

# 15780 - Speech Technologies

## fall semester, 2008/09

### Course Information

---

#### Professors:

Eduardo Lleida (office: D3.11, [lleida@unizar.es](mailto:lleida@unizar.es) )  
Tuesday, 17-19 / Wednesday, 11-13 / Friday, 11-13

Antonio Miguel (office: D2.06, [amiguel@unizar.es](mailto:amiguel@unizar.es))

Web page (<http://physionet.cps.unizar.es/~eduardo>)

Classroom A.14,  
Monday 16-18  
Tuesday 16-17

### Course Agenda

### Hours

	Hours
<b>Part I: Fundamental Theory ..... Eduardo Lleida M-T 11-12</b>	<b>27</b>
0. Fundamentals of Acoustics	3
Sound and Noise	
Sound and Noise Level Measure	
Sound Propagation	
Harmonic Plane Waves	
Pipes and Cavities: Acoustic Circuits	
1. Speech production	4
1.1. Anatomy and Physiology of the Speech Organs	
1.2. Acoustic Phonetics	
1.3. Acoustic Theory of Speech Production	
2. Hearing and Speech Perception	4
2.1. Anatomy and Physiology of the Ear	
2.2. Sound Perception	
3. Speech Analysis and Representation	8
3.1. Short-Time Speech Analysis	
3.2. Time-Domain Parameters	
3.3. Short-Time Fourier Analysis	
3.4. Linear Predictive Coding Analysis	
3.5. Cepstral Analysis	
3.6. Speech enhancement: Additive and Convolutional noise	
3.7. Perceptually motivated representations	
4. Speech and Audio Coding	8
4.1. Speech Coders Attributes	
4.2. Frequency Domain Coders	
4.3. Analysis by Synthesis Coders	
<b>Part II: Speech Technologies ..... Antonio Miguel M 12-13</b>	<b>18</b>
5. Pattern Recognition	6
5.1. Bayes' Decision Theory	
5.2. Classifiers	
5.3. Feature Extraction and Selection	
6. Speech Recognition	12

- 6.1. Why is it so difficult?
- 6.2. A Basic Pattern Recognition Approach
- 6.3. Statistical Speech Recognition
- 6.4. Acoustic Modelling: Hidden Markov Models
- 6.5. Language Modelling
- 6.6. Basic Search Algorithms
- 6.7. Voice Portal Development: VoiceXML

### Part III: Laboratory Work

- 1. The speech signal, acoustic-phonetic features: Generation 4
- 2. Short time analysis, pitch and formant estimation 4
- 3. Voice portals: VoiceXML 2
- 4. Speech recognition engine: command & control application 4

### Bibliography

#### Acoustics:

L.E. Kinsler  
 "Fundamentals of acoustics"  
 3 edition, De. J. Wiley & Sons, 1992

G. Ballou  
 "Handbook for Sound Engineers, the New Audio  
 Cyclopedia"  
 2 edition, Howard W. Sams & Company, 1987

#### Advanced books:

Frederick Jelinek  
 "Statistical Methods for Speech Recognition"  
 MIT Press, 1997

Steve Young, Gerrit Bloothoof  
 "Corpus-Based Methods in Language and Speech  
 Processing"  
 Kluwer Academic Publishers, 1997

#### Speech:

Xuedong Huang, Alex Acero, Hsiao-Wuen Hon  
 "Spoken Language Processing"  
 Prentice Hall PTR, 2001

D. O'Saughnessy  
 "Speech Communication: Human and Machine"  
 IEEE PRESS, 2000

J.R. Deller, J.G. Proakis, J.H.L. Hansen  
 "Discrete-Time Processing of Speech Signals"  
 MacMillan Publishing Co., 1993

L. Rabiner, B.H. Juang  
 "Fundamentals of Speech Recognition"  
 Prentice Hall, Signal Processing Series, 1993

L.R. Rabiner, R.W. Schafer  
 "Digital Processing of Speech Signals"  
 Prentice Hall, Signal Processing Series, 1978

#### On the web:

<http://svr-www.eng.cam.ac.uk/comp.speech>

#### about VoiceXML

<http://www.w3.org/Voice/>

### Evaluation

Lab Work ..... 40 % evaluated over

    homework ..... 25 %

    labwork ..... 25 %

    written examination ..... 50 %

Theory .....60 % written examination

**Final Mark: Theory + Lab Work (if Theory mark > 4 over 10)**