



# Automatic Speech Recognition

*Kjell Elenius*

*Speech, Music and Hearing, KTH*



# Overview

- Introduction
- Problems and challenges
- Acoustic analysis methods
- Speech recognition techniques
  - Pattern matching
  - Time warping (dynamic programming)
  - Neural networks
  - Hidden Markov Models
  - Viterbi search
  - Training techniques
  - Adaptation techniques
- Some results



# Why speech?

- Natural way of communication
  - Requires no special training
- Leaves hands and eyes free
  - Replaces keyboards
- Effective
  - High data rate
- Can be transmitted/received inexpensively
  - (mobile) phones



# Automatic Speech Recognition, ASR

- An ASR system converts a speech signal into words
- The recognized words can be
  - The final output, or
  - The input to natural language processing



# Application Areas for Speech Based Interfaces

- Mostly input (recognition only)
  - Simple command and control
  - Simple data entry (over the phone)
  - Dictation
- Interactive conversation (understanding needed)
  - Information kiosks
  - Transactional processing
  - Intelligent agents



# Challenges – variability

