



# Procesado de la Voz y el Habla

## Guía de Aprendizaje – Información al estudiante

### 1. Datos Descriptivos

<b>Titulación</b>	Máster Universitario en Ingeniería Informática
<b>Módulo</b>	Dirección y Gestión
<b>Materia</b>	Dominios de Aplicación de las TI a los Servicios
<b>Asignatura</b>	Procesado de la Voz y el Habla
<b>Carácter</b>	Optativo
<b>Créditos ECTS</b>	2
<b>Departamento responsable</b>	Arquitectura y Tecnología de Sistemas Informáticos
<b>Especialidad</b>	

<b>Curso académico</b>	2011-2012
<b>Semestre en que se imparte</b>	1er o 2º semestre del curso
<b>Idioma en él que se imparte</b>	Español o Inglés (si lo solicitan al menos 5 alumnos)
<b>Página Web</b>	<a href="http://www.datsi.fi.upm.es/">http://www.datsi.fi.upm.es/</a>



**POLITÉCNICA**



UNIVERSIDAD POLITÉCNICA DE MADRID  
**FACULTAD DE INFORMÁTICA**  
Campus de Montegancedo  
Boadilla del Monte. 28660 Madrid

## 2. Profesorado

NOMBRE Y APELLIDO	DESPACHO	Correo electrónico
Agustín Álvarez Marquina (Coord)	4211	agustin@junipera.datsi.fi.upm.es
Rafael Martínez Olalla	4208	rmolalla@junipera.datsi.fi.upm.es
Víctor Nieto Lluís	4208	victor@fi.upm.es
Pedro Gómez Vilda	4209	pedro@pino.datsi.fi.upm.es

## 3. Conocimientos previos requeridos para poder seguir con normalidad la asignatura

<b>Asignaturas superadas</b>	<ul style="list-style-type: none"><li>•</li></ul>
<b>Otros resultados de aprendizaje necesarios</b>	<ul style="list-style-type: none"><li>•</li></ul>



## 4. Objetivos de Aprendizaje

<b>COMPETENCIAS ESPECÍFICAS ASIGNADAS A LA ASIGNATURA Y SU NIVEL DE ADQUISICIÓN</b>		
<b>Código</b>	<b>Competencia</b>	<b>Nivel</b>
CE4	Capacidad para modelar, diseñar, definir la arquitectura, implantar, gestionar, operar, administrar y mantener aplicaciones, redes, sistemas, servicios y contenidos informáticos.	A
CE13	Capacidad para utilizar y desarrollar metodologías, métodos, técnicas, programas de uso específico, normas y estándares de computación gráfica.	A
CE14	Capacidad para conceptualizar, diseñar, desarrollar y evaluar la interacción persona–ordenador de productos, sistemas y servicios informáticos.	A
CE15	Capacidad para la creación y explotación de entornos virtuales, y para la creación y distribución de contenidos multimedia.	A
CE16	Habilidad para hacer conexiones entre los deseos y necesidades del consumidor o cliente y lo que la tecnología puede ofrecer.	A
CE17	Capacidad para decidir entre adquirir, desarrollar o aplicar tecnología a lo largo de la amplia gama de categorías de procesos, productos y servicios de una empresa o institución.	A
CE18	Capacidad para comprender el mercado, sus hábitos y necesidades de productos o servicios tecnológicos.	A
CE19	Capacidad para desarrollar e implantar una solución informática en un entorno empresarial.	A

Nivel de competencia: conocimiento (C), comprensión (P), aplicación (A) y análisis y síntesis (S),



<b>RESULTADOS DE APRENDIZAJE DE LA ASIGNATURA</b>			
<b>Código</b>	<b>Resultado de aprendizaje</b>	<b>Competencias asociadas</b>	<b>Nivel de adquisición</b>
RA1	Manejo de ficheros en forma de onda	CE13	
RA2	Capacidad para determinar las características espectrales y temporales de una señal de voz.	CE13, CE14	
RA3	Diferenciación de las características de la señal de voz debidas al pulso glótico de aquellas causadas por el tracto vocal.	CE13, CE14	
RA4	Interpretación de las representaciones de armónicos y formantes.	CE13, CE14	
RA5	Capacidad para registrar y manipular bases de datos de voz.	CE4, CE13	
RA6	Extracción de características de la señal de voz.	CE13	
RA7	Manejo de sistemas de reconocimiento de voz	CE 4, CE13, CE14, CE15, CE16, CE17, CE18, CE19	
RA8	Manejo a nivel usuario de sistemas de edición e indexado de audio, video y texto	CE 4, CE13, CE14, CE15, CE16, CE17, CE18, CE19	
RA9	Creación de interfaces de usuario para aplicaciones de voz.	CE 4, CE14, CE15, CE16, CE18, CE19	



## 5. Sistema de evaluación de la asignatura

INDICADORES DE LOGRO		
Ref	Indicador	Relacionado con RA
I1	Representaciones temporales y frecuenciales de la señal de voz.	RA1, RA2
I2	Representación de espectrogramas. Creación de una interfaz sencilla.	RA1, RA2, RA3, RA4, RA9
I3	Creación de una base de datos de voz.	RA1, RA5
I4	Segmentación de señal. Detección de actividad de voz.	RA1, RA6
I5	Cálculo del tono fundamental.	RA1, RA6
I6	Realización de un extractor de característica sonora sordo.	RA1, RA6
I7	Realización de un extractor de plantillas de características de la señal de voz.	RA1, RA6
I8	Realización de un sintetizador de vocales	RA1, RA7
I9	Creación de un sistema sencillo de reconocimiento a nivel fonético.	RA1, RA7
I10	Creación de un detector de hitos fonéticos	RA1, RA8
I11	Creación de una interfaz de indexado de audio.	RA1, RA9



<b>EVALUACION SUMATIVA</b>			
<b>Breve descripción de las actividades evaluables</b>	<b>Momento</b>	<b>Lugar</b>	<b>Peso en la calif.</b>
AE1. Representación de señales en el tiempo y en la frecuencia	Semana 2	Laboratorio	10%
AE2. Representación de espectrogramas	Semana 3	Laboratorio	10%
AE3. Registro de una base de datos de voz	Semana 5	Laboratorio	10%
AE4. Realización de un módulo de detección de actividad de voz y extracción de tono fundamental	Semana 7	Laboratorio	15%
AE5. Realización de un módulo de extracción de plantillas MFCC	Semana 10	Laboratorio	15%
AE6. Integración de los módulos de detección de actividad de voz y extracción de plantillas MFCC con un reconocedor a nivel fonético.	Semana 11	Laboratorio	25%
AE7. Integración de los módulos de detección de actividad de voz y extracción de plantillas MFCC con un sistema de etiquetado fonético.	Semana 14	Laboratorio	15%
			<b>Total: 100%</b>



## CRITERIOS DE CALIFICACIÓN

A lo largo del desarrollo de la asignatura, el estudiante deberá ir completando una serie de módulos señalados en la tabla de la evaluación sumativa. El objetivo final es la realización de diferentes módulos autónomos que puedan funcionar como API's en el desarrollo e implantación de sistemas más complejos. Para superar la asignatura es condición necesaria el haber ido completando de forma satisfactoria cada uno de los módulos que comprenden el sistema completo y haberlos integrado en una interfaz de usuario.

Para la calificación final, el alumno deberá mostrar y defender el sistema desarrollado. Los principales aspectos de su trabajo, así como los resultados obtenidos deberán reflejarse en un documento o memoria explicativa. Para la exposición y defensa del sistema, el alumno se ayudará de una presentación en PowerPoint®. En la calificación final se tendrá en cuenta:

- Realización de todos los módulos de que consta el sistema.
- Explicación de los fundamentos del sistema realizado.
- Presentación de los principales resultados.
- Aportación de soluciones particulares aparte de las explicadas en la asignatura.



## 6. Contenidos y Actividades de Aprendizaje

<b>CONTENIDOS ESPECÍFICOS</b>		
<b>Bloque / Tema / Capítulo</b>	<b>Apartado</b>	<b>Indicadores Relacionados</b>
<b>Tema 1: Introducción</b>	1.1 Visión general de la asignatura	I1 a I11
	1.2 Sonido y resonancia	I1, I2
	1.3 Tipos de representación	I1, I2
<b>Tema 2: Fundamentos de la ciencia del habla</b>	2.1 Anatomía y fisiología del sistema de producción de voz	I1, I8
	2.2 Mecanismo de audición	I1, I9
	2.3 Caracterización espectral y temporal del sonido	I1, I2
<b>Tema 3: Fundamentos de procesado de señal</b>	3.1 Manipulación de ficheros en forma de onda	I1 a I11
	3.2 Conceptos básicos de procesado de señal	I2, I4, I5, I6, I7, I8, I9, I10
	3.3 Análisis en el dominio de la frecuencia	I2
	3.4 Fundamentos de filtrado digital	I4, I8, I9, I10
<b>Tema 4: Registro de sonido</b>	4.1 Formatos de ficheros de sonido	I3
	4.2 Registro de bases de datos de voz	I3
<b>Tema 5: Detección de actividad de voz</b>	5.1 Procesado en bloques de la señal de voz	I4
	5.2 Pre-acondicionamiento de la señal y cálculo de energía	I4
<b>Tema 6: Extracción de características de la señal de voz</b>	6.1 Detección de tono fundamental	I4, I5
	6.2 Detección de característica sonoro/sordo	I4, I6
	6.3 Extracción de plantillas MFCC	I4, I7



<b>Tema 7: Síntesis de voz</b>	7.1 Descripción general de los sistemas de conversión texto voz.	18
	7.2 Síntesis basada en el modelo fuente filtro	18
<b>Tema 8: Reconocimiento de habla</b>	8.1 Descripción general de un sistema de reconocimiento de habla.	19
	8.2 Modelos ocultos de Markov	19
	8.3 Modelado del lenguaje	19
	8.4 Fuentes de variabilidad	19
<b>Tema 9: Aplicaciones del Procesado de la Voz</b>	9.1 Conversión Voz-Texto y Texto-Voz	110
	9.2 Voz sobre IP	110
	9.3 Indexado de Video y Audio	110

## 7. Breve descripción de las modalidades organizativas utilizadas y de los métodos de enseñanza empleados

Tabla 7. Modalidades organizativas de la enseñanza

MODALIDADES ORGANIZATIVAS DE LA ENSEÑANZA		
Escenario	Modalidad	Finalidad
	Clases Teóricas	<i>Hablar a los estudiantes</i>
	Seminarios-Talleres	<i>Construir conocimiento a través de la interacción y la actividad de los estudiantes</i>
	Clases Prácticas	<i>Mostrar a los estudiantes cómo deben actuar</i>
	Prácticas Externas	<i>Completar la formación de los alumnos en un contexto profesional</i>
	Tutorías	<i>Atención personalizada a los estudiantes</i>
	Trabajo en grupo	<i>Hacer que los estudiantes aprendan entre ellos</i>
	Trabajo autónomo	<i>Desarrollar la capacidad de autoaprendizaje</i>

Tabla 5. Métodos de enseñanza

MÉTODOS DE ENSEÑANZA		
	Método	Finalidad
	Método Expositivo/Lección Magistral	Transmitir conocimientos y activar procesos cognitivos en el estudiante
	Estudio de Casos	Adquisición de aprendizajes mediante el análisis de casos reales o simulados
	Resolución de Ejercicios y Problemas	Ejercitar, ensayar y poner en práctica los conocimientos previos
	Aprendizaje Basado en Problemas (ABP)	Desarrollar aprendizajes activos a través de la resolución de problemas
	Aprendizaje orientado a Proyectos	Realización de un proyecto para la resolución de un problema, aplicando habilidades y conocimientos adquiridos
	Aprendizaje Cooperativo	Desarrollar aprendizajes activos y significativos de forma cooperativa
	Contrato de Aprendizaje	Desarrollar el aprendizaje autónomo

Se conoce como método expositivo "la presentación de un tema lógicamente estructurado con la finalidad de facilitar información organizada siguiendo criterios adecuados a la finalidad pretendida". Esta metodología -también conocida como lección (lecture)- se centra fundamentalmente en la exposición verbal por parte del profesor de los contenidos sobre la materia objeto de estudio. El término "lección magistral" se suele utilizar para denominar un tipo específico de lección impartida por un profesor en ocasiones especiales.

Análisis intensivo y completo de un hecho, problema o suceso real con la finalidad de conocerlo, interpretarlo, resolverlo, generar hipótesis, contrastar datos, reflexionar, completar conocimientos, diagnosticarlo y, en ocasiones, entrenarse en los posibles procedimientos alternativos de solución.

Situaciones en las que se solicita a los estudiantes que desarrollen las soluciones adecuadas o correctas mediante la ejercitación de rutinas, la aplicación de fórmulas o algoritmos, la aplicación de procedimientos de transformación de la información disponible y la interpretación de los resultados. Se suele utilizar como complemento de la lección magistral.

Método de enseñanza-aprendizaje cuyo punto de partida es un problema que, diseñado por el profesor, el estudiante ha de resolver para desarrollar determinadas competencias previamente definidas.

Método de enseñanza-aprendizaje en el que los estudiantes llevan a cabo la realización de un proyecto en un tiempo determinado para resolver un problema o abordar una tarea mediante la planificación, diseño y realización de una serie de actividades, y todo ello a partir del desarrollo y aplicación de aprendizajes adquiridos y del uso efectivo de recursos.

Enfoque interactivo de organización del trabajo en el aula en el cual los alumnos son responsables de su aprendizaje y del de sus compañeros en una estrategia de corresponsabilidad para alcanzar metas e incentivos grupales. Es tanto un método, a utilizar entre otros, como un enfoque global de la enseñanza, una filosofía.

Un acuerdo establecido entre el profesor y el estudiante para la consecución de unos aprendizajes a través de una propuesta de trabajo autónomo, con una supervisión por parte del profesor y durante un periodo determinado. En el contrato de aprendizaje es básico un acuerdo formalizado, una relación de contraprestación recíproca, una implicación personal y un marco temporal de ejecución.



<b>BREVE DESCRIPCIÓN DE LAS MODALIDADES ORGANIZATIVAS UTILIZADAS Y METODOS DE ENSEÑANZA EMPLEADOS</b>	
<b>CLASES DE TEORIA</b>	Cada uno de los 9 temas de que consta la asignatura se impartirán a modo de seminario en un laboratorio equipado con ordenadores con micrófono y altavoces. Se desarrollarán clases introductorias en que se expondrán los conceptos básicos de un modo teórico, aunque apoyados en ejemplos prácticos sobre la plataforma de ordenador, para pasar progresivamente a aumentar el componente de trabajo práctico en modalidad “learning by doing” reduciendo paulatinamente la componente teórica explicativa e incorporando módulos de autoaprendizaje.
<b>CLASES DE PROBLEMAS</b>	...
<b>PRÁCTICAS</b>	A partir del tercer tema, en cada una de las sesiones se comenzará a desarrollar, con la ayuda del profesor, una de las actividades evaluables (AE) que forman parte del proyecto final de la asignatura, dejando como tarea para el alumno el completar esa actividad.
<b>TRABAJOS AUTONOMOS</b>	El alumno deberá completar y ampliar las actividades evaluables iniciadas junto al profesor en el laboratorio.
<b>TRABAJOS EN GRUPO</b>	...
<b>TUTORÍAS</b>	Durante todo el semestre se realizará un seguimiento de la evolución del trabajo del alumno mediante tutorías personalizadas. Estas tutorías se intensificarán a partir de la semana 11, en la que comenzará la integración de los diferentes bloques dentro del proyecto final. En las semanas 13 y 14, se revisará en el laboratorio la evolución de los proyectos de forma común (cumplimiento de los hitos AE9 y AE10). La semana 15, la revisión de esa evolución será personalizada.



## 8. Recursos didácticos

<b>RECURSOS DIDÁCTICOS</b>	
<b>BIBLIOGRAFÍA</b>	Deller, J. R., Proakis, J. G., Hansen, J. H. L., <i>Discrete-Time Processing of Speech Signals</i> . Ed. Macmillan, 1993.
	S. Furui. <i>Digital Speech Processing Synthesis and Recognition</i> . CRC Press, 2001
	G. M. Ballou, <i>Handbook for Sound Engineers</i> , Focal Press, 2003
<b>RECURSOS WEB</b>	Página web de la asignatura: <a href="http://www.datsi.fi.upm.es/">http://www.datsi.fi.upm.es/</a>
	Sitio Moodle de la asignatura ( <a href="http://">http://</a> )
<b>EQUIPAMIENTO</b>	Laboratorio
	Aula Informática con equipos Multimedia
	Sala de trabajo en grupo



### 9. Cronograma de trabajo de la asignatura

Semana	Actividades en Aula	Actividades en Laboratorio	Trabajo Individual	Trabajo en Grupo	Actividades de Evaluación	Otros
Semana 1 (3 horas)	•	• Tema 1 (2 horas)	• Estudio personal (2 horas)	•	•	•
Semana 2 (3 horas)	•	• Tema 2 (2 horas)	• Estudio personal (2 horas)	•	•	•
Semana 3 (3 horas)	•	• Tema 3 (2 horas)	• Completar AE1 (2 horas)	•	•	•
Semana 4 (3 horas)	•	• Tema 3 (2 horas)	• Completar AE2 (2 horas)	•	•	•
Semana 5 (3 horas)	•	• Tema 4 (2 horas)	• Completar AE3 (2 horas)	•	•	•
Semana 6 (3 horas)	•	• Tema 5 (2 horas)	• Completar AE4 (2 horas)	•	•	•
Semana 7 (3 horas)	•	• Tema 6 (2 horas)	• Iniciar AE5 (2 horas)	•	•	•
Semana 8 (3 horas)	•	• Tema 6 (2 horas)	• Completar AE5 (2 horas)	•	•	•
Semana 9 (3 horas)	•	• Tema 7 (2 horas)	• Iniciar AE6 (2 horas)	•	•	•
Semana 10 (3 horas)	•	• Tema 8 (2 horas)	• Completar AE6 (2 horas)	•	•	•
Semana 11 (3 horas)	•	• Tema 9 (2 horas)	• Completar AE9 (2 horas)	•	•	•
Semana 12 (3 horas)	•	• Tutoría común (2 horas)	• Iniciar AE7 (2 horas)	•	•	•
Semana 13 (3 horas)	•	• Tutoría común (2 horas)	• Desarrollar AE7 (4 horas)	•	•	•
Semana 14 (3 horas)	•	• Tutoría personalizada (1 hora)	• Desarrollar AE7 (4 horas)	•	•	•
Semana 15 (3 horas)	•	•	• Completar AE7 (4 horas)	•	•	•
Semana 16 (3 horas)	•	•	• Preparar presentación (6 horas)	•	• Presentación sistema completo (0.5 horas)	•

Nota: Para cada actividad se especifica la dedicación en horas que implica para el alumno.



**POLITÉCNICA**



UNIVERSIDAD POLITÉCNICA DE MADRID  
**FACULTAD DE INFORMÁTICA**  
Campus de Montegancedo  
Boadilla del Monte. 28660 Madrid