M.Sc. LST
Speech Science

Psychoacoustics

Jan 25, 2024

Bernd Möbius & Omnia Ibrahim

Language Science and Technology
Saarland University

UNIVERSITÄT DES SAARLANDES

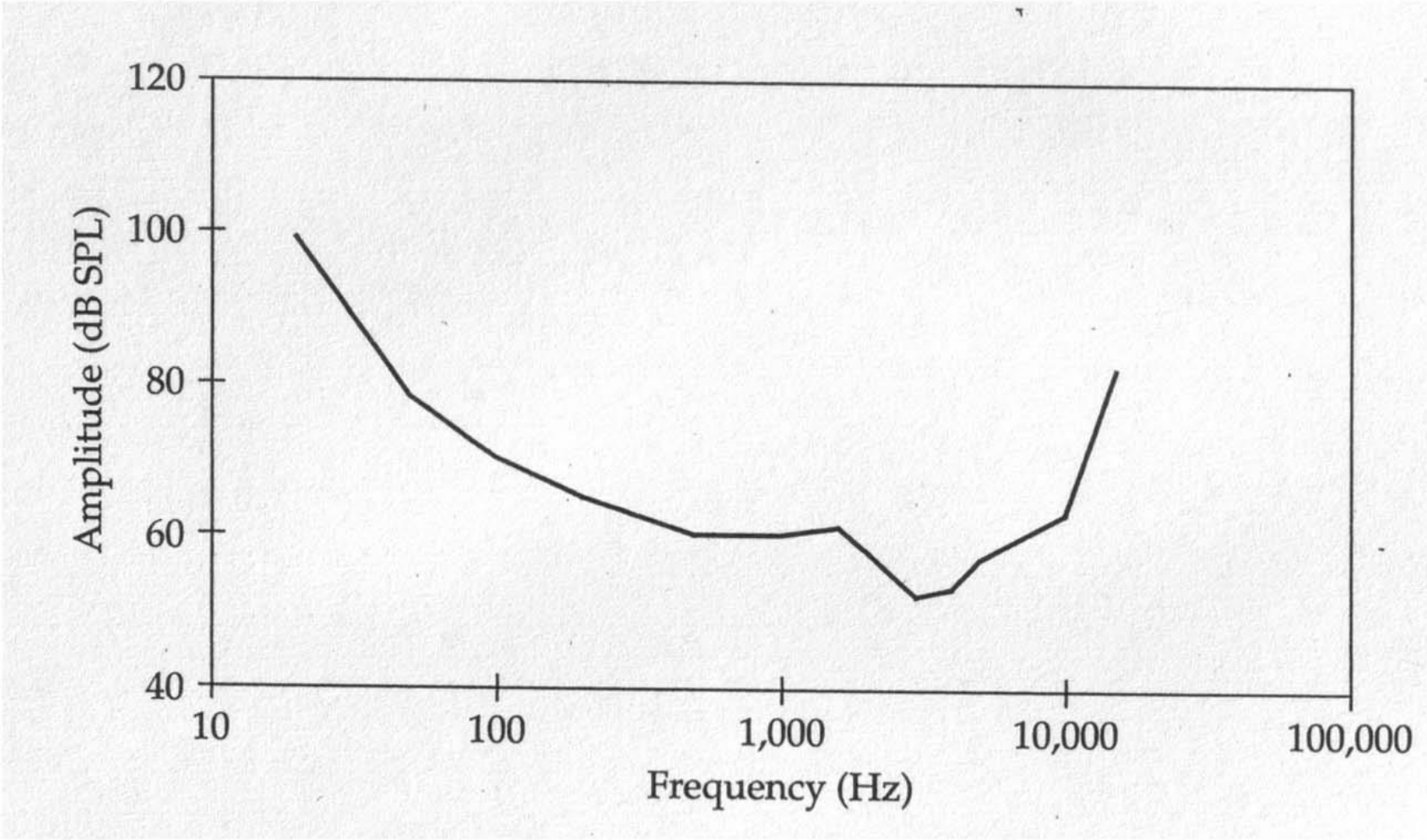
Overview

- Anatomy and physiology of the auditory system
- Speech perception
  - auditory perception
  - psychoacoustics
  - auditory-perceptual phonetics

Perception of loudness

- Loudness: perceptual correlate of acoustic sound intensity
  - differences in loudness are perceived on a logarithmic scale (e.g., decibel/[dB]) by the auditory system
    - 0 dB is equivalent to sound pressure level of a reference signal (at perceptual threshold at 1 kHz)
    - doubling of loudness is equivalent to increase by 10 dB
  - perceptual threshold ("just noticeable difference", JND) for pure tones: approx. 1 dB

Equal-loudness contour

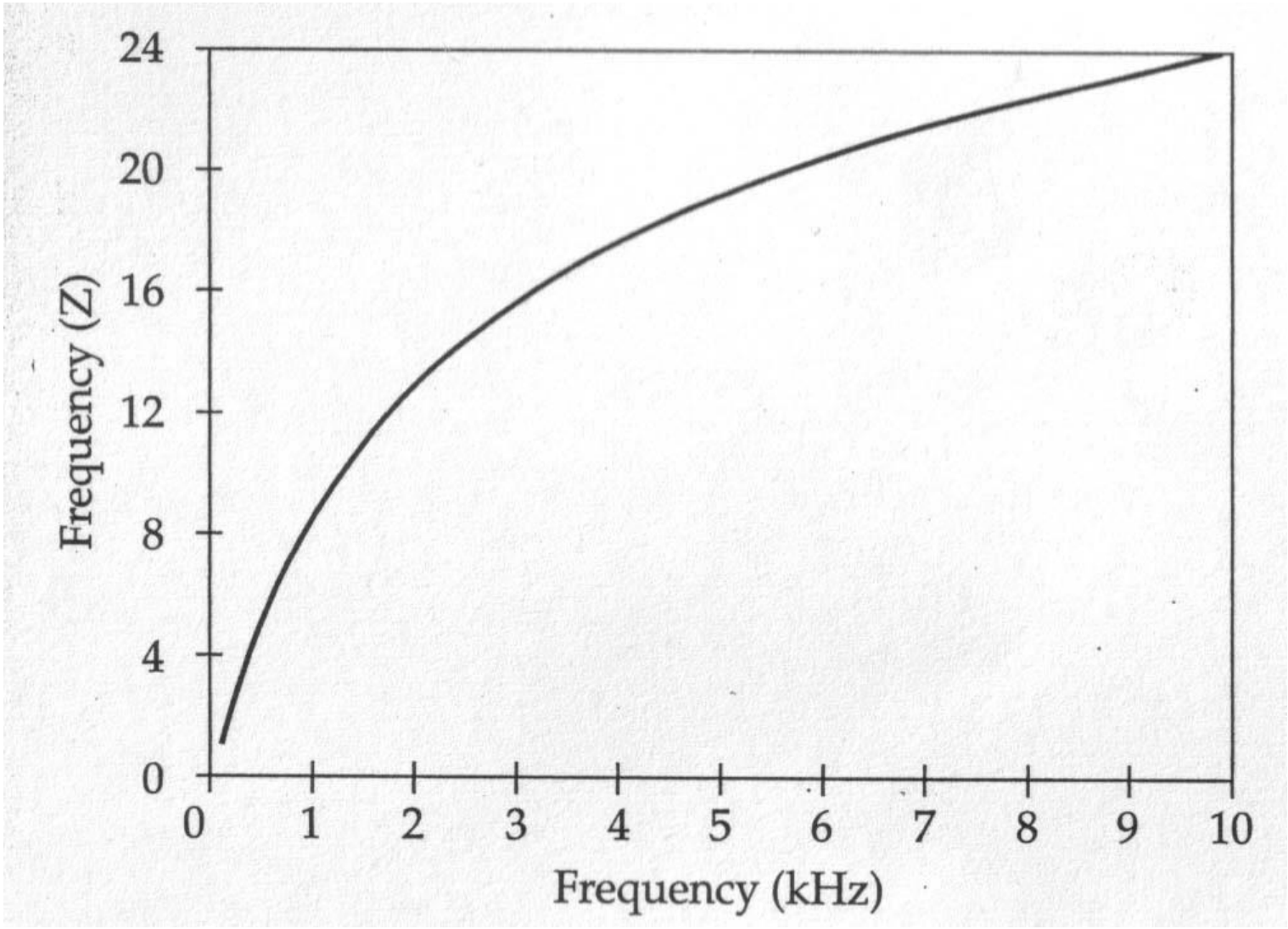


[Johnson, 1997, p.54]

Perception of pitch

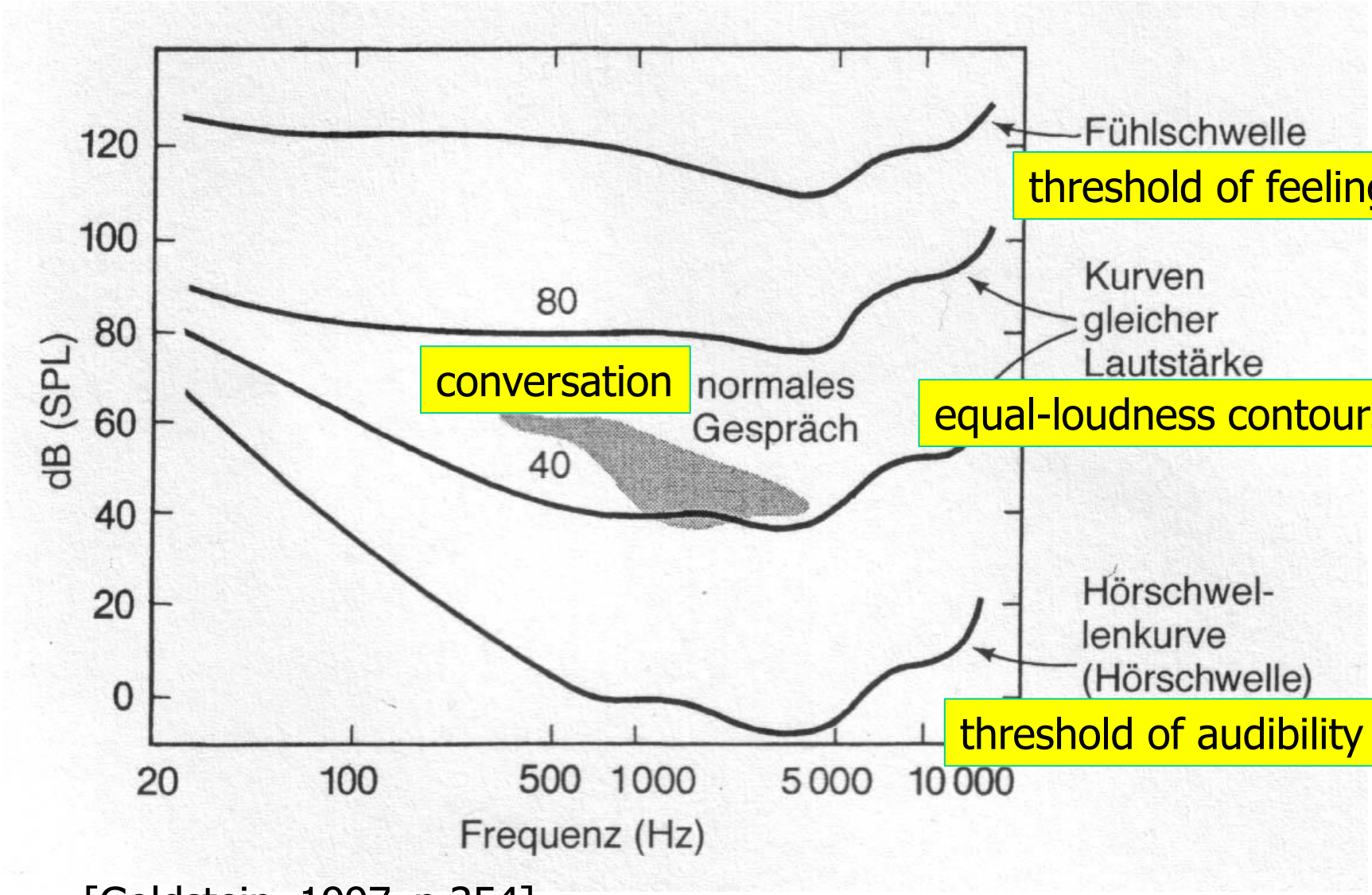
- Pitch: perceptual correlate of acoustic frequency
  - auditory frequency range: approx. 20 – 20,000 Hz
  - frequency selectivity: resolution of frequency components of a complex (e.g., speech) signal:
    - optimal below 500 Hz
    - logarithmically decreasing above 500 Hz
  - JND for pure tones:
    - below 1000 Hz: approx. 0.5%
    - at higher frequencies: approx. 5%
  - auditory frequency scale (e.g., Bark [Z]): auditory system is more sensitive to frequency differences in low frequencies than in high frequencies

Auditory frequency scale



[Johnson, 1997, p.55]

Auditory plane and auditory thresholds



threshold of feeling

conversation

equal-loudness contours

threshold of audibility

[Goldstein, 1997, p.354]

Loudness differences

Sound source	Sound pressure [μPa]	SPL [dB]
audibility threshold	20	0
soft whisper	200	20
quiet office	2 000	40
normal conversation	20 000	60
city bus	200 000	80
subway train	2 000 000	100
heavy thunder	20 000 000	120
pain	200 000 000	140

- Sound pressure (Schalldruck), measured in micro-Pascal
  - objective measurement of sound pressure differences
- Sound pressure level (Schalldruckpegel), measured in dB
  - subjective sensation of loudness differences

Perception of duration

- Duration: perceptual correlate of physical property "time"
  - temporal processing and resolution:
    - over which temporal interval can the auditory system integrate information?
    - detection of gaps in otherwise continuous signals
  - JND:
    - duration differences: >20 ms at 500 - 1500 Hz
    - detection of gaps of 6 – 8 ms
    - fast spectral changes within <30 ms are not analyzed, but perceptually integrated

Speech perception

- Psychoacoustic properties of auditory system are compliant with requirements of speech perception; e.g.:
  - very good frequency resolution in low-frequency range
    $\rightarrow$ fundamental frequency analysis (voicing, intonation)
  - medium frequency resolution in mid-frequency range
    $\rightarrow$ formant analysis and tracking (vowels, sonorants)
  - poor frequency resolution in high-frequency range
    $\rightarrow$ rough spectral patterns (fricatives)
  - temporal integration supports capturing coarticulation, and good temporal resolution supports recognition of stop releases (but stop bursts are too short for spectral analysis)
- Co-evolution, or adaptation of production system to auditory system?

Speech intelligibility

- Intelligibility of speech depends on many factors, including:
  - frequency band (e.g., telephone 350 – 3500 Hz)
  - loudness
  - duration of segments of speech and gaps
  - semantic content (top-down processing) and semantic predictability
  - robustness of speech signals
    - gaps <200 ms hardly disturb intelligibility
    - gaps >500 ms destroy intelligibility
  - disturbing noise (signal-to-noise ratio, SNR)

Thanks!