

King Saud University

College of Engineering

IE – 341: “Human Factors”

Fall – 2015 (1st Sem. 1436-7H)

Human Capabilities

Part - B. Speech Communications

(Chapter 7)

Prepared by: Ahmed M. El-Sherbeeney, PhD

Lesson Overview

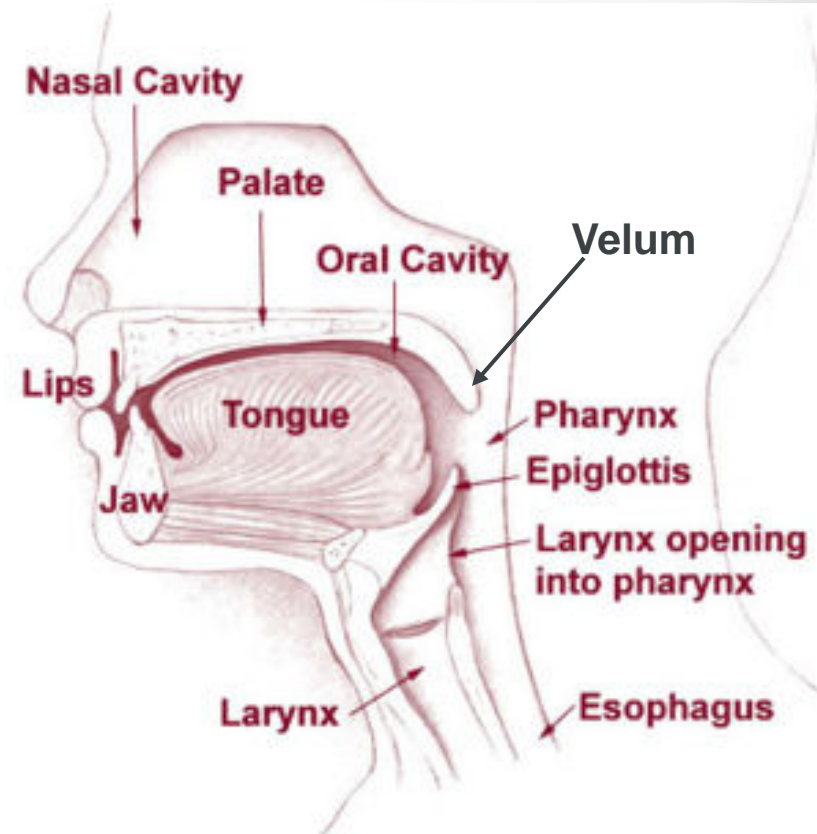
- Introduction
- The Nature of Speech
- Criteria for Evaluating Speech
- Components of Speech Communication Systems

Introduction

- Speech is form of “display”
 - i.e. form of auditory information
- Source of speech
 - Mostly human (focus of this lesson)
 - Could also be *synthesized*
 - i.e. machine; e.g. voice mail, access confirmation)
- Receiver of speech
 - Mostly human
 - Could also be machine: “voice recognition”
 - not advanced as synthesized sound

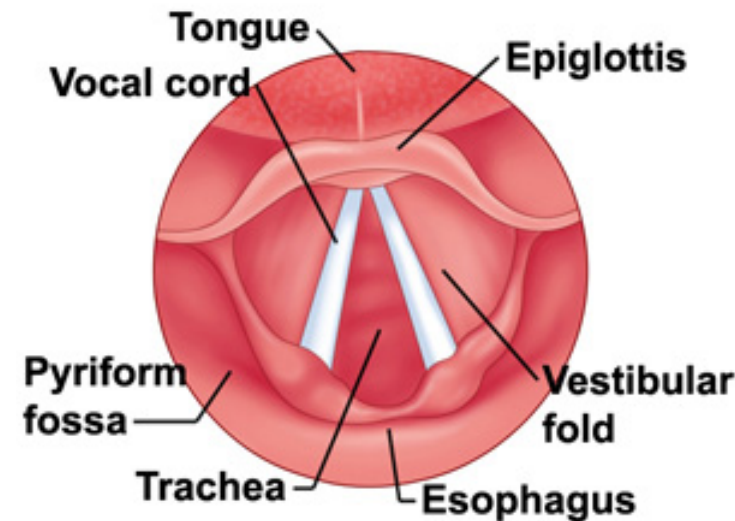
The Nature of Speech

- Speech: closely associated with breathing
- Organs associated with speech:
 - Lungs
 - Larynx
 - contains vocal cords
 - Pharynx
 - channel bet. larynx & mouth
 - Mouth (AKA: oral cavity):
 - tongue, lips, teeth, velum
 - Nasal cavity



Cont. The Nature of Speech

- Vocal cords
 - Contains vibrating folds
 - Opening between folds: glottis / epiglottis
 - Vibrates 80-400 times/sec.
 - Rate of vibration of vocal cords: controls freq. of resulting speech sounds
 - Watch "Vocal Cords in Action": www.youtube.com/watch?v=iYpDwhpILkQ
 - Speech/sound waves:
 - Produced by: vocal cords
 - Further modified by "resonators":
 - pharynx, oral cavity, nasal cavity
 - Further articulated by "manipulators":
 - Mouth: tongue, lips, velum
 - Nasal cavity: velum, pharynx muscles

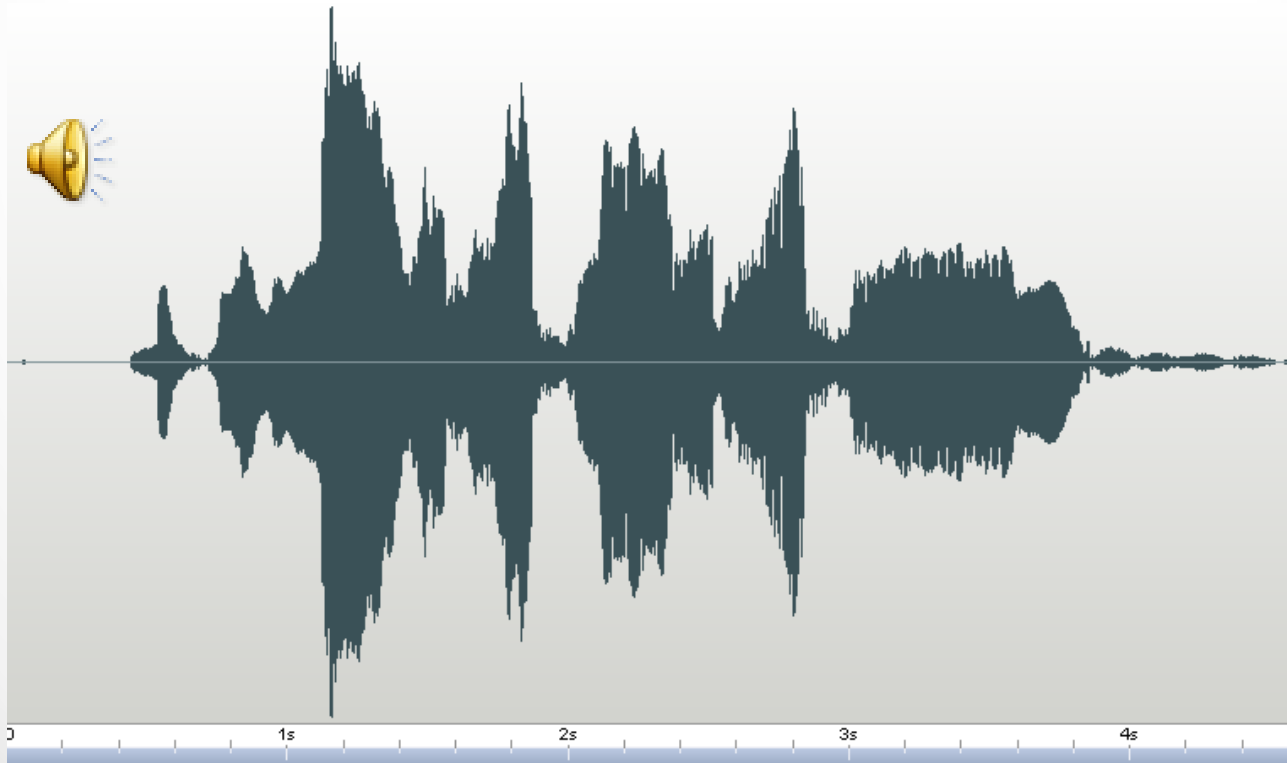


Cont. The Nature of Speech

- Types of Speech sounds
 - Phonemes
 - Basic unit of speech
 - Defⁿ: “shortest segment of speech which, if changed, would change the meaning of a word”
 - Phonemes in English language:
 - Vowel sounds: 13 (e.g. **u** sound in *put*, **u** sound in *but*)
 - Consonant sounds: 25 (e.g. **g** sound in *gyp*, **g** in *gale*)
 - Diphthongs (i.e. sound combinations):
e.g. **oy** sound in *boy*; **ou** sound in *about*
 - Can you compare these to Arabic phonemes?
 - Combining phonemes:
 - Phonemes form syllables ⇒
syllables form words (e.g. **ac·a·dem·ic**) ⇒
words form sentences
 - Note Phonemes > letters (why?): since phonemes change when combined together (e.g. **d** in *di* different than *du*)

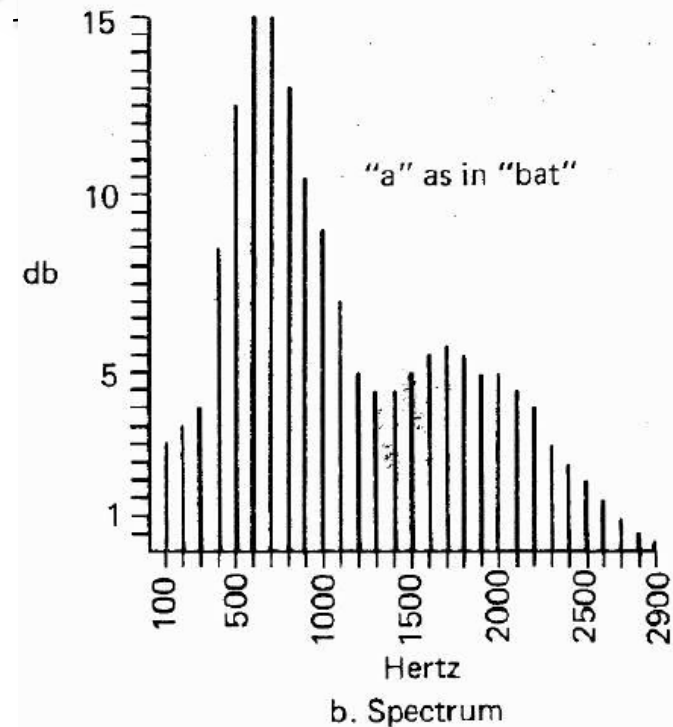
Cont. The Nature of Speech

- Depicting Speech
 - Sound is generated by variations in air pressure
 - This is represented in several graphical ways
 - Method 1: **waveform**
 - Shows intensity variation over time (relative scale)
 - Listen to file below for verse “بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ” *

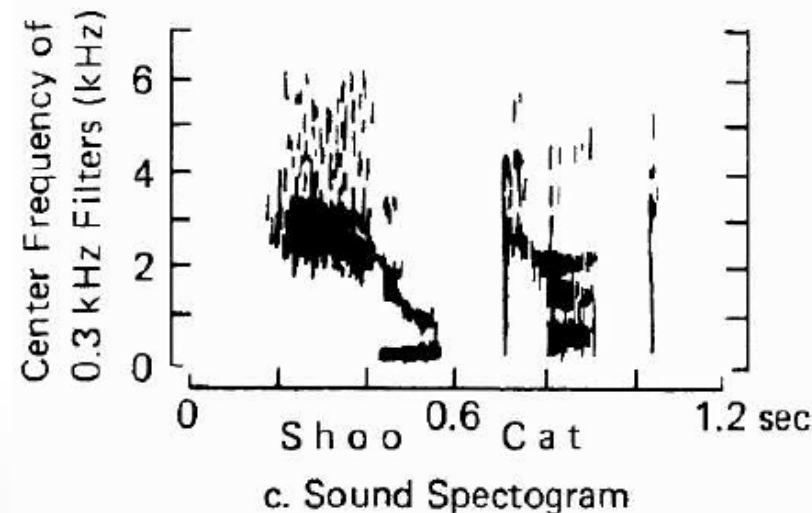


Cont. The Nature of Speech

- Cont. Depicting Speech
 - Method 2: **spectrum**
 - Shows for given phoneme / word: intensity of various frequencies in that sound sample (see right)
 - Which freq. has highest intensity in shown figure?

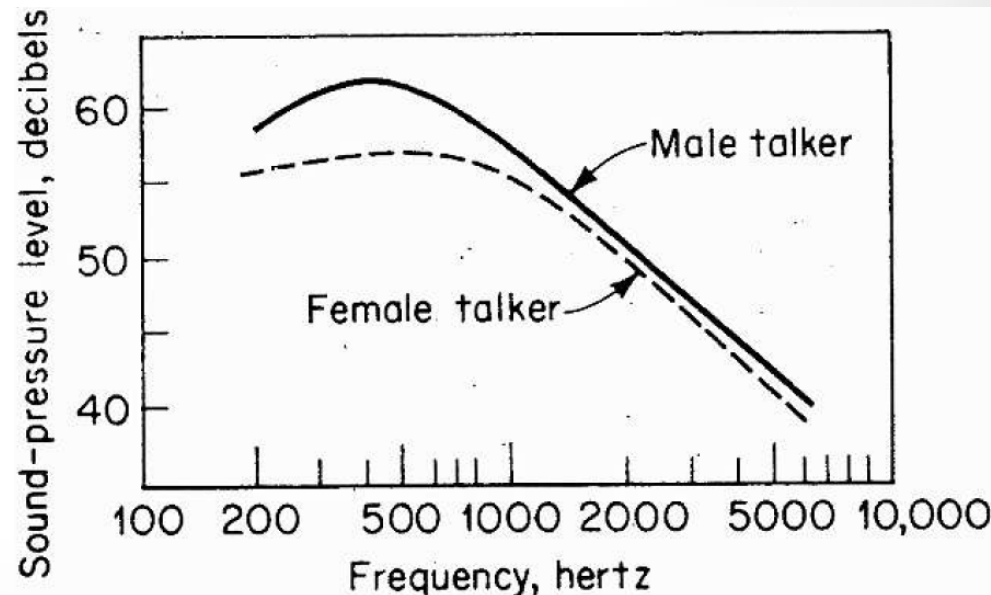


- Method 3: sound **spectrogram**
 - Frequency: vertical scale
 - Time: horizontal scale
 - Intensity: degree of darkness on plot (see right)



Cont. The Nature of Speech

- Intensity of Speech (AKA "Speech Power")
 - Variation among phonemes
 - Vowels speech power » consonants
 - e.g. **a** in "talk" has speech power: 680 times > **th** in *then* (i.e 28 dB difference)
 - Variation among speech types
 - conversational speech: 45-55 dBA*
 - Telephone/lecture speech: 65 dBA
 - Loud speech: 75 dBA
 - Shouting: 85 dBA
 - Variation: Male & Female
 - Male > female by 3-5 dB (in general)
 - Men in lower freq. has higher intensity than women (see right)



Criteria for Evaluating Speech

- Speech Intelligibility
 - Defⁿ: “degree/percentage to which a speech message (e.g. group of words) is *correctly* recognized”
 - This’s major criterion for evaluating speech
 - Assessment of speech intelligibility:
 - Either repeating back read material
 - Or answering questions regarding material
 - Speech Intelligibility tests:
 - Nonsense syllables (e.g. un, us, mus, sub, sud, ...)
 - these have least intelligibility
 - Phonetically balanced (PB) word lists
 - Nonsense syllables < words Intelligibility < sentences
 - Complete sentences
 - These have highest intelligibility, even when some words are not recognized (i.e. depends on context)
 - e.g. “Did you go to the store” may sound as “Dijoo ...”

Cont. Criteria for Evaluating Speech

- Speech Quality
 - Another criterion for evaluating speech
 - May be important in identifying a specific speaker
e.g. on phone (i.e. absolute identification)
 - Also important to choose bet. different products
e.g. speaker phone on home phones, mobile phones
 - Assessment of speech quality
 - Usually done using rating system
 - e.g. people listen to speech and asked to rate quality:
excellent, fair, poor, unacceptable, etc.
 - May also be done by comparing to some standard speech quality

Components of Speech Communication Systems

- Components
 1. Speaker
 2. Message
 3. Transmission System
 4. Noise Environment
 5. Hearer
- Discussed here in terms of
 - Effects on intelligibility of speech communications
 - Methods to improve intelligibility of system

Cont. Speech Communication Systems

1. Speaker

- Intelligibility of speaker usu. called "**enunciation**"
- Research found higher intelligibility is caused by:
 - **Longer syllable** duration
 - Speaking with **high intensity**
 - Making **use of speech time** with spoken words and little pauses
 - Variation of **speech frequencies**
- Differences bet. Intelligibilities generate from:
 - Structure of articulators (sound-producing organs)
 - Speech habits that people acquire
 - Speech training may improve speech intelligibility (but not very much)

Cont. Speech Communication Systems

2. Message

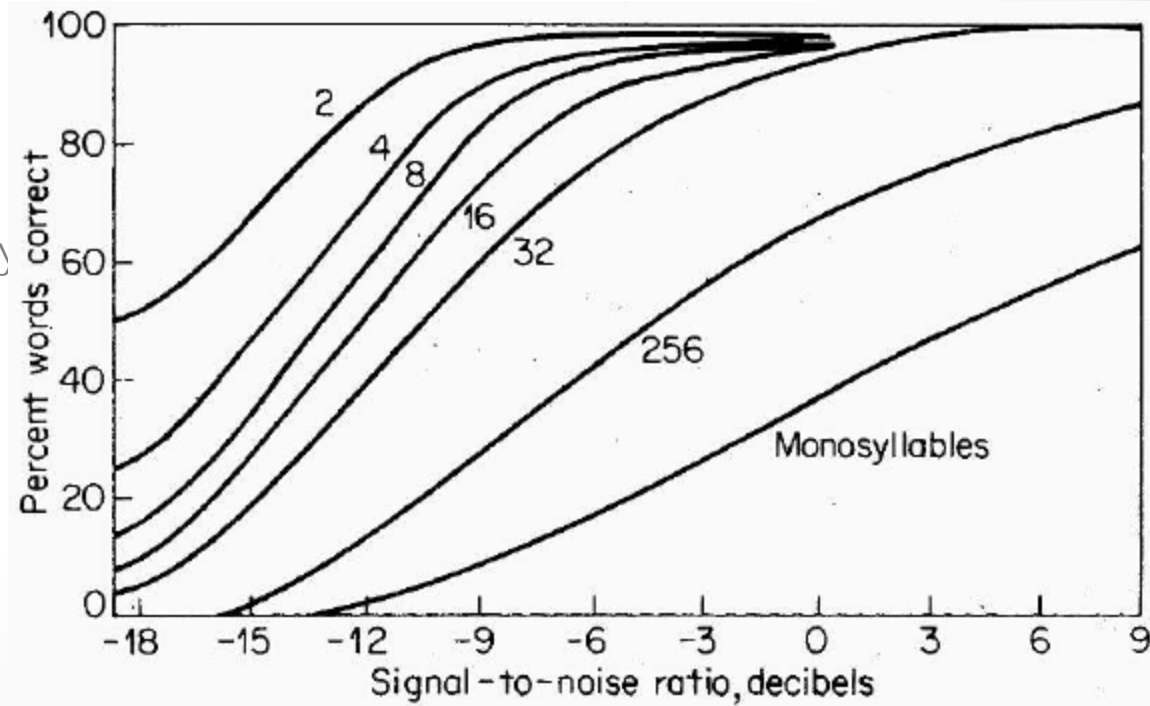
Affected by: phonemes used, words, context

- Phoneme Confusions
 - Some speech sounds more easily confused than others
 - e.g. letters in each group (consonants) can be confused with each other: **DVPBGCET, FXSH, KJA, MN**
 - Avoid using single letters in presence of noise
- Word Characteristics: for higher intelligibility use:
 - More **familiar words**
 - **Longer words**: for longer words even if part of word is dropped, rest can still be figured out
 - e.g. "word-spelling" alphabet: alpha, bravo, charlie, delta, ... instead of A, B, C, D

Cont. Speech Communication Systems

2. Cont. Message

- Context features: for higher intelligibility use:
 - **Sentences** (rather than words)
 - **Meaningful sentences** (rather than non-sense phrases)
 - e.g. "This book is great" rather than "is great book this"
 - **Less vocabulary** (words) in the presence of noise
 - More words with noise \Rightarrow less intelligibility (see below)
 - Note, -ve SNR means noise is more intense than signal
 - Also note, monosyllable: words with only one syllable (e.g. hit, ant, cube, fish)



Cont. Speech Communication Systems

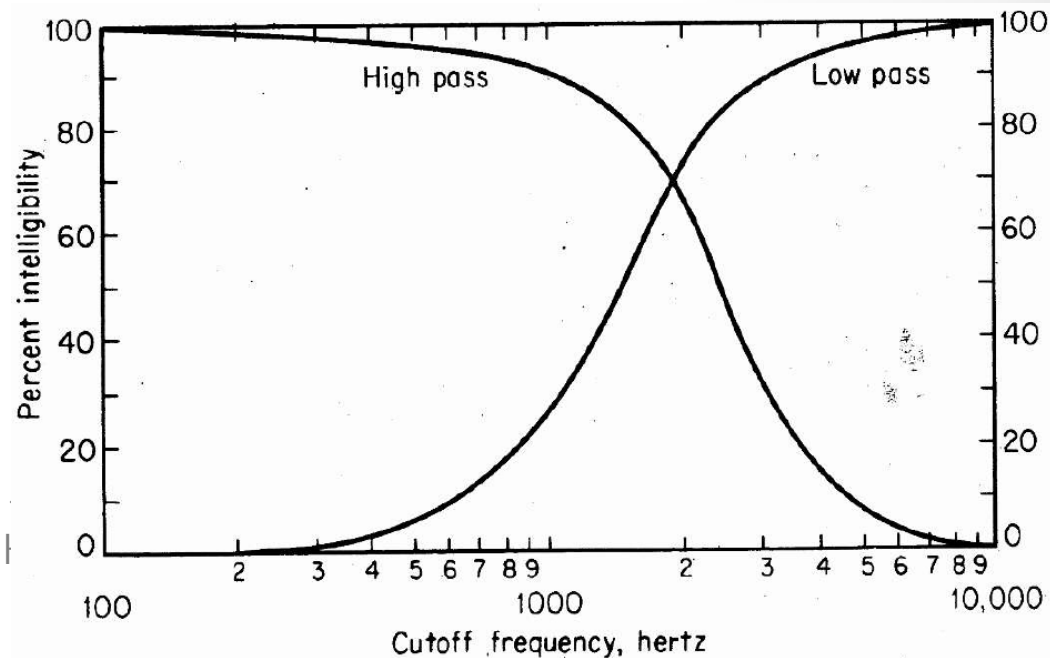
3. Transmission System

- Transmission Systems

- Natural: air
- Artificial: telephone, radio, etc.

- Artificial systems cause distortions, e.g.

- Frequency distortion
- Amplitude distortion
- Filtering
 - Low-pass filter: eliminates freq. above some level
 - High-pass filter: eliminates freq. Below level
 - Filtering: freq. > 4000 Hz, about > 1000 Hz, < 3000 Hz



Cont. Speech Communication Systems

4. Noise Environment

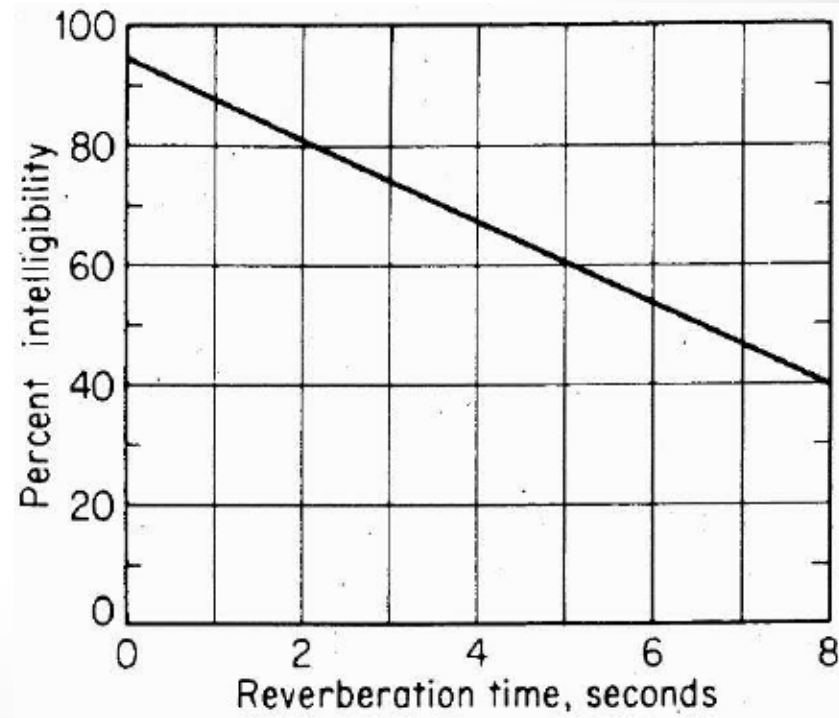
- causes biggest harm to speech intelligibility
- **SNR** (signal to noise ratio):
 - Simplest way to evaluate impact of noise on intelligibility
 - Study: for noise level of 35-100 dB \Rightarrow SNR = 12 dB for threshold of intelligibility (what to do for loud noise?)
 - However, SNR does not take frequency into consideration (only intensity)
- Other measures (taking freq. into consideration):
 - Articulation index (AI): a measure (0-1) of speech intelligibility while knowing the noise environment
 - Preferred-octave speech interference level (PSIL): rough measure of effect of noise on speech reception
 - Preferred noise criteria (PNC) curves: suggest acceptable noise level for different work environments (e.g. offices)

Cont. Speech Communication Systems

4. Cont. Noise Environment

o Reverberation:

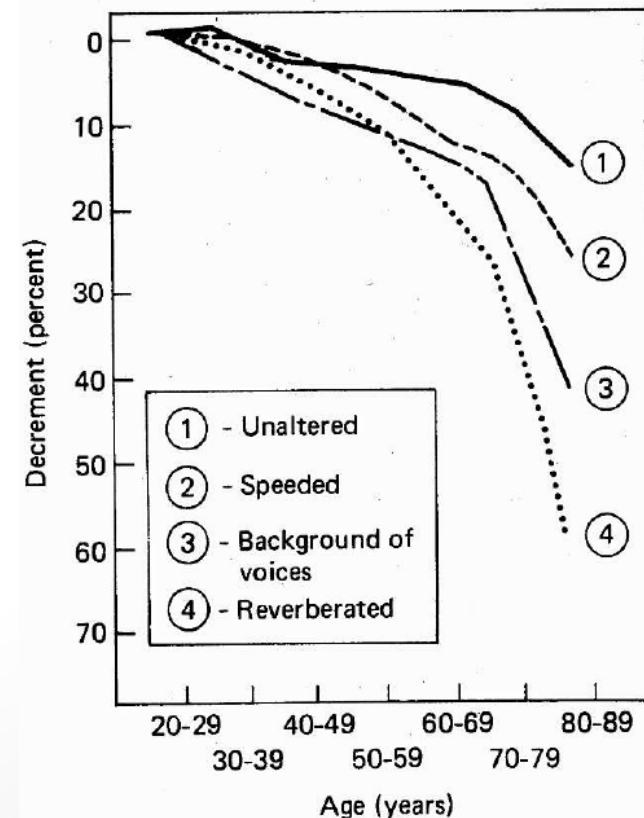
- Bouncing effect of noise from walls, floor, ceiling in a closed room
- Greatly decreases speech intelligibility (e.g. classrooms)
- In general, the longer the reverberation time, the more the speech intelligibility decreases
- Examine the linear relation (right) for decaying a 60 dB noise



Cont. Speech Communication Systems

5. Hearer

- To receive speech under noise: hearer should
 - Have normal hearing
 - Be trained to receive messages
 - Be able to withstand stress of situation
- Age
 - Also affects speech reception (i.e. intelligibility); see right
 - 20-29 age group: base level
 - Note, unaltered speech: 120 wpm vs. speeded speech: 300 wpm
- Hearing protection
 - Prevents hearing loss
 - May improve SI for noise >80 dBA
 - Decreases SI for noise <80 dBA



References

- *Human Factors in Engineering and Design*. Mark S. Sanders, Ernest J. McCormick. 7th Ed. McGraw: New York, 1993. ISBN: 0-07-112826-3.