

# Guia docent

## 320123 - TP - Tecnologia de la Parla

Última modificació: 22/04/2021

**Unitat responsable:** Escola Superior d'Enginyeries Industrial, Aeroespacial i Audiovisual de Terrassa

**Unitat que imparteix:** 739 - TSC - Departament de Teoria del Senyal i Comunicacions.

**Titulació:** GRAU EN ENGINYERIA DE SISTEMES AUDIOVISUALS (Pla 2009). (Assignatura optativa).

**Curs:** 2021

**Crèdits ECTS:** 6.0

**Idiomes:** Català, Castellà

### PROFESSORAT

**Professorat responsable:** Ignasi Esquerra

**Altres:** Albino Nogueiras

### CAPACITATS PRÈVIES

És convenient haver cursat les assignatures de "Processament digital d'àudio" i "Algorísmica i programació audiovisual", i tenir un cert coneixement de programació en entorns Linux.

### COMPETÈNCIES DE LA TITULACIÓ A LES QUALS CONTRIBUEIX L'ASSIGNATURA

**Específiques:**

CE26. AUD: Coneixements i capacitats per aprofundir en tecnologies específiques de l'àmbit. (OBSOLETA)

### METODOLOGIES DOCENTS

Aprentatge dirigit :

a) Sessions en el laboratori. Presentació dels continguts teòrics i realització de demostracions. Exercicis pràctics. Exercicis d'avaluació.

Aprentatge autònom:

b) Estudi individual, resolució d'exercicis i preparació dels treballs.

### OBJECTIUS D'APRENTATGE DE L'ASSIGNATURA

Conèixer les tecnologies utilitzades en els sistemes d'interacció amb els ordinadors a través de la veu. Els estudiants en acabar el curs coneixeran diverses tècniques de tractament dels senyals de veu i seran capaços d'aplicar-los en sistemes de reconeixement de veu i de síntesi de veu.

### HORES TOTALS DE DEDICACIÓ DE L'ESTUDIANTAT

Tipus	Hores	Percentatge
Hores grup gran	30,0	20.00
Hores grup petit	30,0	20.00
Hores aprenentatge autònom	90,0	60.00

**Dedicació total:** 150 h

## CONTINGUTS

### 1. INTERACCIÓ ORAL AMB ELS ORDINADORS

**Descripció:**

Tecnologies de la parla i del llenguatge. Perspectiva històrica de la interacció persona-ordinador. Aplicacions i línies de recerca actuals.

**Dedicació:** 10h

Grup gran/Teoria: 2h

Grup petit/Laboratori: 2h

Aprenentatge autònom: 6h

### 2. BASES DE DADES

**Descripció:**

Bases de dades de text i senyals de veu. Fonètica i lingüística. Producció i percepció de la parla. Representació i anàlisi dels senyals de veu.

**Dedicació:** 30h

Grup gran/Teoria: 8h

Grup petit/Laboratori: 8h

Aprenentatge autònom: 14h

### 3. SÍNTESI DE VEU

**Descripció:**

Anàlisi del text. Transcripció fonètica. Prosòdia. Síntesi per selecció d'unitats. Síntesi per models probabilístics.

**Dedicació:** 46h

Grup gran/Teoria: 8h

Grup petit/Laboratori: 8h

Aprenentatge autònom: 30h

### 4. REONEIXEMENT DE VEU

**Descripció:**

Parametrització. Models acústics. Models ocults de Markov. Models de llenguatge.

**Dedicació:** 64h

Grup gran/Teoria: 12h

Grup petit/Laboratori: 12h

Aprenentatge autònom: 40h

## SISTEMA DE QUALIFICACIÓ

La qualificació final és la suma ponderada de les qualificacions parcials dels temes. Cada tema és avaluat amb diversos treballs de laboratori o controls. Cap dels elements de qualificació té un pes superior a 25% sobre el total.

Tema 1 (10%), Tema 2 (20%), Tema 3 (35%), Tema 4 (35%)



## BIBLIOGRAFIA

---

### Complementària:

- O'Shaughnessy, D. Speech communications: human and machine. 2nd ed. New York: IEEE Press, 2000. ISBN 978-0780334496.
- Huang, X.; Acero, A.; Hon, H-W. Spoken language processing: a guide to theory, algorithm and system development. Upper Saddle River: Prentice Hall, 2001. ISBN 0130226165.