

Departamento de Computación, FCEyN, UBA

Procesamiento del Habla

Agustín Gravano

1er Cuatrimestre 2017

Procesamiento del Habla

Objetivo: Construir sistemas informáticos capaces de manipular efectivamente el lenguaje oral.

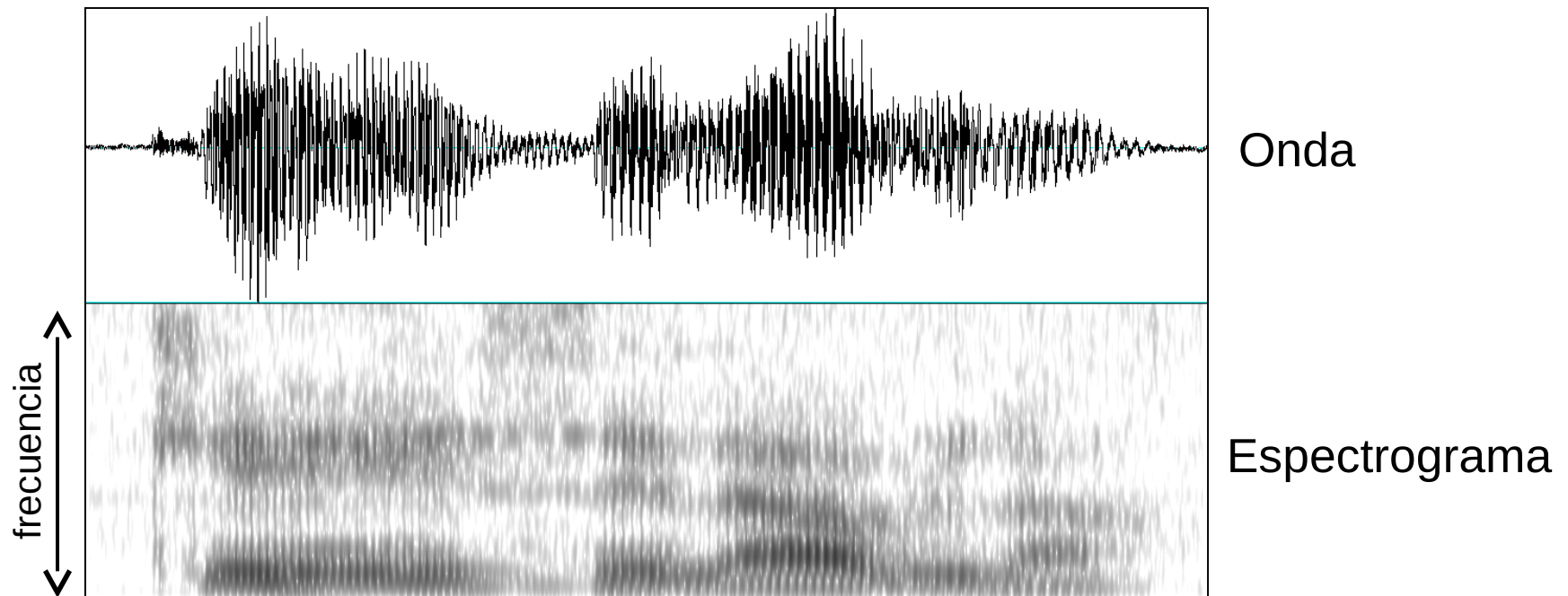
Disciplina fuertemente interdisciplinaria.

Objetivos de la Materia

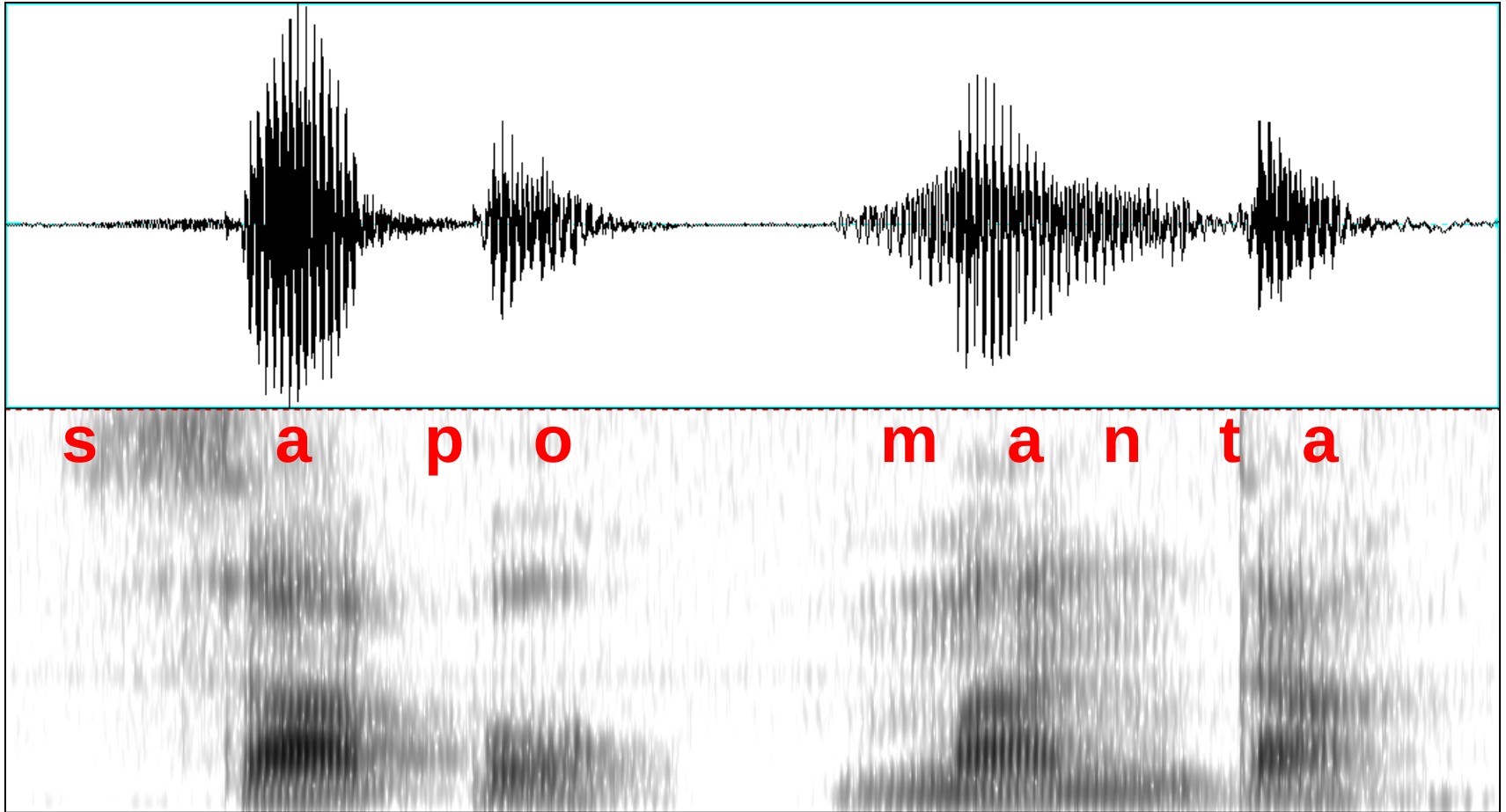
- Estudiar las bases necesarias para construir sistemas simples de procesamiento del habla.
 - Reconocimiento automático.
 - Síntesis del habla.
 - Detectores de características del hablante.
- Presentar (más superficialmente) temas avanzados de procesamiento del habla.

Reconocimiento Automático

- Habla → Secuencia de palabras
- ¿Qué es el habla?

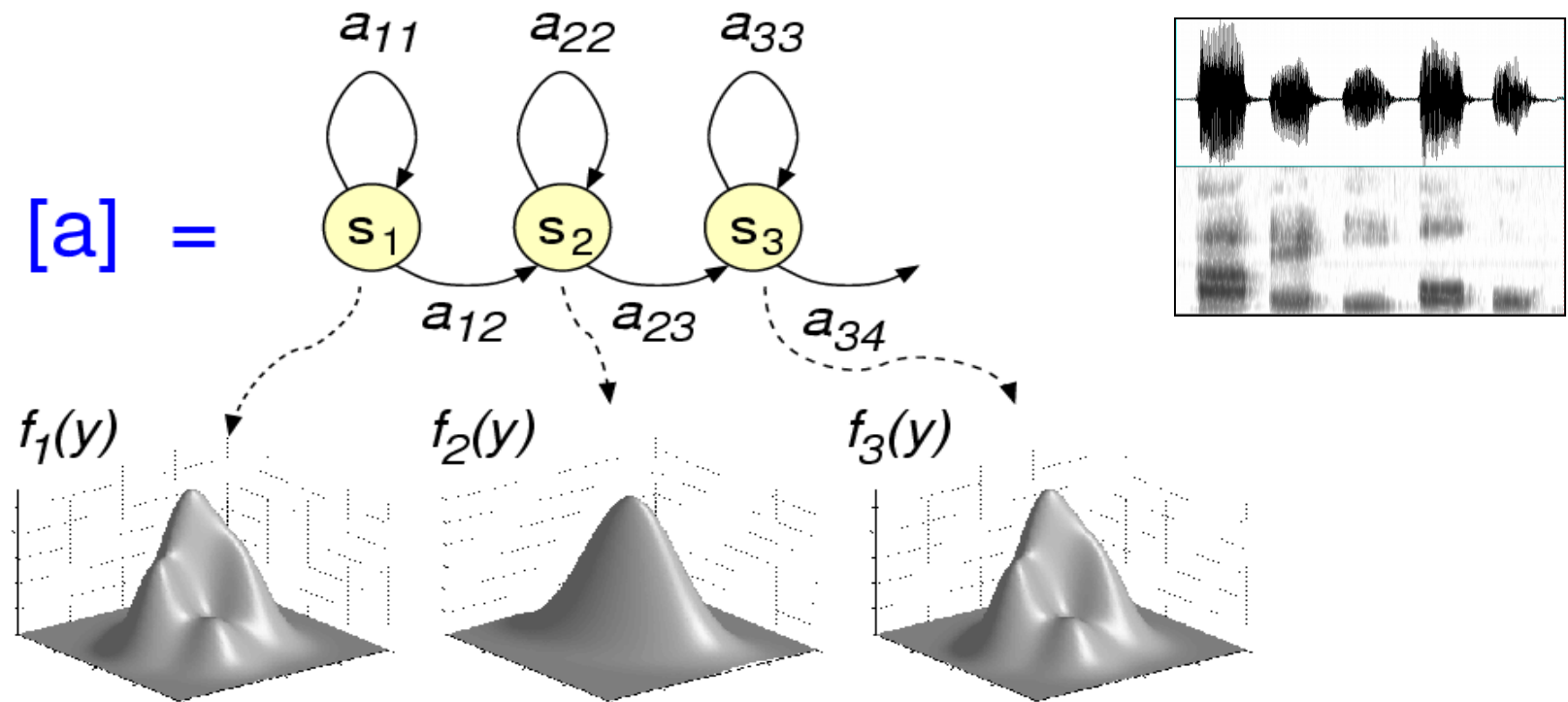


Reconocimiento Automático



Reconocimiento Automático

Hidden Markov Models



Ultimos años: Redes Neuronales Profundas

Reconocimiento Automático

- Pre-procesamiento de la entrada:
 - Filtro de ruidos.
 - Segmentación en hablantes.
- Post-procesamiento de la salida:
 - Puntuación, mayúsculas, formato de números.
 - *setenta y siete* → 77; *siglo dieciséis* → *siglo XVI*
- Comprensión del lenguaje natural:
 - “*Quiero un pasaje de Rosario a Mendoza para el veinte de noviembre a las seis de la tarde.*”
de:Rosario a:Mendoza fecha:20/11/09 hora:18:00

Síntesis del Habla

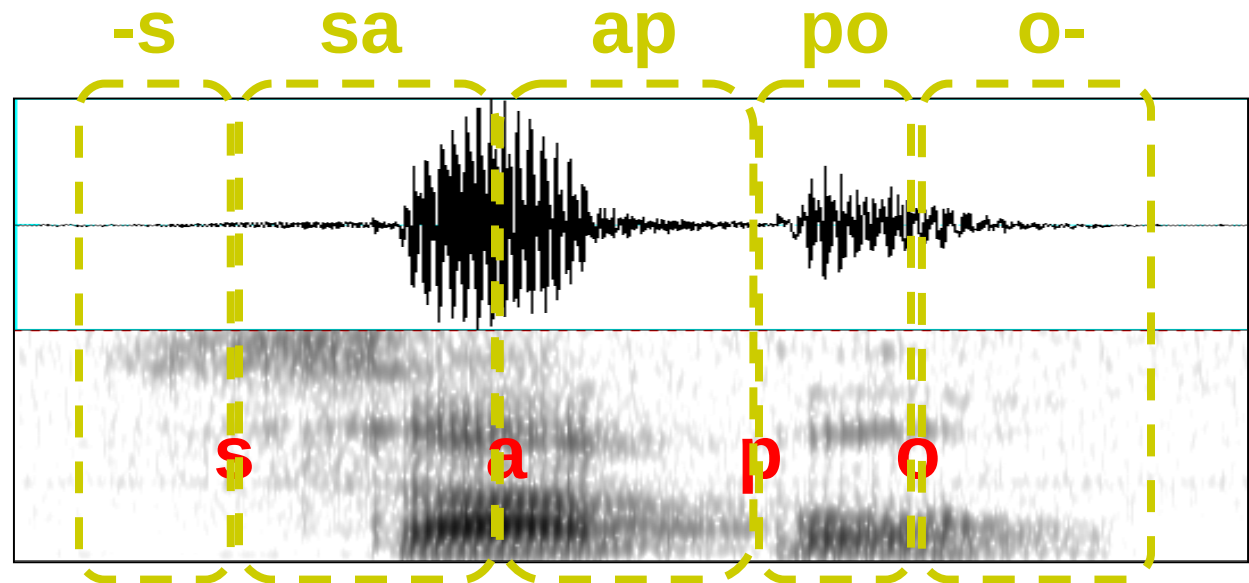
Secuencia de palabras → Habla

Hora oficial: 113 [113.wav]

<i>cero horas</i> <i>una hora</i> <i>dos horas</i> <i>...</i> <i>veintitrés horas</i>	<i>cero minutos</i> <i>un minuto</i> <i>dos minutos</i> <i>...</i> <i>cincuenta y nueve minutos</i>	<i>cero segundos</i> <i>diez segundos</i> <i>veinte segundos</i> <i>...</i> <i>cincuenta segundos</i>				
24	x	60	x	6	=	8640

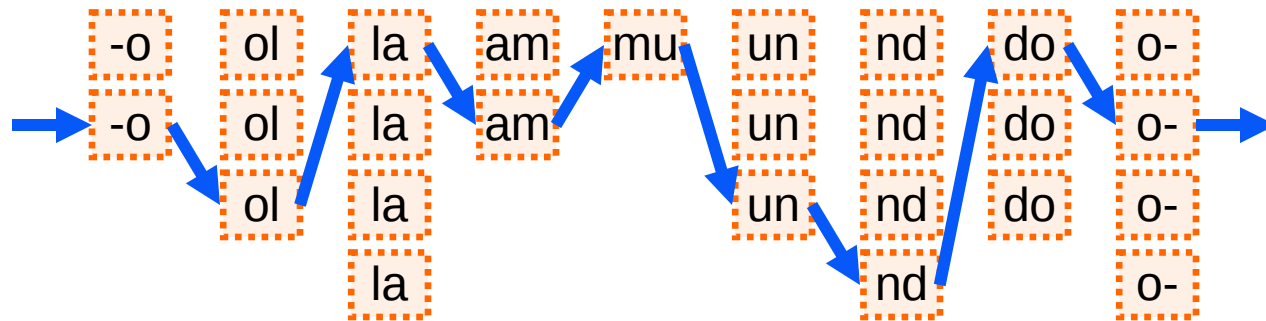
Síntesis del Habla

Difonos



Síntesis del Habla

- Base de datos de difonos.
 - Parámetros acústicos de cada difono.
- Ejemplo de síntesis
 - *hola mundo* → -o ol la am mu un nd do o-



Secuencia que minimiza penalidades.

Síntesis del Habla

- Demos

- Nuance: <http://www.nuance.com/.../tts-demo/spanish/>
- Cepstral: <http://www.cepstral.com/en/demos>

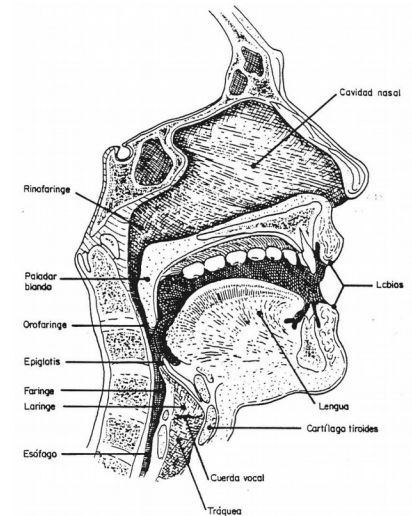
- TP1 (de años anteriores)

- *[tp1-*.wav]*

Síntesis del Habla

Otros enfoques

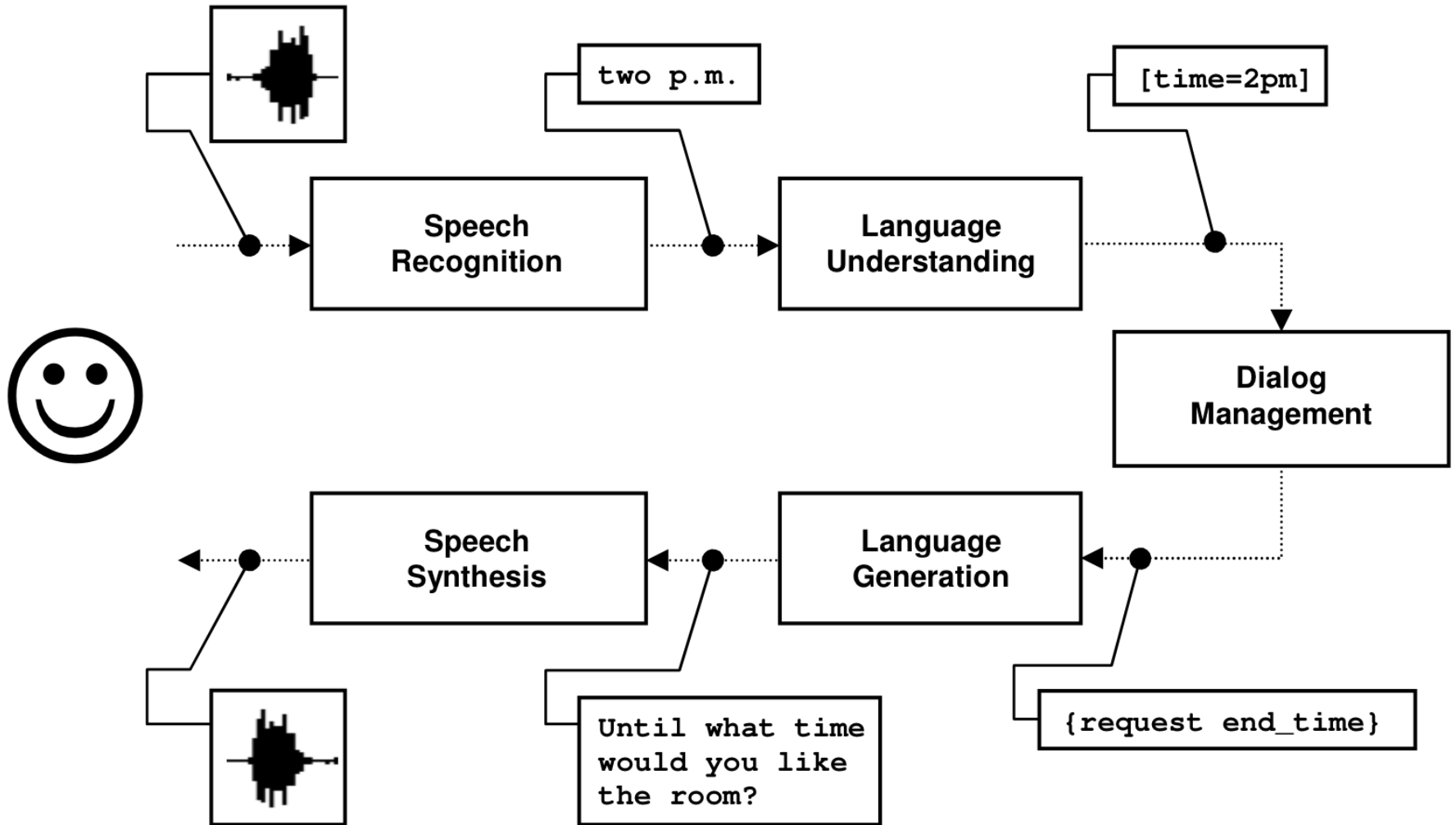
- Síntesis articulatoria.
 - Simulación del tracto vocal y de los procesos articulatorios.
- Síntesis por formantes.
- Síntesis basada en HMMs.
 - *[uba_secyt-hsmm*.wav]*



Síntesis del Habla

- Pre-procesamiento de la entrada:
 - Normalización
 - *110* → *ciento diez / uno uno cero*
 - *DGI* → *de ge i*; *AFIP* → *afip*
 - Palabras fuera de vocabulario; extranjeras.
 - *Monroe, Wilde.*
- Generación de lenguaje natural:
 - Rosario-Mendoza 2014/11/29 18:00, ASIENTOS=[].
“No quedan pasajes de Rosario a Mendoza para el veinte de noviembre a las dieciocho horas.”

Sistemas de Diálogo Hablado



Prosodia del Habla

- Mucha información más allá de las palabras.
- ¿Cómo varía la prosodia?
 - “*hace frío [. ? !]*”
 - “*no dije Corea del Sur, dije Corea del **Norte***”
 - “*por un lado ... por otro lado...*”
 - “*no renuncié por el sueldo*”
 - “*¿vamos al cine o al teatro?*”
 - “*no cantes victoria*”

Prosodia del Habla

- Tono de voz.
 - Frecuencia fundamental (f_0).
- Intensidad.
 - Volumen, energía, amplitud de onda.
- Velocidad.
 - Palabras/seg, sílabas/seg, fonemas/seg.
- Calidad de la voz.
 - Susurro, voz tensa, voz rasposa, etc.

Procesamiento del Habla

- Reconocimiento automático.
- Síntesis del habla.
- Sistemas de diálogo hablado.
- Traducción automática.
- Identificación del hablante.
- Interfaces de usuario (PC, auto, celular).
- Indexación de bases de datos de audio (YouTube).
- Detección y generación de emociones.

Procesamiento del Habla

- Interfaces para personas con capacidades especiales.
 - [Stephen Hawking](#)
- Manipulación de voces (cambio de identidad).
 - [vocalid.org](#) - TED Talk de Rupal Patel
- Procesamiento automático de la música.
 - [Hatsune Miku](#)
- “Habla” de animales (delfines, ballenas, pájaros).
 - [Laboratorio de Sistemas Dinámicos](#) (DF)
- Adquisición del lenguaje.
- ...

Temas del programa

- Acústica y procesamiento digital de señales.
- Fonética y prosodia.
- Modelo del lenguaje. Técnicas de PLN.
- Sistemas de texto-a-habla (TTS).
- Sistemas de reconocimiento automático del habla (ASR).
- Sistemas de diálogo hablado.
- Evaluación de sistemas de procesamiento del habla.
- Temas avanzados:
 - reconocimiento de información del hablante (id, edad, sexo);
 - procesamiento del habla afectiva (emociones, mentiras);
 - reconocimiento del idioma o dialecto;
 - traducción automática y generación automática de resúmenes;
 - y otros.

Varios

- Horario de cursada: lunes 9:30-13:30h, labo Turing.
- Puntaje para Computación: 3 puntos para Lic y Doc.
- Correlativas: Teoría de Lenguajes y Métodos Numéricos.
- Modo de evaluación:
 - 2 parciales y 2 trabajos prácticos grupales.
 - Promoción:
 - La materia se promociona si en todas las instancias (P1, P2, TP1, TP2) se obtiene nota 75/100 o superior.
 - Quienes obtengan al menos una nota inferior a 75/100 deben dar el final.
 - Para quienes recuperen un parcial o TP, la nota que cuenta es la del recuperatorio.
 - La promoción es opcional: pueden optar por dar final para levantar la nota.
- Mails: gravano@dc ; ph-alu@dc
- <http://www.dc.uba.ar/materias/ph/>

Bibliografía

- Jurafsky & Martin, “*Speech and Language Processing*”, (2nd ed.). Prentice Hall. 2009. En Biblioteca Central.
- Johnson, “*Acoustic & Auditory Phonetics*”, (2nd ed.). Blackwell. 2003. En Biblioteca Central.
- Benesty, Mohan Sondhi & Huang (Eds.), “*Springer Handbook of Speech Processing*”. Springer-Verlag, 2008. Versión electrónica disponible desde de la red de la UBA (ver web de la materia)

Prerrequisitos técnicos

- **Linux**

- manejo de archivos (ls, cd, pwd, cp, mv, rm, mkdir, rmdir, ...); permisos (chmod, chown, chgrp); procesamiento de archivos de texto (cat, grep, less, sed, awk, cut); man.
- <http://www.linux.org/forums/beginner-tutorials.53/>

- **Python 2.7.x**

- expresiones, variables, funciones, listas, listas por comprensión, estructuras de control (if, for, while), iteradores, lectura/escritura de archivos, cómo importar módulos.
- <https://docs.python.org/2/tutorial/>

Para hacer ahora...

- 1) Buscar el nombre de usuario (“ph-NN”) en la planilla.
 - Es para usar en Linux, no en Windows.
- 2) Recordar el nombre de usuario!!!
- 3) Ingresar y cambiar el password AHORA.
 - Ingresar con usuario y contraseña (ver el pizarrón).
 - Abrir una Terminal.
 - Ingresar “passwd”, ENTER. Seguir las instrucciones.
- 4) (opcional) Tutorial Linux.
 - <http://www.linux.org/forums/beginner-tutorials.53/>
- 5) (opcional) Tutorial Python.
 - <https://docs.python.org/2/tutorial/>