BE 514. **Speech Signal Processing**
Oded Ghitza, Biomedical Engineering

The goal of this course is to provide the basic concepts and theories of speech production, speech perception and speech signal processing, and their applications to contemporary speech technology. The course is organized in a manner that builds a strong foundation of basics first, and then concentrates on a range of signal processing methods for representing and processing the speech signal.

Prerequisites:
General: A familiarity with signals and systems, including continuous-time and discrete-time frequency analysis, sampling and filtering theory. A basic familiarity with probability, including Bayes' theory.
Courses:
ENG BE 401 or equivalent (e.g., ENG EC 401)
ENG BE 200 or equivalent (e.g., ENG EC 481, or ENG ME 308)

Course Schedule:
Lecture 4hrs/week

Textbook:
Theory and Applications of Digital Speech Processing (Rabiner and Schafer, 2011)

References:
Speech and Audio Signal Processing (Gold and Morgan, 2000)
Acoustics of American English Speech (Olive, Greenwood and Coleman, 1993)

Course Topics:

Week	*Class*
01	Introduction to Digital Speech Processing
02	Fundamentals of Digital Signal Processing
03	Fundamentals of Human Speech Production
04	Acoustic Properties of American English Speech
05	Fundamentals of Speech Perception
06	Computational Models of Speech Perception
07	Time-Domain Methods for Speech Processing
08	Frequency-Domain Representations
09	The Cepstrum and Homomorphic Speech processing; Linear Predictive Analysis of Speech
10	Algorithms for Estimating Speech Parameters
11	Digital Coding of Speech and Audio
12	Text-to-Speech Synthesis Methods
13	Automatic Speech Recognition and Natural Language Understanding
14	Finale

Topics will be selected out of the following list (page numbers: lead textbook):

WEEK 1 Introduction to Digital Speech Processing 1
1.1 The Speech Signal 3
1.2 The Speech Stack 8
1.3 Applications of Digital Speech Processing 10

WEEK 2 Fundamentals of Digital Signal Processing 18
2.2 Review of Discrete-Time Signals and Systems 18
2.3 Review of Transform Representation of Signals and Systems (DFT, STFT) 22
2.4 Fundamentals of Digital Filters 33
2.5 Review of Sampling Theory 44

WEEK 3 Fundamentals of Human Speech Production 67
(Sound Propagation in the Human Vocal Tract 170)
3.2 The Process of Speech Production 68
3.3 Short-Time Fourier Representation of Speech 81
5.1 The Acoustic Theory of Speech Production 170
5.2 Lossless Tube Models of the Vocal Tract 200
5.3 Digital Models for Sampled Speech Signals 219

WEEK 4 Acoustic Properties of American English Speech
--- Reference book by Olive et al.
3.4 Acoustic Phonetics 86
3.5 Distinctive Features of the Phonemes of American English 108

WEEK 5 Fundamentals of Speech Perception 124
4.2 The Speech Chain 125
4.3 Anatomy and Function of the Ear 127
4.4 The Perception of Sound; Masking; Pitch 133
4.7 Measurement of Speech Quality and Intelligibility 162

WEEK 6 Computational Models of Speech Perception
(Objective assessment of Speech Quality and Intelligibility)
--- EIH (Ensemble Interval Histogram Representation)
--- MOCB (Medial OlivoCochlear Bundle); Closed-loop Cochlear Processing
--- Modulation Spectrum
--- Objective assessment of Speech Quality and Intelligibility (AI, STI)

WEEK 7 Time-Domain Methods for Speech Processing 239
6.2 Introduction to Short-Time Analysis of Speech 242
6.3 Short-Time Energy and Short-Time Magnitude 248
6.4 Short-Time Zero-Crossing Rate 257
6.5 The Short-Time Autocorrelation Function 265
6.6 The Modified Short-Time Autocorrelation Function 273
6.7 The Short-Time Average Magnitude Difference Function 275

WEEK 8 Frequency-Domain Representations 287
7.2 Discrete-Time Fourier Analysis 289
7.3 Short-Time Fourier Analysis 292
7.4 Spectrographic Displays 312
7.5 Overlap Addition (OLA) Method of Synthesis 319
7.6 Filter Bank Summation (FBS) Method of Synthesis 331
7.7 Time-Decimated Filter Banks 340
7.8 Two-Channel Filter Banks 348
7.9 Implementation of the FBS Method Using the FFT 358
7.10 OLA Revisited 365
7.11 Modifications of the STFT 367

WEEK 9 Selected topics from:
The Cepstrum and Homomorphic Speech Processing 399
8.2 Homomorphic Systems for Convolution 401
8.3 Homomorphic Analysis of the Speech Model 417
8.4 Computing the Short-Time Cepstrum and Complex Cepstrum of Speech 429
8.5 Homomorphic Filtering of Natural Speech 440
8.6 Cepstrum Analysis of All-Pole Models 456
8.7 Cepstrum Distance Measures 459

Linear Predictive Analysis of Speech Signals 473