Speech Science

WiSe 2024

Spoken Language

Oct 24, 2024

Bernd Möbius & Valentin Kany

Language Science and Technology
Saarland University

UNIVERSITÄT DES SAARLANDES

Vowels and consonants: written

- __e _ea__e_ _o_e_a__ _o_ _o_o__o_:
  _a__e_ __ou_y i_ __e _o__i__ _i__
  a _e_ _u__y __e___ i_ __e a__e__oo_.

Vowels and consonants: written

- Th_ w__th_r f_r_c_st f_r t_m_rr_w: r_th_r cl__d_ _n th_ m_rn_ng w_th _ f_w s_nn_ sp_lls _n th_ _ft_rn__n.

Vowels and consonants: written/spoken

- The weather forecast for tomorrow:
  rather cloudy in the morning with
  a few sunny spells in the afternoon.

- Consonants apparently are more informative than vowels for comprehending a **written** utterance (in languages such as DE, EN, or similar).

- Does this pertain to **spoken** language, too?

Vowels and consonants: spoken

- cf. spoken language:
  - consonants only 
  - vowels only 
  - original 

Vowels and consonants: spoken

- Vowels are apparently more informative than consonants, but we also need to have access to the temporal structure (and rhythm) of utterances.
- **Speech rhythm**: a combination of syllable structure and the weight (duration, prominence) of vowels.
- Demo:
  - vowels only, without silence
  - vowels only, with silence
  - vowels only, monotonous
  - original

Continuous speech

- We perceive continuous speech by chunks
- **Syllables** are prominent vowels, surrounded by (less prominent) consonants.
- **Sentences/utterances** consist of phrases that consist of words that consist of syllables that consist of vowels and the consonants surrounding them.
- **Prosody** (intonation, duration, intensity) contributes to making important chunks more prominent than others.

Continuous utterances

- "The president will be elected for a period of four years."
- Spoken language:
  - natural, continuous
  - chain of isolated words
  - natural, pauses between words
  - chain of isolated words, no pauses
  - isolated vs. continuous function words





- Production effort ↔ Informativity of words
  (longer+louder+unreduced = more effort and precision)

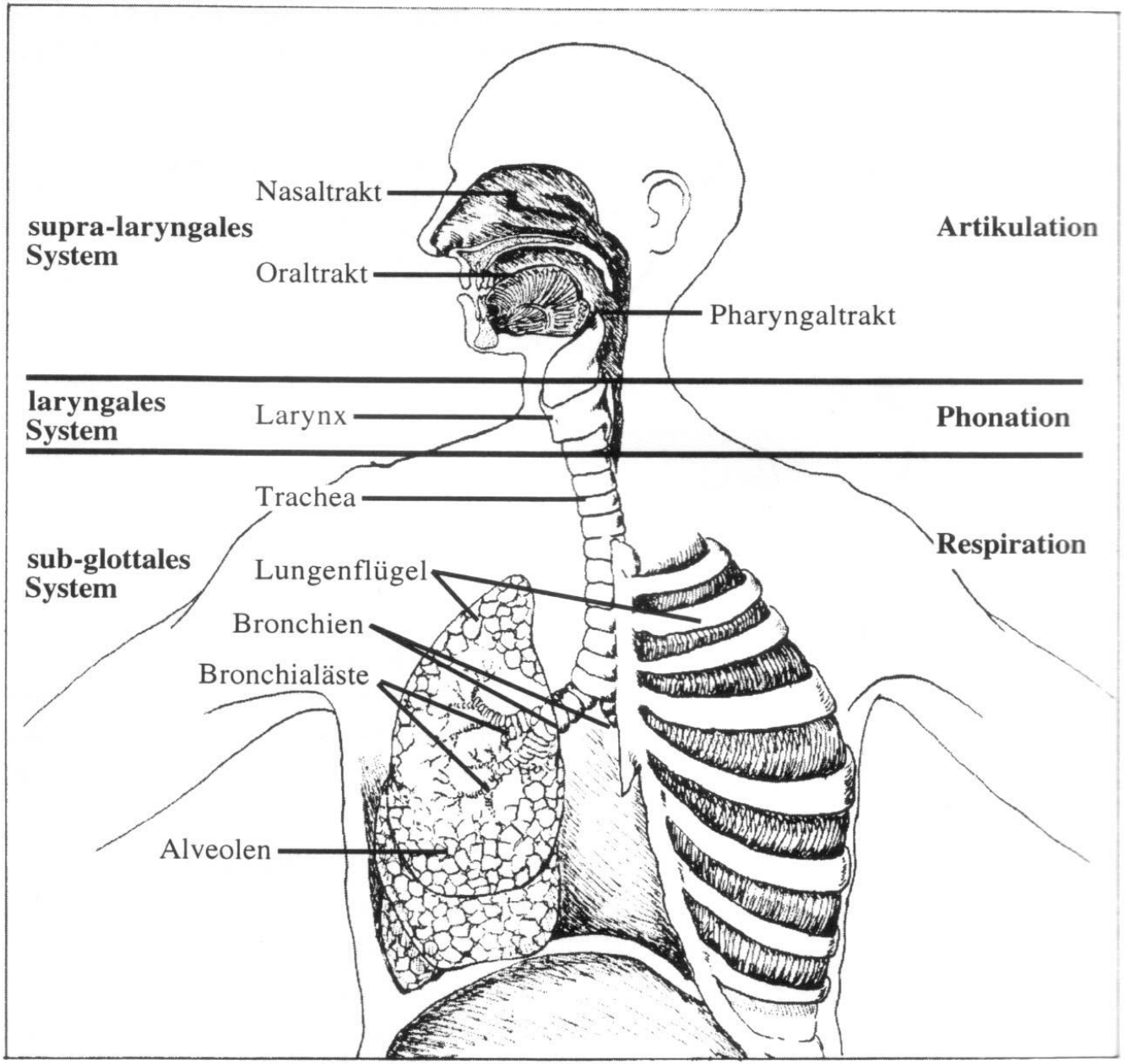
Continuous speech

- Normal, everyday communication: spoken language is not comprised of speech sounds produced in isolation but of **continuous utterances**.
- We do not identify the speech sounds that reach our ears as individual speech sounds.
  - But we can **demonstrate** individual speech sounds.
- The syllable (C*VC*) is arguably the minimal unit of speech perception...
  - and the planning unit in speech production
  - and the reference frame in speech acquisition.

Phonetics

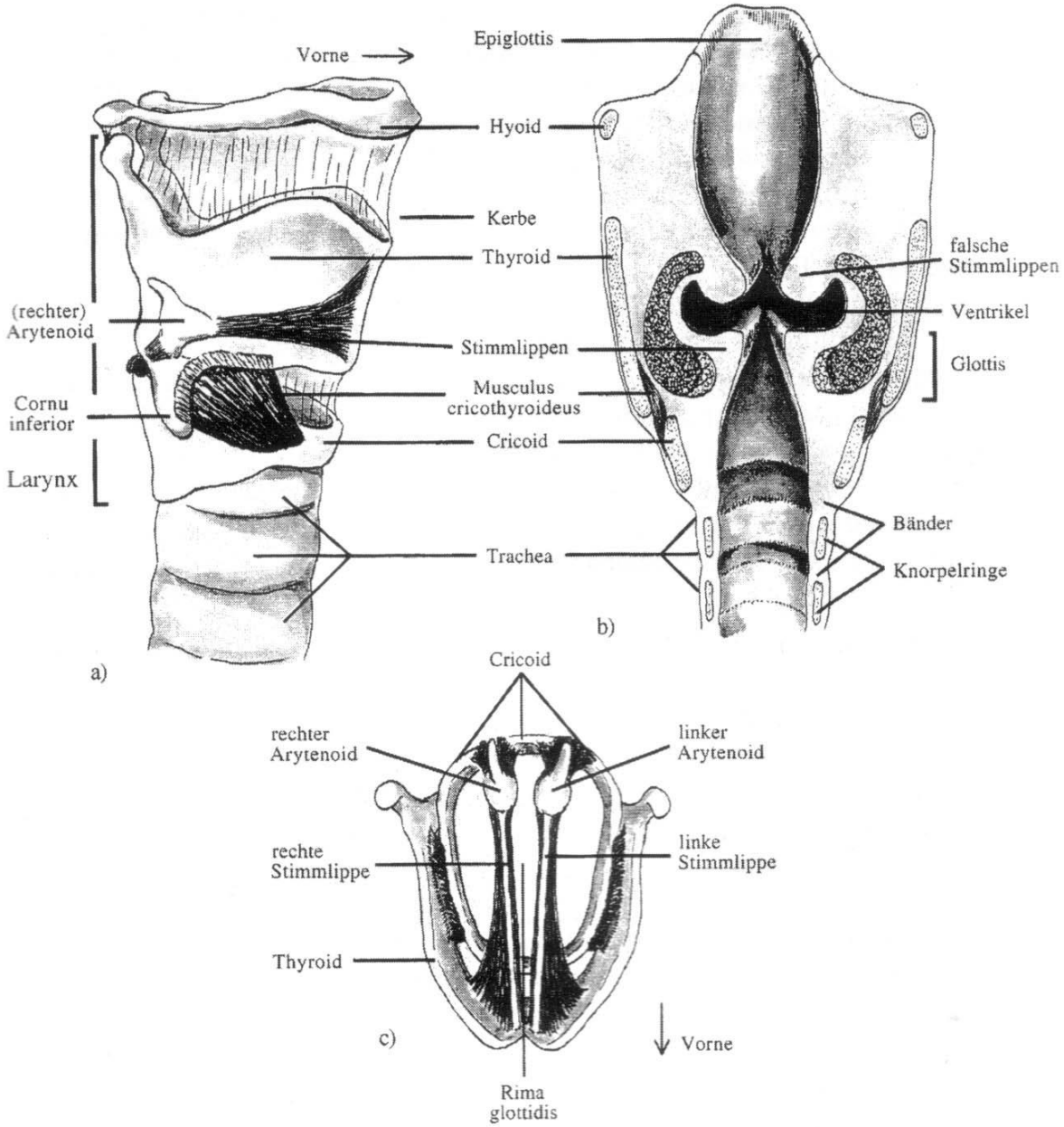
- **Scientific study of *spoken language***
- Basic conditions and constraints of human speech production and perception
- How is spoken language produced and perceived?
  - anatomy and physiology
  - speech production, phonation, articulation
  - speech acoustics, speech signals
  - speech perception
- Articulatory phonetics, Acoustic phonetics, Auditory-perceptual phonetics, Neurophonetics

Speech production organs



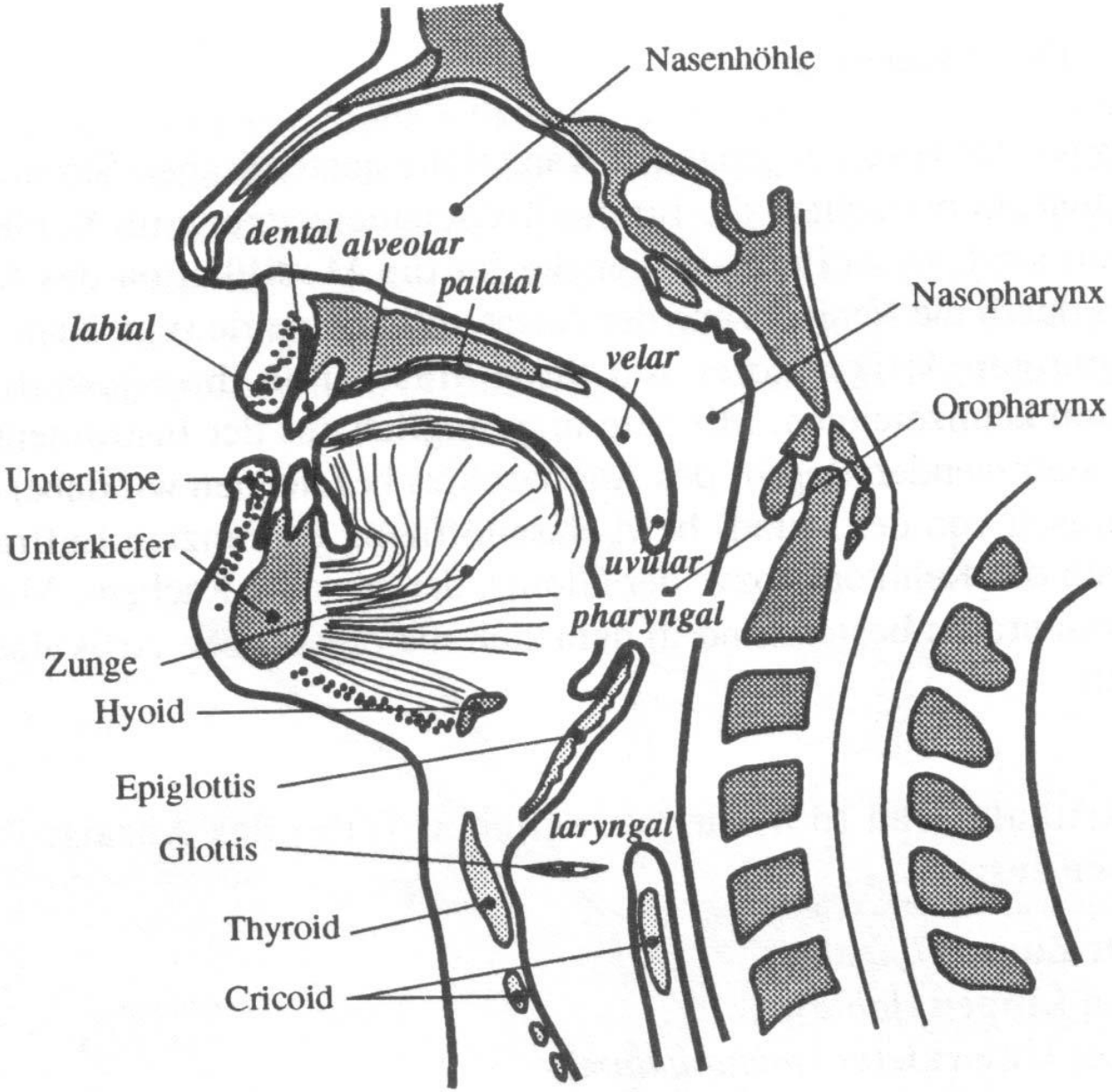
[Reetz,1999]

Speech production organs



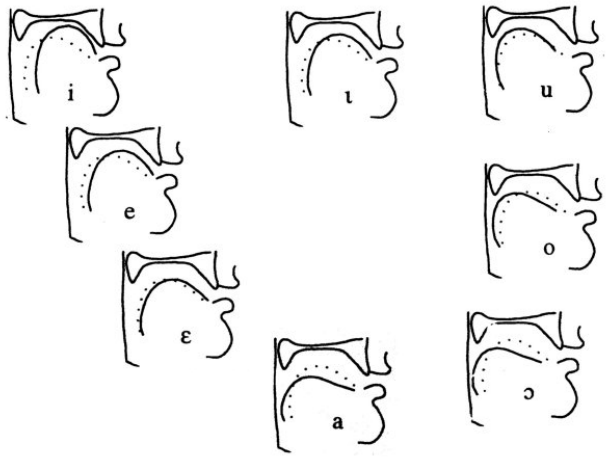
[Reetz,1999]

Speech production organs

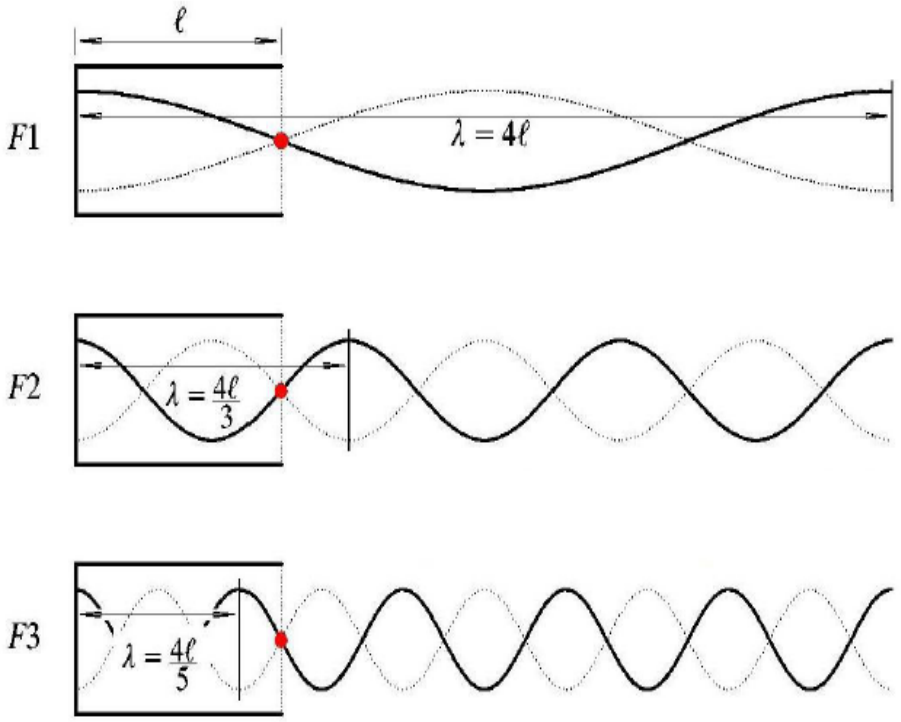


[Reetz,1999]

Acoustic Phonetics



Vocal tract geometry (tongue position)
of some English vowels



Sound pressure waves
of the first three formants

2.3.1 Ausgangspunkt Webster'sche Horngleichung (nach Ungeheuer, 1962)

Wir gehen nun von der Wellengleichung des Schnellenpotentials Φ für die Wellenausbreitung in einem Rohr veränderlichen Querschnittes, der sog. Webster'schen Horngleichung aus

$$\frac{\partial^2\Phi}{\partial x^2} + \frac{1}{A}\frac{\partial\Phi}{\partial x}\frac{dA}{dx} = \frac{1}{c^2}\frac{\partial^2\Phi}{\partial t^2} \tag{45}$$

mit den bekannten Randbedingungen:

$$v(t) = 0 \quad \Rightarrow \quad \frac{\partial\Phi}{\partial x} = 0 \quad [\text{Glottis}, x = 0] \tag{46}$$

$$p(t) = 0 \quad \Rightarrow \quad \Phi = 0 \quad [\text{Mundöffnung}, x = l] \tag{47}$$

Mit Hilfe der Trennung der Variablen

$$\Phi(x,t) = \varphi(x)\cdot\psi(t) \tag{48}$$

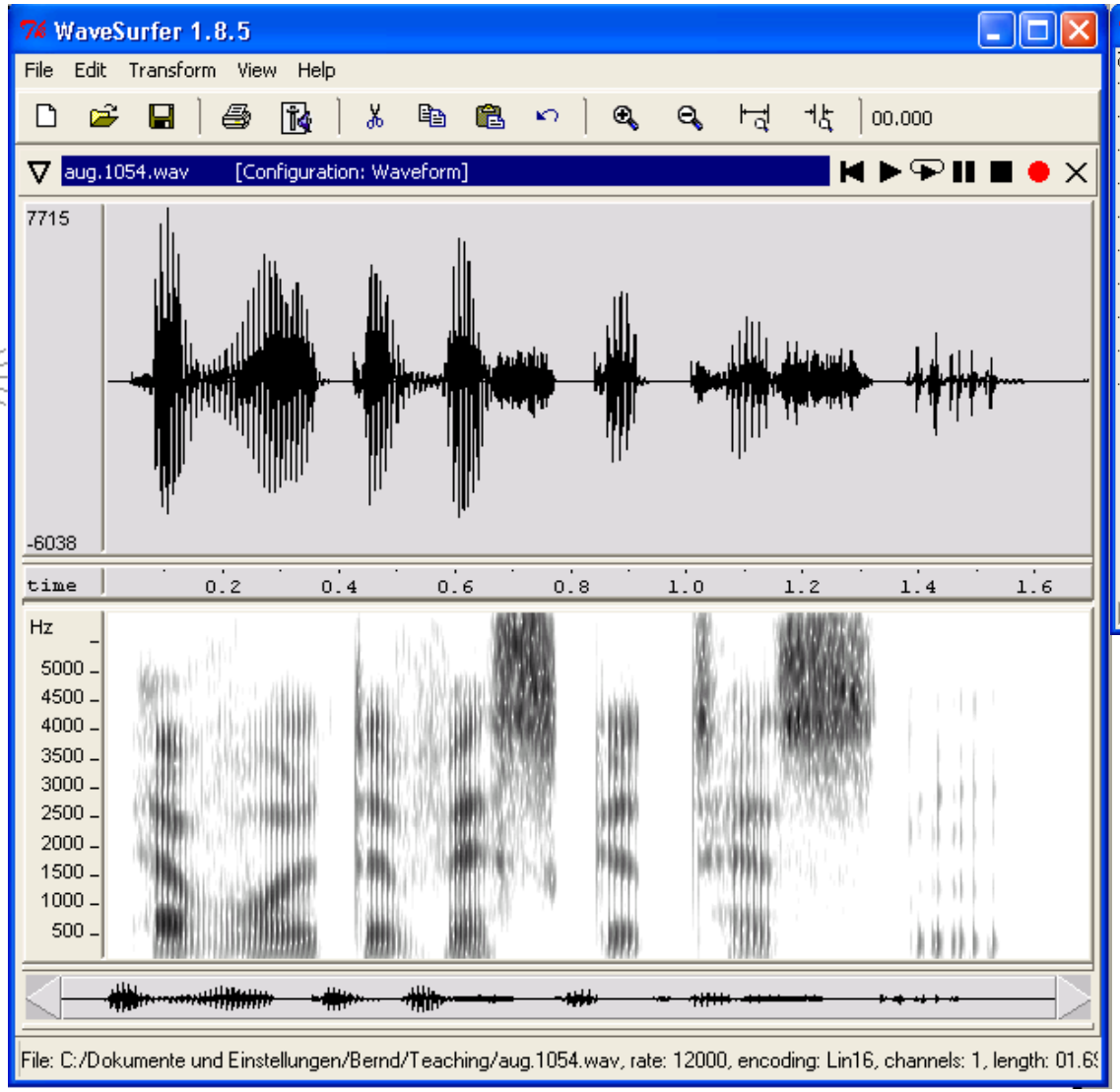
können wir (45) schreiben

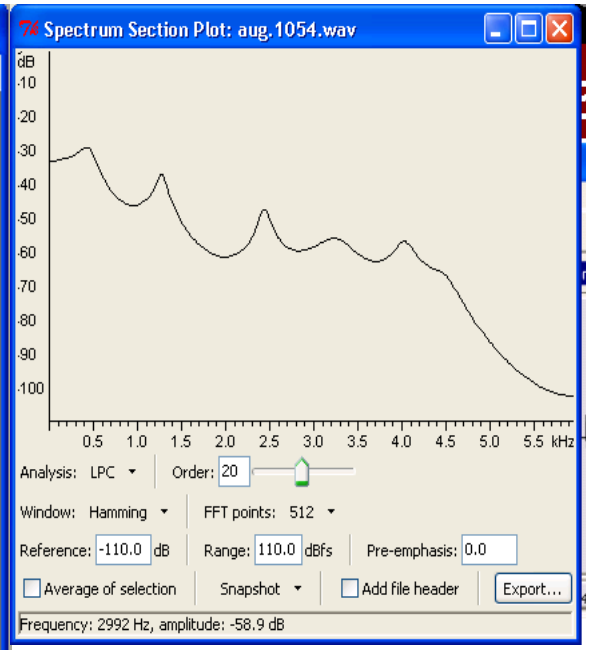
$$\frac{1}{\varphi}\left[\frac{d^2\varphi}{dx^2} + \frac{1}{A}\frac{d\varphi}{dx}\frac{dA}{dx}\right] = \frac{1}{c^2\psi}\frac{d^2\psi}{dt^2} \tag{49}$$

Die linke Hälfte hängt nur von x ab, die rechte nur von t. Damit können beide als gleich einer Konstante gesehen werden, die mit $-\Lambda$ bezeichnet sei:

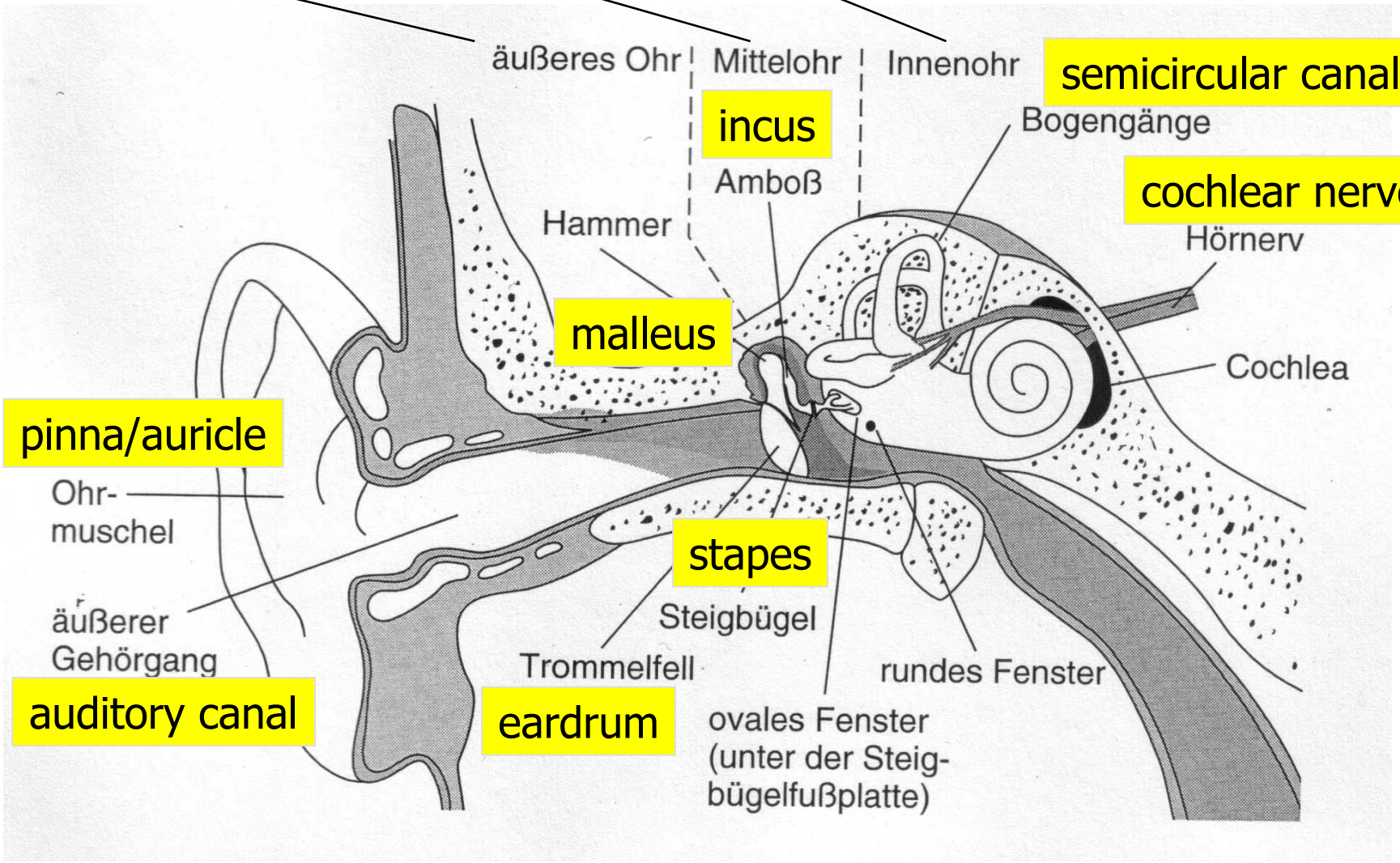
$$\frac{1}{\varphi}\left[\frac{d^2\varphi}{dx^2} + \frac{1}{A}\frac{d\varphi}{dx}\frac{dA}{dx}\right] = -\Lambda = \frac{1}{c^2\psi}\frac{d^2\psi}{dt^2} \tag{50}$$

Acoustic speech signal





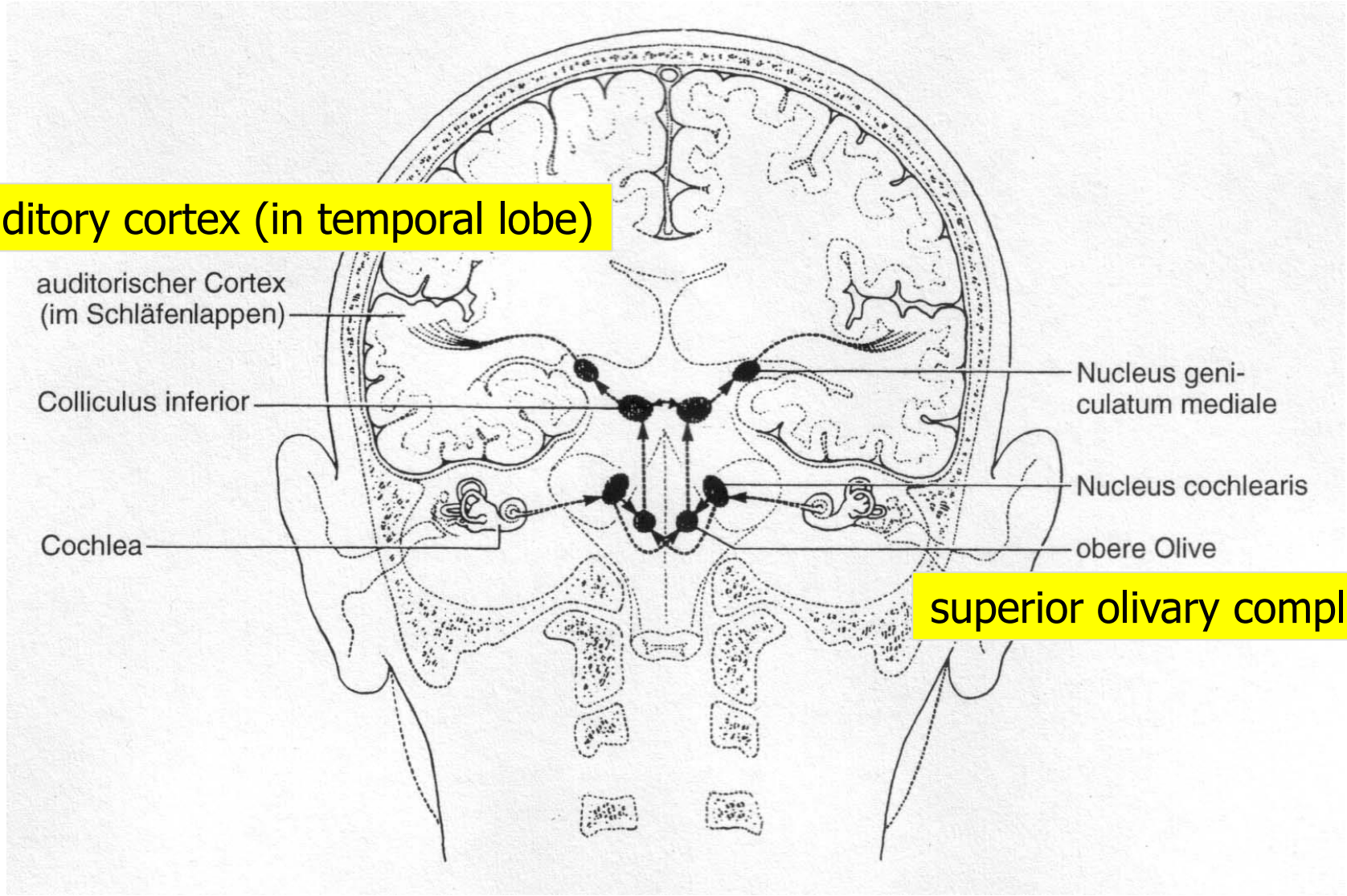
Outer ear, middle ear, inner ear



[Goldstein, 1997, p.322]

Connections in auditory system

auditory cortex (in temporal lobe)



superior olivary complex

[Goldstein, 1997, p.327]

Phoneticians

- What do phoneticians do, anyway?
  - *observe* how people say things
  - *describe* language on the level of pronunciation
  - *measure* properties of spoken language, pronunciation events
  - *model* pronunciation behavior and speech processing
  - *explain* the communicative contribution of pronunciation patterns
  - construct *theories, hypotheses and models* of phonetic events – and test them experimentally

Phoneticians and speech corpora

- Perform technical recordings of spoken language.
  - Choice: language/variety, speaker, type of signal.
  - This choice determines the types of analysis.
    - *Language:* speech sounds, precise or informal speech; monologue, discourse, dialogue
    - *Speaker:* (e.g., dialectal, regional or "standard" speaker)
    - *Signal:* acoustic=microphone, electromyographical, physiological, neurological (EGG, EPG, MEG, fMRI)
  - The type of signal determines the experimental design: only the acoustic signal, and perhaps not even that, makes natural recordings possible.

Some application areas of phonetics

- An understanding of the mechanisms of spoken language, i.e., of the processes of speech production and perception, is indispensable for
  - learning and teaching foreign languages
  - pronunciation dictionaries
  - speech pathology and language and speech disorders, clinical phonetics
  - forensic phonetics
  - speech technology (automatic speech and speaker recognition, speech synthesis, speech-to-speech translation, dialog systems)

Phonetic transcription - IPA

- Phonetic transcription (DE, standard text)
  - "Einst stritten sich Nordwind und Sonne…"
  - [ˈʔaɪns ʃtʁɪtn zɪç ˈnɔɐtvɪnt ʊnt ˈzɔnə]
- IPA = International Phonetic Association
  - established 1886 in Paris
  - *Aim:* universal classification system for all speech sounds
  - *Aim:* universal phonetic alphabet, to describe all speech sounds of all languages
  - most recent major revision: Kiel 1989 (alphabet: 2005)
  - IPA home page

More on the IPA system in Omnia's exercise session on Nov 6!

Suggested readings

- John Clark, Colin Yallop, Janet Fletcher (^{3}2007): An Introduction to Phonetics and Phonology. Blackwell.
- Victoria Fromkin, Robert Rodman, Nina Hyams (^{9}2011): An Introduction to Language. Wadsworth. Chapter 4.
  [covers basic articulatory phonetics only]
- IPA (ed.) (1999): Handbook of the International Phonetic Association. Cambridge University Press. IPA Handbook

Thanks!