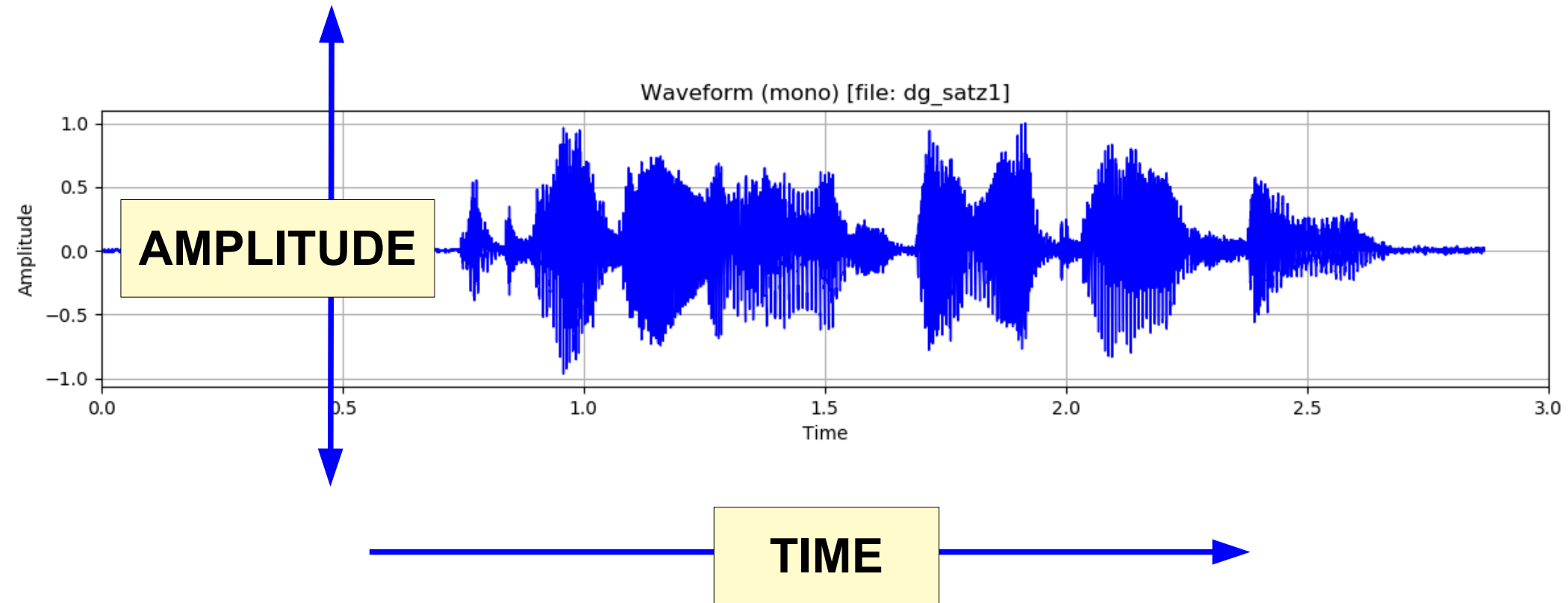
AMPLITUDE

TIME
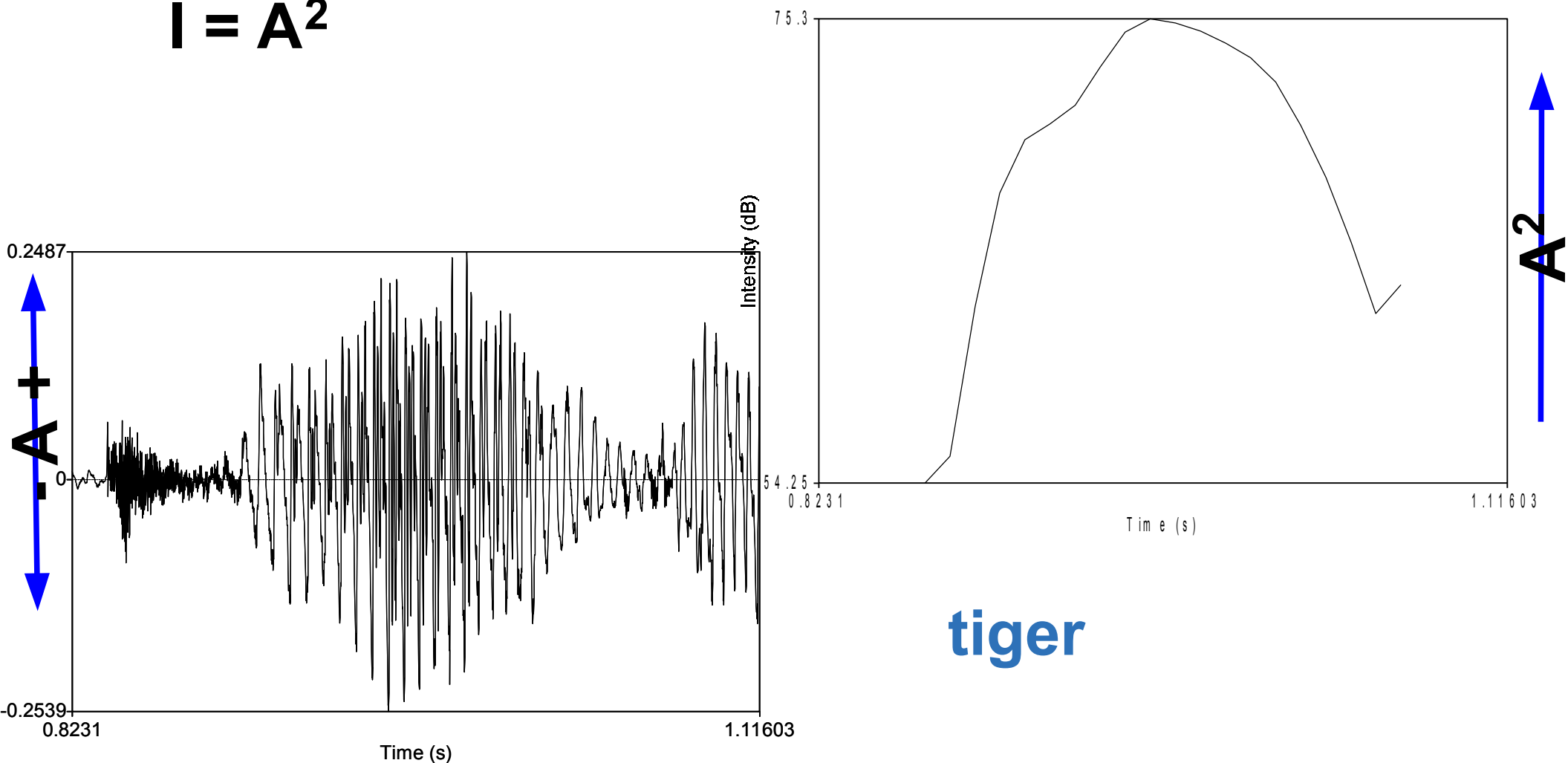
The Time Domain: the Speech Wave-Form

Positive or negative amplitude A of the speech signal:

Derived parameter *INTENSITY*

- The *intensity* of the speech signal at any given point in time is the *square of the amplitude* of the wave from zero at this point in time:
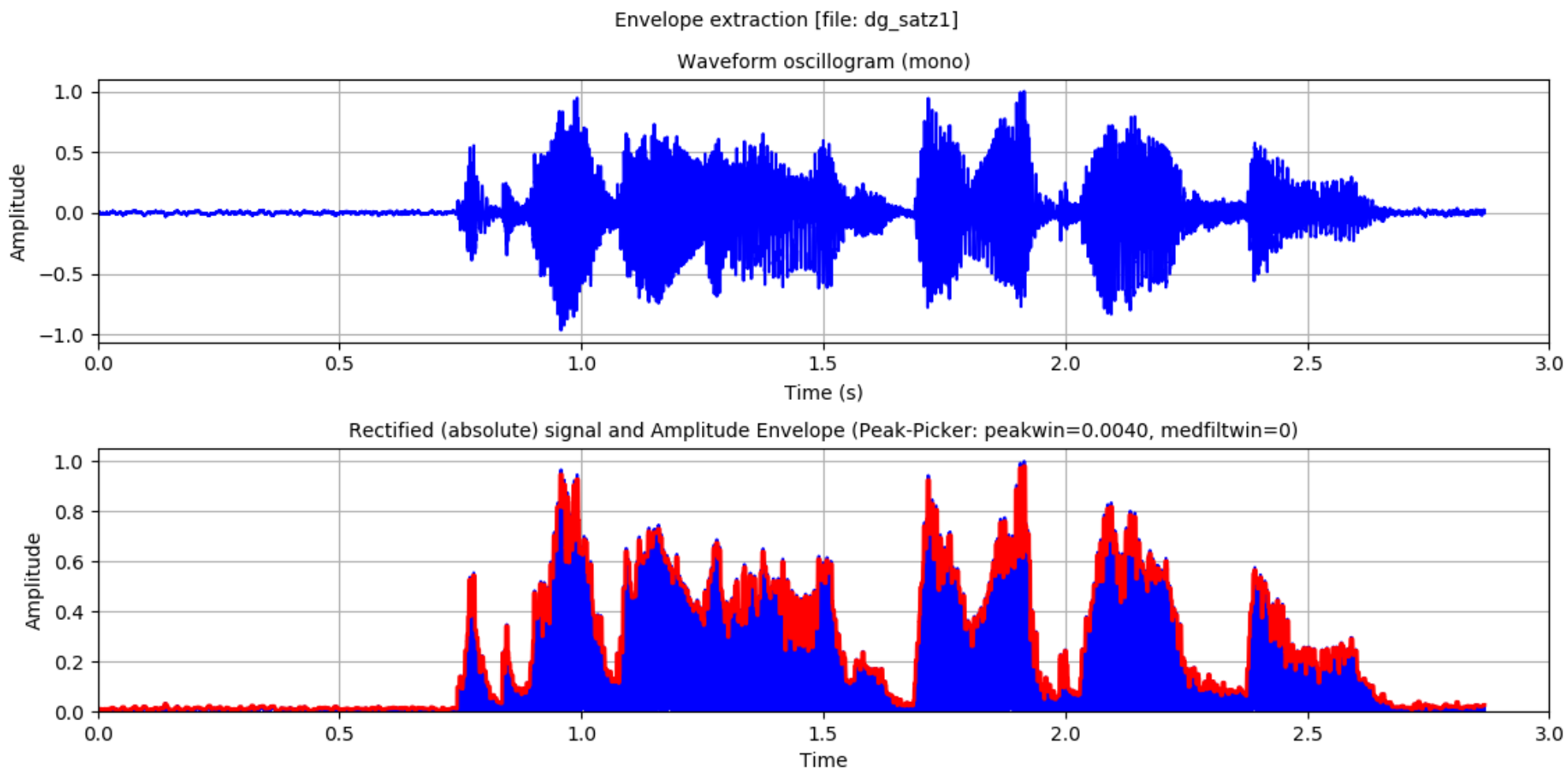
$$I = A^2$$

A +

0.2487

0

-0.2539

0.8231

1.11603

Time (s)

Intensity (dB)

75.3

54.25

0.8231

1.11603

Time (s)

A²

tiger

Derived parameter *ENERGY*

- The energy *E* (root-mean-square energy) is
  - the square root of the mean of a sequence of intensity values $I_1, ..., I_n$ (remember: intensity is amplitude squared)
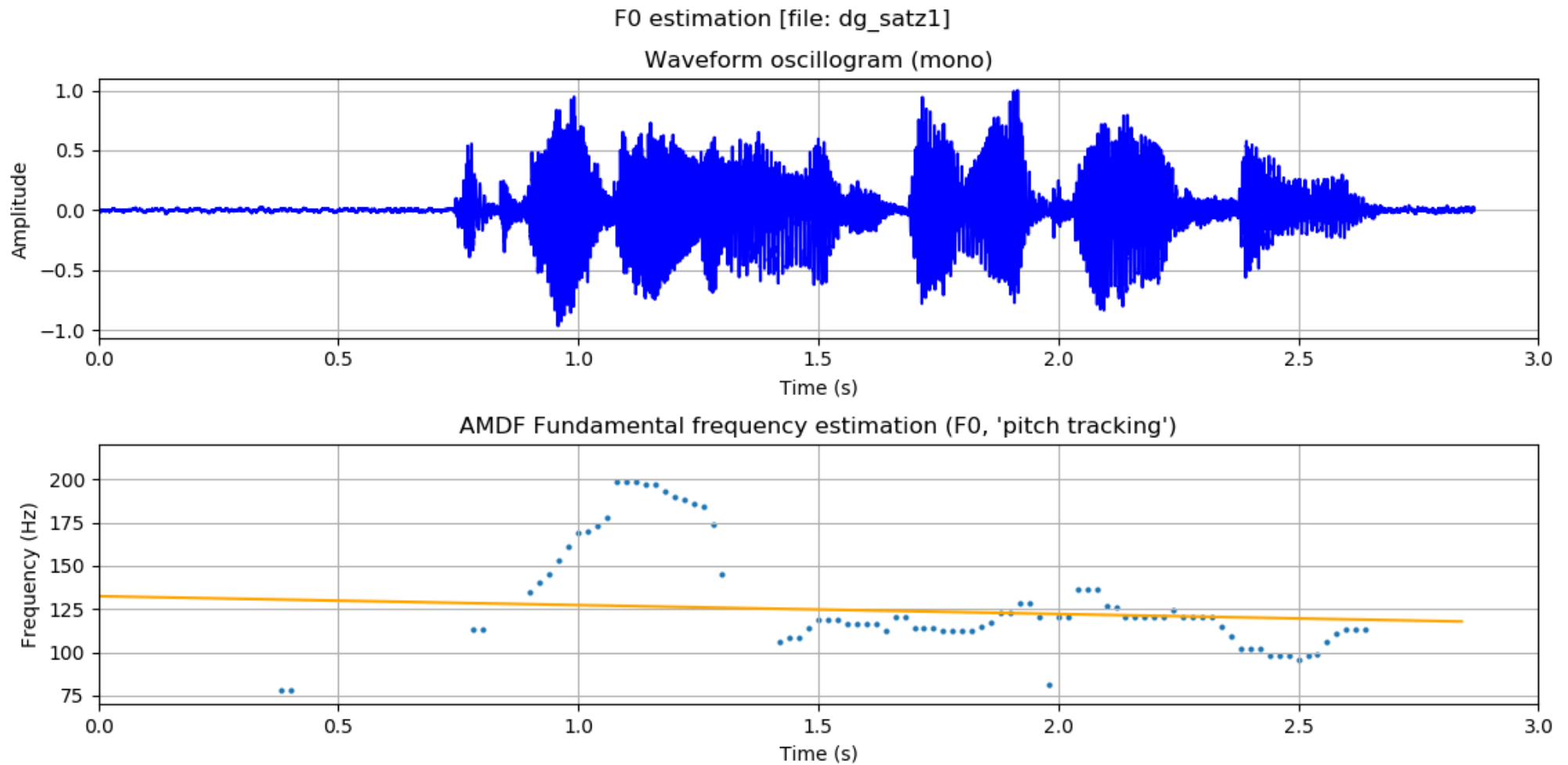
$$E = \sqrt{\frac{\Sigma_{i=1...n} A(x_i)^2}{n}}$$

- Energy is intensity averaged over time
  - In fact, intensity measurements are, in practice, energy measurements over very short periods of time
- Compare other measurement units per time unit:
  - miles per hour
  - kilowatts per hour

The Time Domain: the Speech Wave-Form

Envelope extraction [file: dg_satz1]

Waveform oscillogram (mono)

Rectified (absolute) signal and Amplitude Envelope (Peak-Picker: peakwin=0.0040, medfiltwin=0)

The Time Domain: the Fundamental Frequency

F0 estimation [file: dg_satz1]
Waveform oscillogram (mono)
AMDF Fundamental frequency estimation (F0, 'pitch tracking')

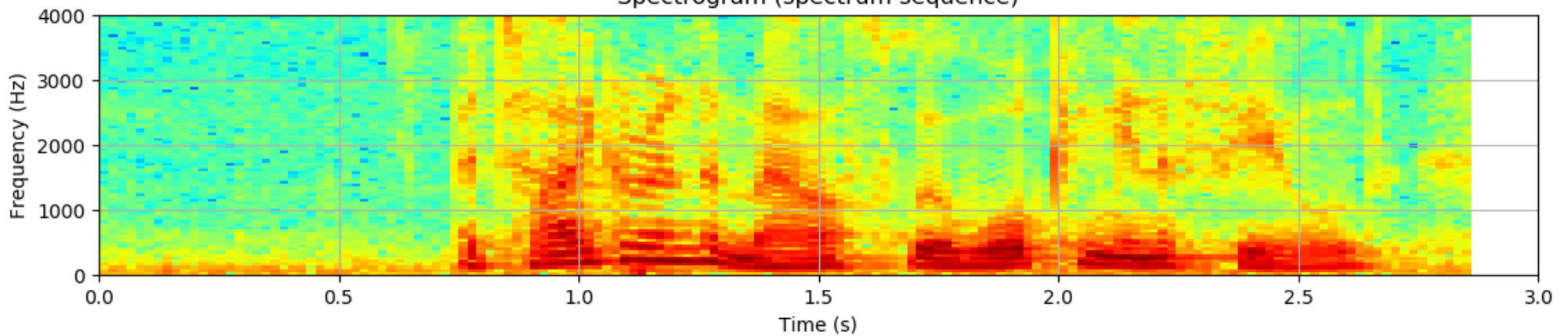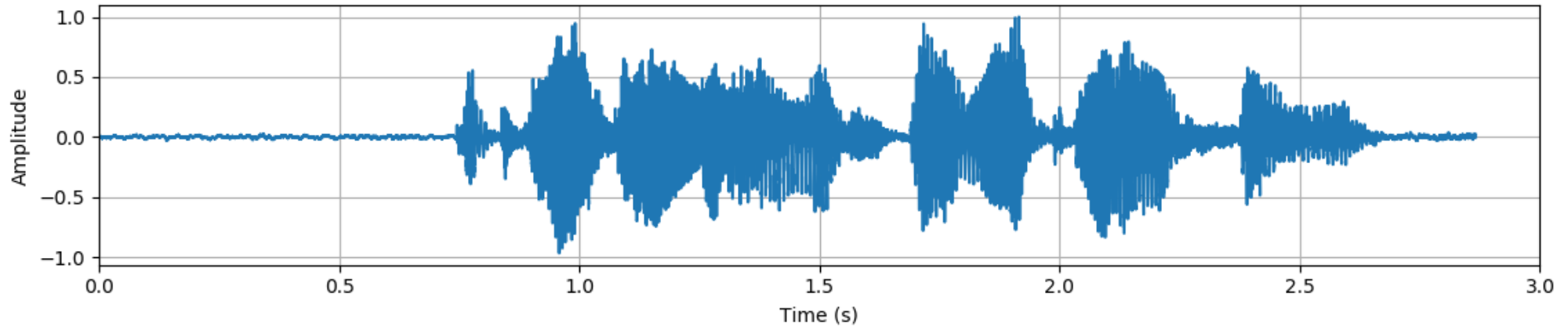# Time and Frequency Domains: the Spectrogram
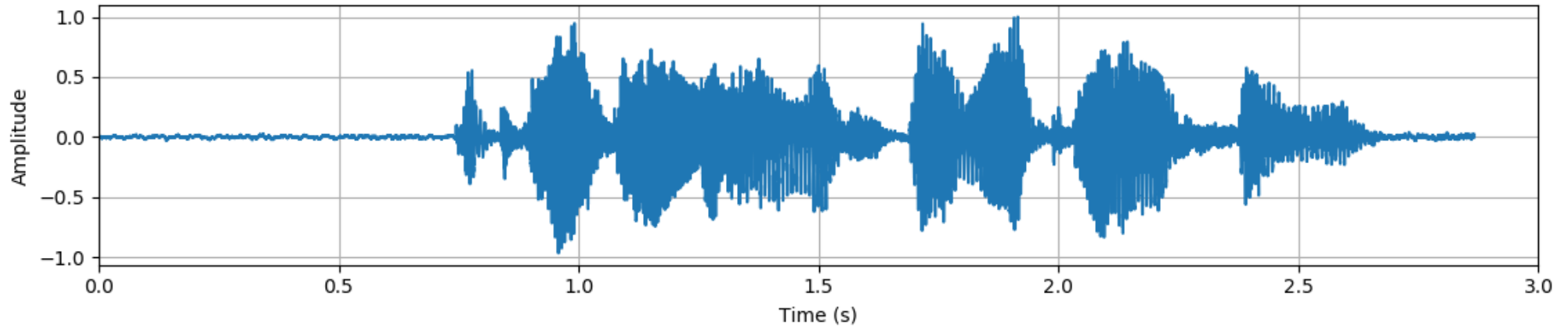
Spectrogram [file: dg_satz1]

Time and Frequency Domains: the Spectrogram

Spectrogram [file: dg_satz1]

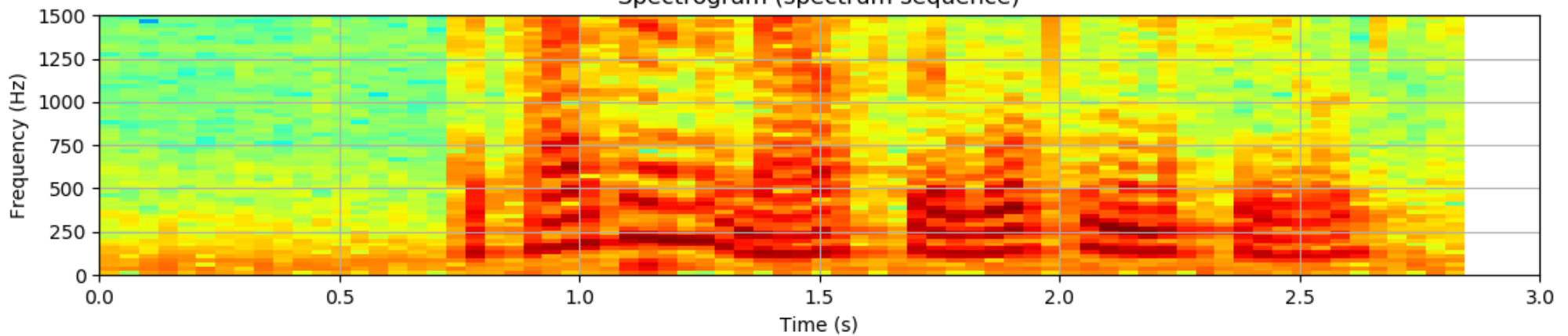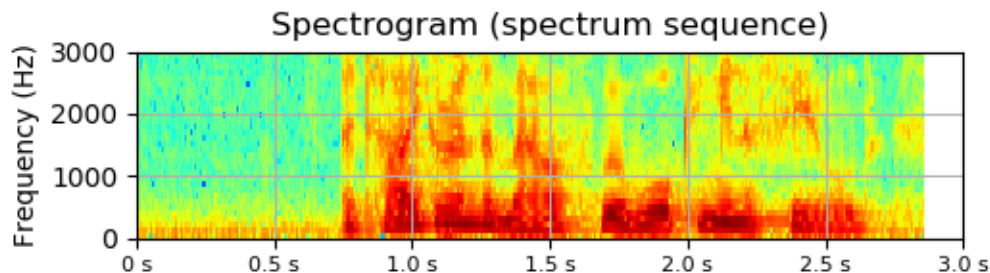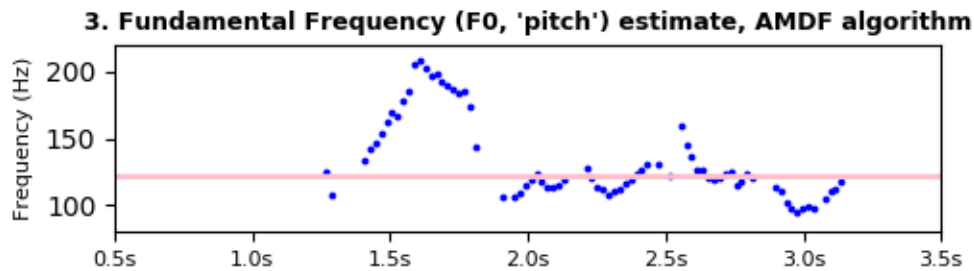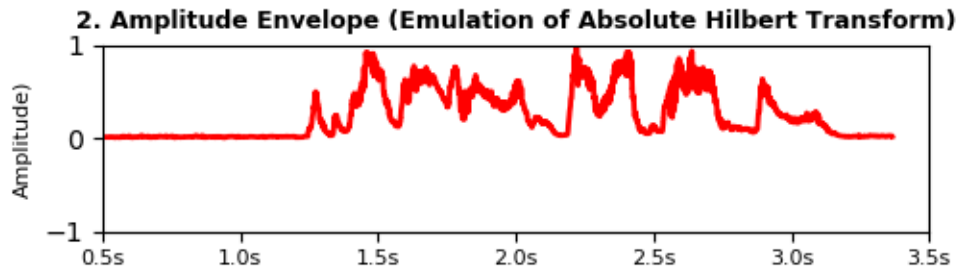Waveform oscillogram (mono)

Spectrogram (spectrum sequence)

Four Acoustic Phonetic Visualisations

1. Waveform (mono)

2. Amplitude Envelope (Emulation of Absolute Hilbert Transform)

3. Fundamental Frequency (F0, 'pitch') estimate, AMDF algorithm

Spectrogram (spectrum sequence)
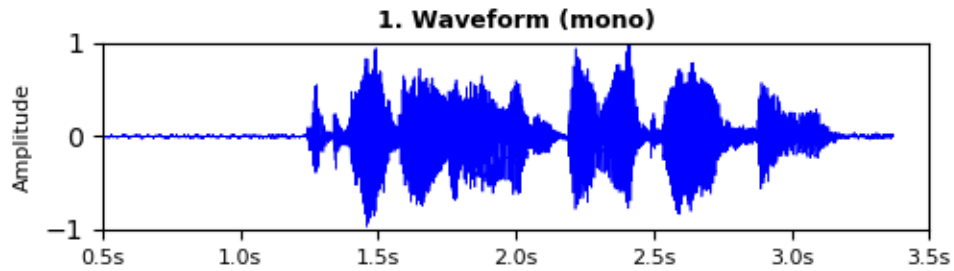
… almost all you need to know about the rhythms and melodies of speech …

Summary: Acoustic Phonetics

- What are the basic parameters of the speech signal?
- Define the following terms:
  - amplitude
  - intensity
  - energy
- How are time-domain representations of speech signal converted to frequency domain representations?
- Define the following terms:
  - Spectrum, spectrogram
  - fundamental frequency, F0, pitch
  - harmonic
  - formant
  - analog-to-digital conversion

Auditory Phonetics
(Speech Perception)

The Auditory Domain: Anatomy of the Ear
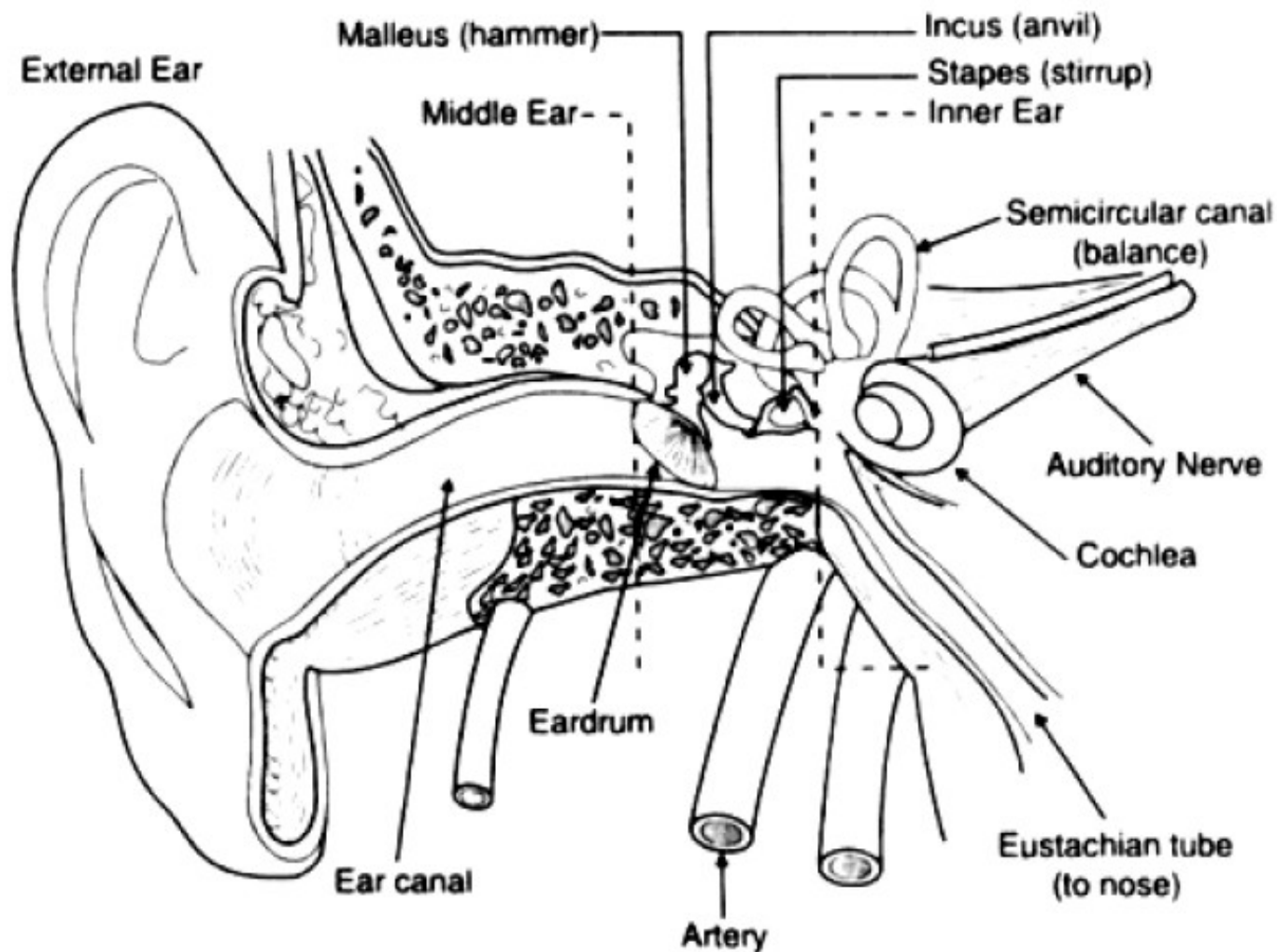
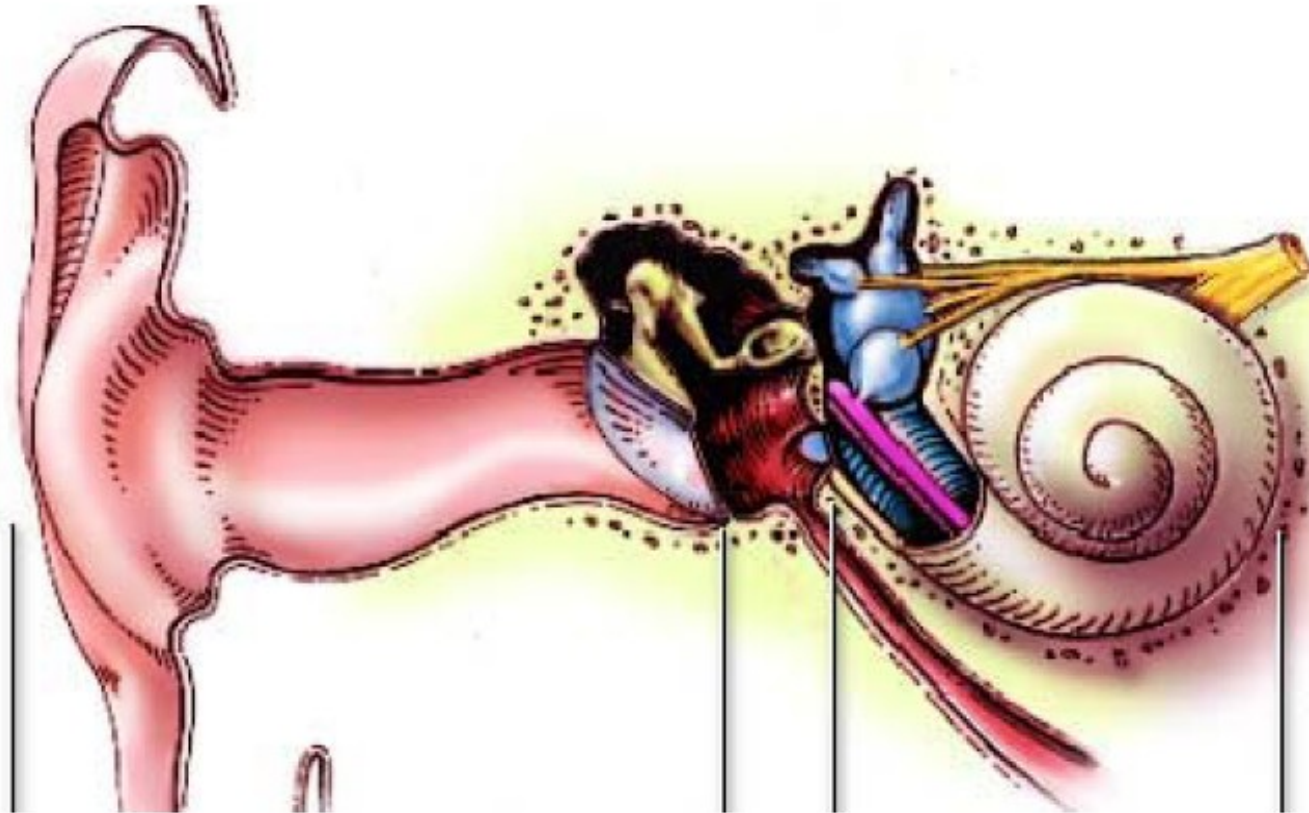# The Auditory Domain: Anatomy of the Ear
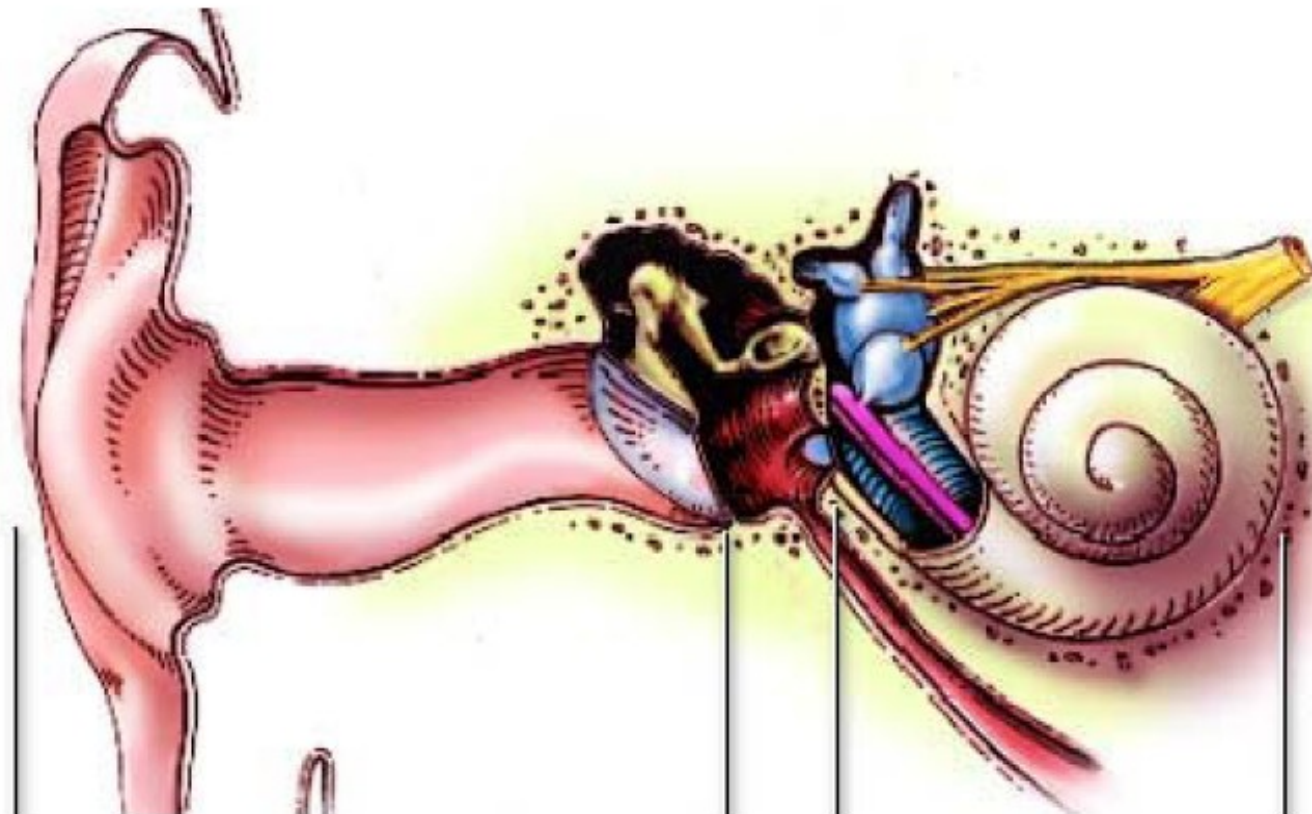
outer ear

middle ear

inner ear

The Auditory Domain: Anatomy of the Ear

| microphone | amplifier | Fourier transform |

outer ear

middle ear

inner ear

Summary: Auditory Phonetics

- What are the functions of
  - the outer ear?
  - the middle ear?
  - the inner ear?
- What are
  - the ossicles?
  - the oval window?
  - the cochlea?
  - the basilar membrane?

Conclusion

After studying this unit ...

… you should know the basic semiotic and physical foundations on which rhythms and melodies of speech are based