COM4211/6451
SPEECH PROCESSING

Lecture 1
INTRODUCTION

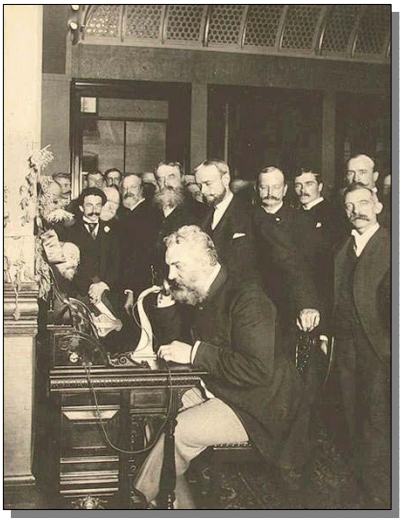
The University Of Sheffield.

The Course

- Two parts
  - speech
  - speech processing
- Both halves running in parallel
  - Mondays: *Prof. Roger Moore*
  - Tuesdays: *Dr. Thomas Hain*
- Assessment
  - exam at the end of Semester 1
- Lecture notes (these slides)
  - will be on-line after the lecture

The University Of Sheffield.

What is Speech Processing ?



What is Speech Processing ?

- Speech processing is the study of speech signals and the processing methods of these signals
- The signals are usually processed in a digital representation whereby speech processing can be seen as the intersection of digital signal processing and natural language processing

http://en.wikipedia.org/wiki/Speech_processing

Speech Processing Technologies

Digital Speech Coding

Automatic Speech Recognition

Text-to-Speech Synthesis

Spoken Language Dialogue Systems

Why Speech Processing ?

- Lots of applications

  … especially in Science Fiction

The 'Conversational' Computer

Why use Speech ?

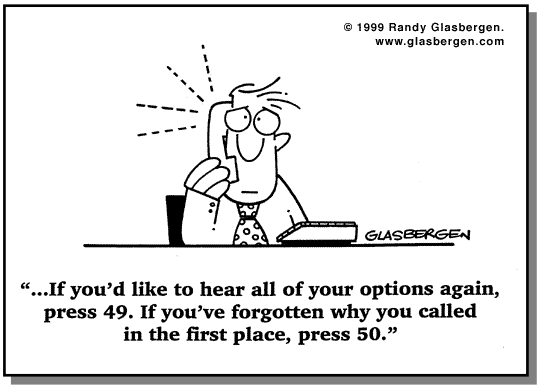
"Speech is the 'natural' way to interact with your computer."

Speech *may* be a more intuitive way of accessing information, controlling things and communicating …

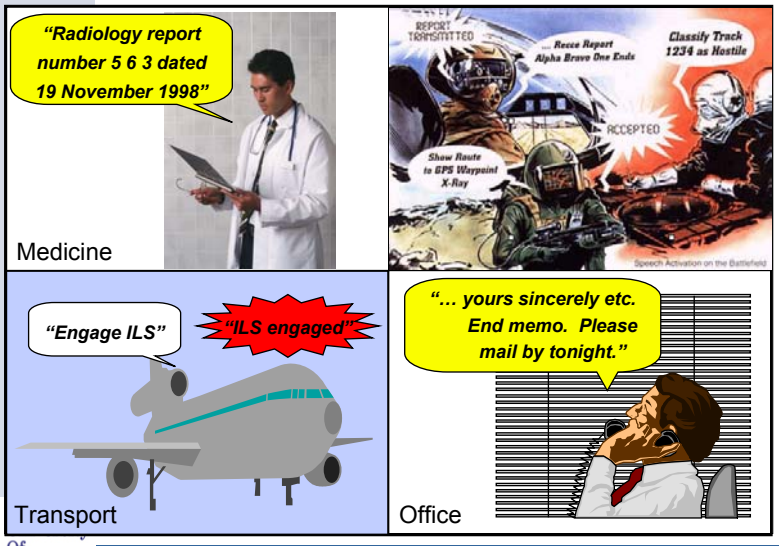
but there may be viable alternatives.

Why use Speech ?

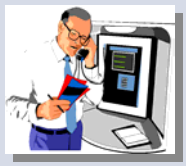
Some alternatives can be problematic …

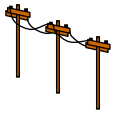


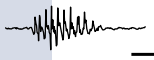
Markets & Applications



Extracting Information from Speech







Speech Signal

Speaker Recognition → Speaker Name *John Q. Public*

Gender Identification → Gender *Male* or *Female*

Language Recognition → Language Name *English*

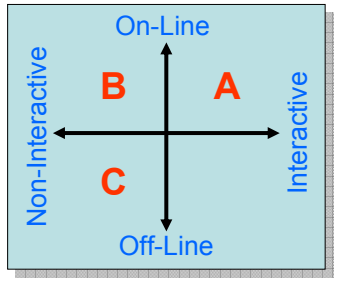
Speech Recognition → Words *"How are you?"*

Slide courtesy of Doug Reynolds, MIT Lincoln Lab.

PHILIPS

Taxonomy of Applications

- Interactive *vs* non-interactive
  - conversational human-machine interface (A)
  - monitoring speech communications (B)
- On-line *vs* off-line
  - 'live' speech processing (B)
  - speech data mining (C)



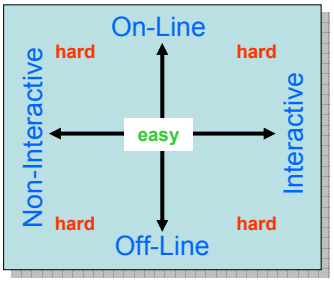
Question …

What current or future applications of speech processing can you think of ?

Speech Processing Application

Describe your proposed application ….

Place an 'X' on this diagram that corresponds to your application …



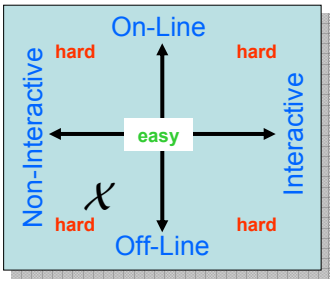
Name: _____

Speech Processing Application

Describe your proposed application ….

Being able to search voice notes recorded on my PDA

Place an 'X' on this diagram that corresponds to your application …



Name: *Roger Moore*

To beat the EU ban on right-hand drive cars BMW uninvents the wheel.

By the end of 2007 you will not be allowed to use a right-hand drive car on the roads of mainland Europe.

It's a ruling BMW has vigorously opposed, but our lawyers were eventually routed and it was left to our engineers to fight a rearguard action.

Their riposte was one of startling élan: hands-free s

It uses a combination of sensors and VAT (Voice Ac Technology) and does away with the steering wheel alto

All the dials and controls are mounted in the centre dash on a pivoting section which can be angled towards of the front seats.

While crossing the Channel, simply tilt the instrumen to the left and change seats with your passenger. Ped recessed into both footwells, the relevant set becoming when the position of the central section is fixed.

The sensors work as an intermediary between the

eyes and the road and combine with voice commands to steer the car wherever you want it to go in complete safety.

Early prototypes were prone to sudden U-turns if the driver swung round to shout at the children in the back, but a satellite monitoring system developed by Dr Bitt-Fischi, our head of R&D,

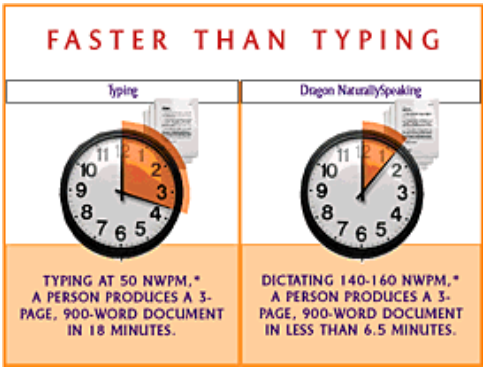
real

Some ~~example~~ applications …

Command & Control

Robust Speaker-Independent Small-Vocabulary Automatic Speech Recognition

Document Creation



Speaker-Dependent Large-Vocabulary Continuous Speech Recognition

Virtual Newsreader

Text-to-Speech Synthesis

Location-Based Services

YEOMAN

Text-to-Speech Synthesis

Information Services

Telephone-Based Speaker-Independent
Medium-Vocabulary
Spoken Language Dialogue System

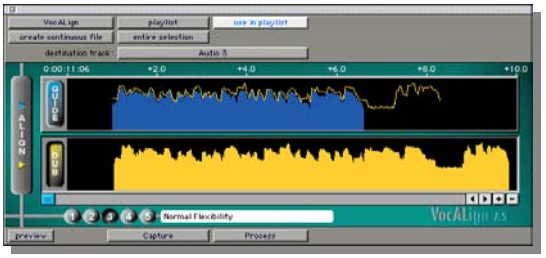
Live Captioning

Language Learning

Structured Conversational Interface

Child's Reading Aid

Film Dubbing



GUIDE	GUIDE	GUIDE
DUB	DUB	DUB
MIX	MIX	MIX
ALIGN	ALIGN	ALIGN

Games

‘Live’ Demonstration

LINGUISTIC PROCESSING IN THE *SPLATWRITER*™ VOICED HYPE WRITER

by Adey Jamuser

President, SPLAT Systems Inc., UK

ABSTRACT

It is well known that a sin tax can act as a useful constraint in speech recognition. Everybody knows egg samples of a coo. Stickly similar pears like open quotes. Wreck a nice beach' and 'wreck an iced bee twitch are semantically different. Close quotes after the second beach. In *SPLATWRITER*™ we have developed a sadistically motivated bottom. Up a brooch, variations in the fanatic lettuce are modelled by a sarcastic grandma witchy, strained by a stimmulated and kneeling parson. Preference is given to words which are. Free Quentin the famous slob! Cor! puss, amazing nudist play! Technology has—been included for awful graphic output with each, you nit.

Aargh! *** this is ***, or full. It's Terry Bull. How do I turn this *** thing off?. Our *** competitors must never get twinned. A fit! Can we chain Jenny Thing before we're illicit publicly? End occupant. No you stupid machine, I said end document.

END-OF-DOCUMENT>

Taken from Inst. Acoustics ‘Speakeasy’, October 1988.

The Advantages of Speech

- hands-free
- eyes-free
- fast
- Intuitive

Fran Capo

World Record Holder

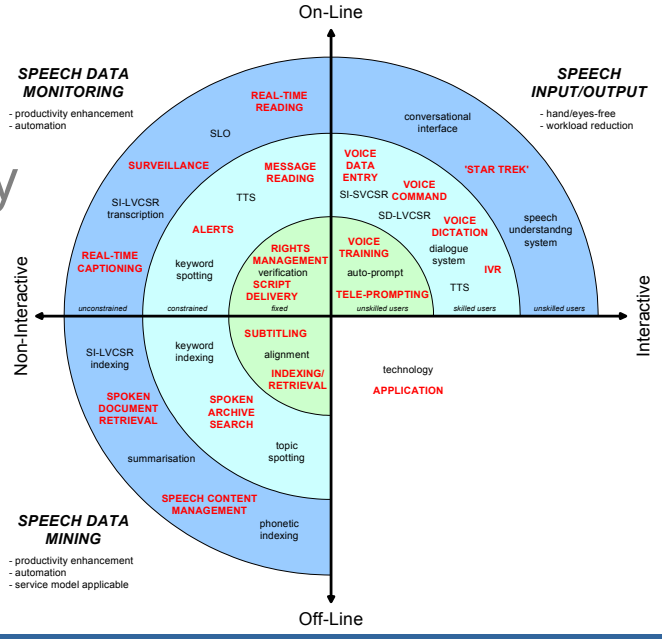
603.32 wpm

"You have been learning since birth the only skill needed to operate our equipment."

Application Benefits

- cost savings through automation
- human resource savings
- increased operational effectiveness
- increased productivity
- workload reduction
- increased security
- increased safety
- increased functionality
- space saving
- improved quality of life
- etc. …

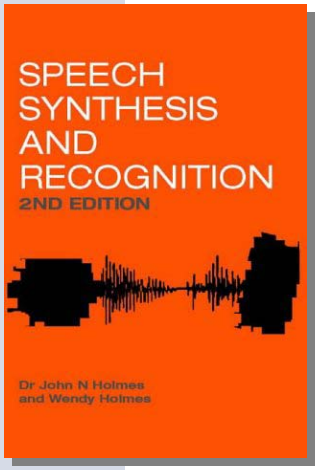
Summary

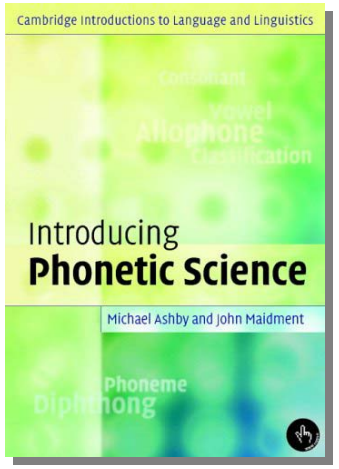


The Rest of the ‘Speech’ Course

- Speaking and listening
- The nature of speech
- Sounds and symbols
- Phonetics
- Phonology
- Prosody
- Models

Recommended Text Books





Resources on the Web

- International Speech Communication Association (ISCA)
  http://www.isca-speech.org/
- Comp.Speech
  '*Frequently Asked Questions*'
  http://svr-www.eng.cam.ac.uk/comp.speech
- Speech Technology Magazine
  http://www.speechtechmag.com

Conferences

- ISCA INTERSPEECH
- IEEE international Conference on Acoustics, Speech and Signal Processing (ICASSP)
- SpeechTEK
- International Conference of the Phonetic Sciences (ICPhS)

Scientific Journals

Software Resources

- Praat: doing phonetics by computer
  http://www.fon.hum.uva.nl/praat/
- WaveSurfer
  http://www.speech.kth.se/wavesurfer/
- Speech Filing System
  http://www.phon.ucl.ac.uk/resource/sfs/
- Hidden Markov Model Toolkit (HTK)
  http://htk.eng.cam.ac.uk/index.shtml
- The Festival Speech Synthesis System
  http://www.cstr.ed.ac.uk/projects/festival/

Any Questions ?