

Academic Year: ( 2024 / 2025 )

Review date: 12-02-2025

Department assigned to the subject: Signal and Communications Theory Department

Coordinating teacher: PELAEZ MORENO, CARMEN

Type: Electives ECTS Credits : 3.0

Year : 1 Semester : 2

## OBJECTIVES

Speech Technologies for Health will provide the students with knowledge on:

- the speech production and perception mechanisms and the linguistic categories of the voice
- the state of the art and fundamentals of speech and audio coders, speech recognition, speech synthesis or text-to-speech, speaker recognition and dialogue systems
- their applications to health

They will also acquire abilities to:

- initiate research work in speech and audio coding, speech recognition, speech synthesis or text-to-speech, speaker recognition and natural language processing
- to apply the previous knowledge to pose research questions on their applications to health

## DESCRIPTION OF CONTENTS: PROGRAMME

Unit 0. Introduction to Speech Technologies

Unit 1. The Auditory System and Speech Perception.

Unit 2. The Speech Production System and Phonation. Speech and Audio Coding

Unit 3. Speech recognition:

- Feature extraction
- Acoustic & Language models
- End-to-end recognition and deep learning models

Unit 4. Speaker recognition and biometrics

Unit 5. Text-to-Speech synthesis

Unit 6. Applications to health

- Affective computing: Emotion and Sentiment analysis
- Pathologies detection

## LEARNING ACTIVITIES AND METHODOLOGY

The following learning activities and methodologies are employed: Combined master and lab classes (AF3) included guided lab assignments (AF4 and AF6), debates about key state of the art issues based on paper readings and audiovisual clips (AF6-AF7-AF8) and final project (AF6).

Teachers are available during 2 hours per week for office hours (AF5).

Methodologically, we will use class lectures (MD1), flipped classrooms (MD2) and problem based learning (with varying amount of supervision and scope, MD3) and gamification (MD4 and MD5).

Combined master and lab classes:

Master classes provide an overview of the main theoretical & mathematical concepts of the representation and processing of speech processing along with the analytic tools employed for their

processing.

In these classes, lab examples will be introduced as part of the theoretical expositions: all the formative sessions (lab availability provided) will take place in the lab to imbricate practical examples within the explanations to add dynamism to the class. This is also beneficial to solve different background issues.

Moreover, every unit will begin with a debate of its technological implications. For this purpose, flipped classroom methodologies will be employed. In particular, students will be provided with some selected videos in advance to motivate the debate together with a list of questions (sometimes controversial) that the instructor will not answer categorically to encourage discussions. In this way, we expect to awake the curiosity of the student on the materials that will be subsequently explained.

## GAMIFICACIÓN

We will illustrate the process of scientific publishing with a role game. The students will join teams with different roles: editor, reviewer and author. Guidelines with requisites and forms to carry out their roles will be provided together with deadline dates (for reviewing, revising and editorial decisions). A paper will be assigned to each team and the outcomes of the process will be publicly debated. In this way, we promote team working, critical and active reading or research papers.

## FINAL PROJECT

Students will work on a project in which they will program a complete modular system of one of the tools explained in class. The students will be provided with some guidelines and some preparatory sessions by using problema-based learning.

## ASSESSMENT SYSTEM

<b>% end-of-term-examination:</b>	0
<b>% of continuous assessment (assignments, laboratory, practicals...):</b>	100

First call:

- Research paper presentation and discussion (30%, SE1-SE2)
- Lab assignment (30%, SE3)
- Class participation, wooclap tests (10%, SE1)
- Final project (30%, SE2)

Second call:

- Research paper presentation and discussion (30%, SE1-SE2)
- Lab assignment (30%, SE3)
- Final project (40%, SE2)

## BASIC BIBLIOGRAPHY

- Ben Gold (Author), Nelson Morgan (Author), Dan Ellis (Author) *Speech and Audio Signal Processing: Processing and Perception of Speech and Music*, Wiley, 2011
- Dan Jurafsky and James H. Martin *Speech and Language Processing (3rd ed.)*, Prentice Hall, 2018
- Yu, Dong, Deng, Li *Automatic Speech Recognition*, Springer, 2015

## ADDITIONAL BIBLIOGRAPHY

- Amy Neustein (Editor), Jenay M. Beer (Contributor), Conrad Bzura (Contributor) et al *Speech and Automata in Health Care*, De Gruyter, 2014
- Amy Neustein (Editor), Hemant A. Patil (Editor) *Acoustic Analysis of Pathologies: From Infancy to Young Adulthood*, De Gruyter, 2020
- Deborah Dahl (Author), Katharine Beals (Author), Marcia Linebarger (Author), Ruth Fink (Author) *Speech and Language Technology for Language Disorders*, De Gruyter, 2015

- Rupayan Chakraborty, Meghna Pandharipande, Sunil Kumar Kopparapu Analyzing Emotion in Spontaneous Speech, Springer, 2018