

Short Title:	Speech Processing Technology APPROVED
Full Title:	Speech Processing Technology
Module Code:	MHLT H6017
ECTS credits:	10
NFQ Level:	9
Module Delivered in	1 programme(s)
Module Contributor:	Barry Kirkpatrick
Module Description:	In this module the learner will gain knowledge and experience of working with speech signal processing technology and learn about about the key signal processing techniques underpinning modern speech technology applications and how they impact the performance of the system.
Learning Outcomes:	
<i>On successful completion of this module the learner will be able to</i>	
<ol style="list-style-type: none"> 1. Identify, describe and apply signal processing techniques to speech processing applications and speech feature extraction. 2. Compare and select suitable features set to be employed for a specified speech application. 3. Describe and evaluate the components of automatic speech recognition systems. 4. Describe and evaluate the components of text-to-speech synthesis systems. 5. Describe and critique emerging applications of speech technology. 	

Module Content & Assessment

Indicative Content
<p>Fundamental Theory Introduction to speech - Sound and human speech. Phonetics and Phonology. Words, syllables and phonemes. Syntax and Semantics. Speech Technology Overview - Technology and applications utilising speech processing. Overview of Digital Signal Processing. Digital signals, systems and sampling. Time and frequency domain representations. Digital filters, the fast Fourier transform, windowing and filterbanks. Auditory system & speech perception - Anatomy of the auditory system. Signal processing models of the auditory system. Psychoacoustics and auditory perception of speech. Speech production - Anatomy of the speech production system. Models of the speech production system. Speech Signal representations - Short time Fourier analysis and feature extraction. Acoustic model of speech production. Linear predictive coding. Perceptually motivated representations. Formant frequencies and pitch. Measurement of speech quality and intelligibility.</p>
<p>Speech Recognition Hidden Markov Models - The Markov chain and hidden Markov models. Continuous and semi continuous HMMs. Practical issues in using HMMs and their limitations. Gaussian Mixture Models. Acoustic Modelling - Variability in the speech signal. How to measure speech recognition errors. Signal processing and feature extraction. Phonetic modelling in speech recognition. Acoustic modelling—scoring acoustic features. Robustness and adaptive techniques—minimizing mismatches. Confidence measures: measuring the reliability.</p>
<p>Text-to-speech synthesis Text and Phonetic Analysis - Modules and data Flow. The lexicon of the synthesiser. Document structured detection. Text normalization. Linguistic analysis. Homograph disambiguation. Morphological analysis. Letter-to-sound conversion. Evaluation. Case study: Festival speech synthesis system. Prosody - Perception of prosody. Prosody generation schematic, speaking style, symbolic prosody and duration assignment. Pitch generation and evaluation. Speech Synthesis - Attributes of speech synthesis. Formant synthesis. Concatenative synthesis and unit selection synthesis. Prosodic modification of speech. Source-filter models for prosody modification. Evaluation of TTS Systems. Statistical parametric speech synthesis. Developing a speech corpus for synthesis. Expressive speech synthesis.</p>
<p>Review of Research Oriented and Emerging Applications A review of emerging fields of study and applications in speech processing for example biometrics and biomedical applications of speech processing.</p>

Indicative Assessment Breakdown	%
Course Work Assessment %	100.00%

Course Work Assessment %				
Assessment Type	Assessment Description	Outcome addressed	% of total	Assessment Date
Project	Fundamentals of Speech Processing and Speech Feature Extraction - In this project the learner will investigate, implement and compare a number of speech feature sets and representations. They will investigate a new or emerging speech based application, conduct a literature review and investigate suitable speech representations and algorithms for the chosen application.	1,2,5	30.00	n/a
Project	Speech Recognition - This project will involve a literature review, system specification and design, implementation and testing of a basic speech recognition system.	1,2,3,5	35.00	n/a
Project	Speech Synthesis - This project will involve a literature review, system design, recording of a speech corpus and implementation and evaluation of a basic speech synthesis system.	1,2,4,5	35.00	n/a

No Final Exam Assessment %

ITB reserves the right to alter the nature and timings of assessment

Indicative Module Workload & Resources

Indicative Workload: Full Time	
Frequency	Indicative Average Weekly Learner Workload
Every Week	2.00
Every Week	2.00
Every Week	3.00

Resources

Recommended Book Resources

Lawrence Rabiner, Ronald Schafer 2010, *Theory and Applications of Digital Speech Processing*, Prentice Hall [ISBN: 0136034284]

Xuedong Huang, Alex Acero, Hsiao-Wuen Hon 2001, *Spoken language processing*, Prentice Hall PTR Upper Saddle River, N.J. [ISBN: 0130226165]

by Thierry Dutoit, *An introduction to text-to-speech synthesis*, Dordrecht ; Kluwer Academic Publishers, c1997. [ISBN: 0792344987]

Lawrence Rabiner, Biing-Hwang Juang 1993, *Fundamentals of speech recognition*, PTR Prentice Hall Englewood Cliffs, N.J. [ISBN: 9780130151575]

This module does not have any article/paper resources

This module does not have any other resources

Module Delivered in

Programme Code	Programme	Semester	Delivery
BN_KMHLT_R	Master of Science in Computing in Multimodal Human Language Technology	2	Elective