



Tema 1. Introducción al Procesamiento del Habla

Procesamiento del Habla

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

Disciplina fuertemente interdisciplinaria.



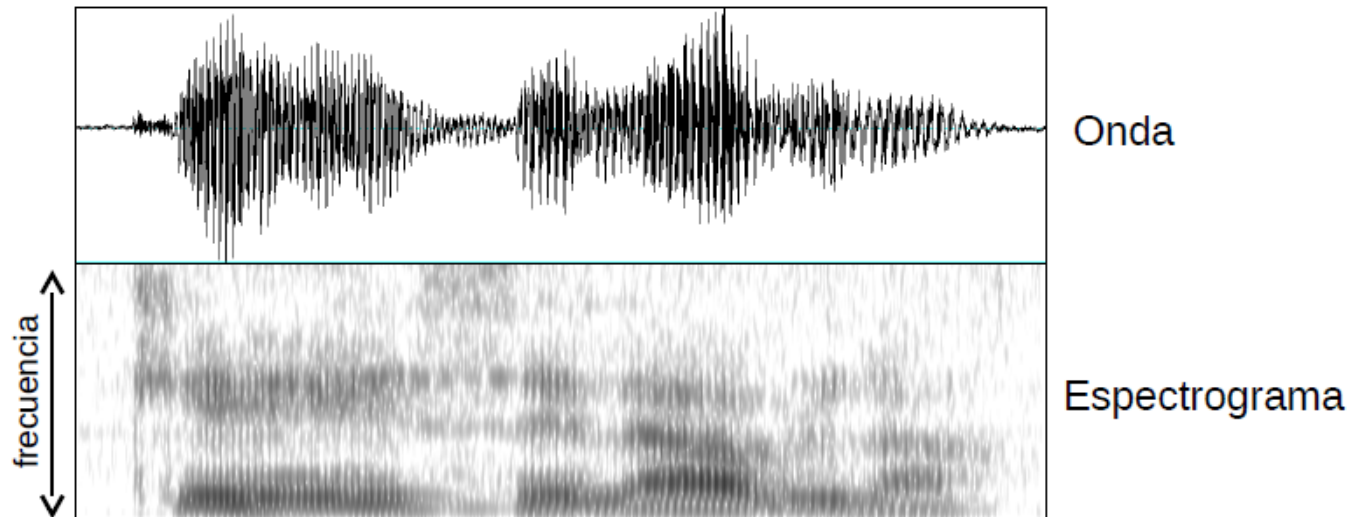
Tema 1. Introducción al Procesamiento del Habla

- 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.

Tema 1. Introducción al Procesamiento del Habla

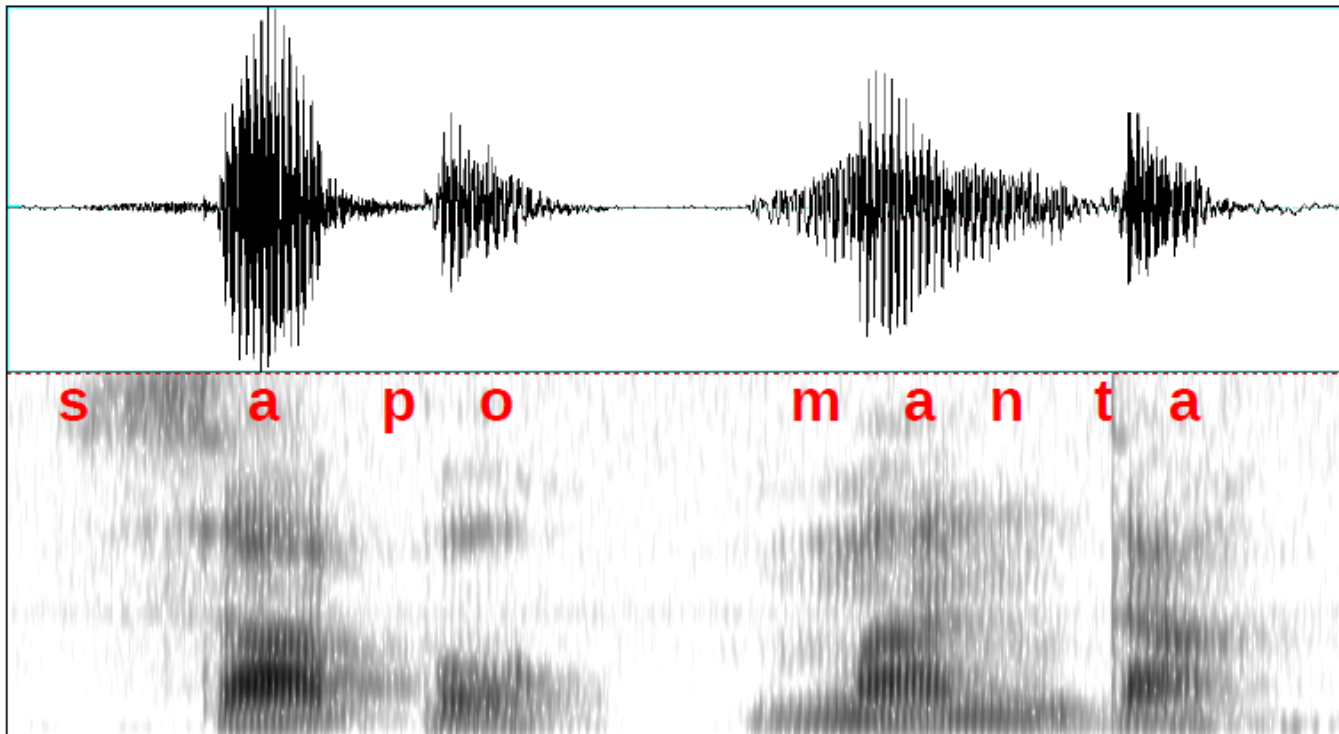
Reconocimiento Automático

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



Tema 1. Introducción al Procesamiento del Habla

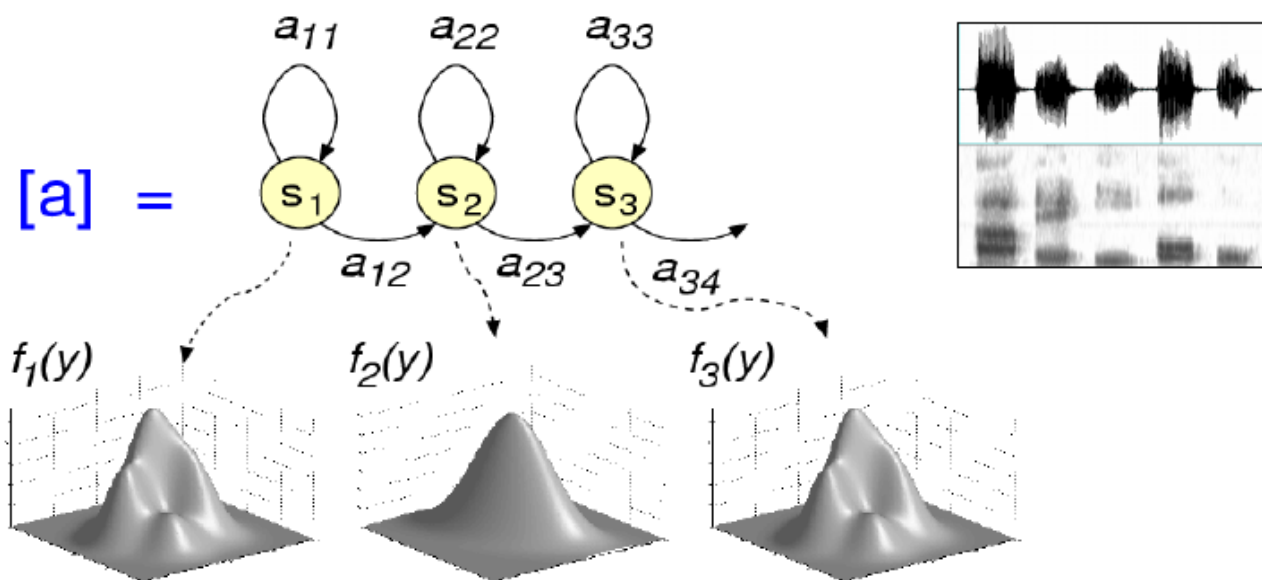
Reconocimiento Automático



Tema 1. Introducción al Procesamiento del Habla

Reconocimiento Automático

Hidden Markov Models



Ultimos años: Redes Neuronales Profundas



Tema 1. Introducción al Procesamiento del Habla

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 billete de Huelva a Madrid para el veinte de noviembre a las seis de la tarde"*
De: Huelva **a:** Madrid **fecha:** 20/11/17 **hora:**18:00

Tema 1. Introducción al Procesamiento del Habla

Síntesis del Habla

Secuencia de palabras → Habla

Hora oficial: 113 [113.wav]



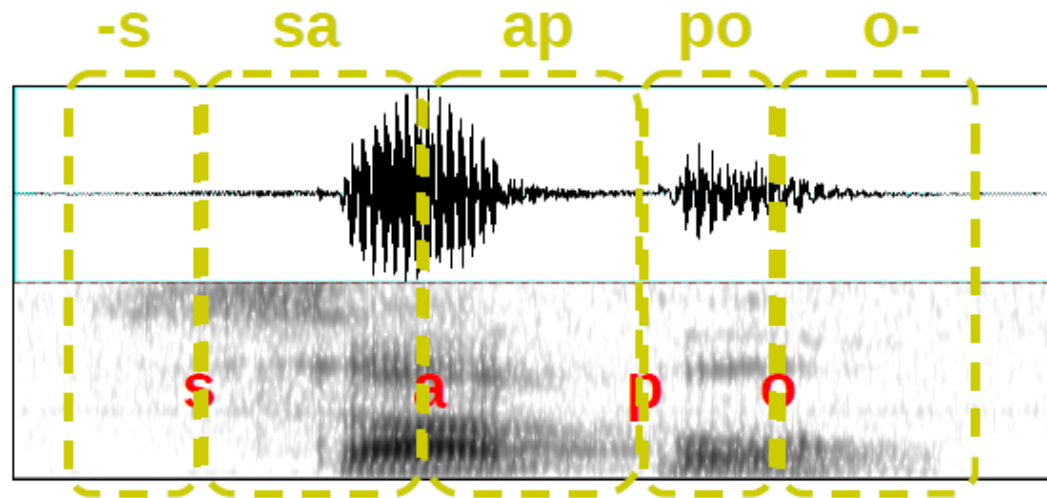
<i>cero horas</i>	<i>cero minutos</i>	<i>cero segundos</i>
<i>una hora</i>	<i>un minuto</i>	<i>diez segundos</i>
<i>dos horas</i>	<i>dos minutos</i>	<i>veinte segundos</i>
<i>...</i>	<i>...</i>	<i>...</i>
<i>veintitrés horas</i>	<i>cincuenta y nueve minutos</i>	<i>cincuenta segundos</i>

24 x 60 x 6 = 8640

Tema 1. Introducción al Procesamiento del Habla

Síntesis del Habla

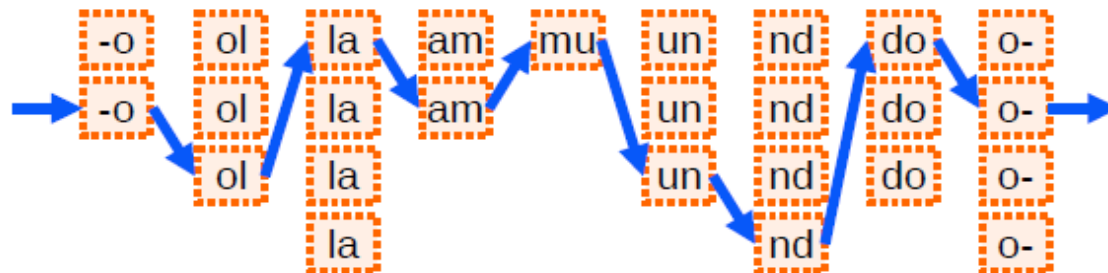
Difonos



Tema 1. Introducción al Procesamiento del Habla

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.

Tema 1. Introducción al Procesamiento del Habla

Síntesis del Habla

Demos:

- Nuance: <http://www.nuance.es/empresas/solucion/soluciones-de-atencion-al-cliente/servicios-y-soluciones/soluciones-de-recepcion-de-llamadas/loquendo-small-business-bundle/tts-demo/spanish/index.htm>
- Cepstral: <http://www.cepstral.com/en/demos>
- TP1(de años anteriores) tp1-.wav

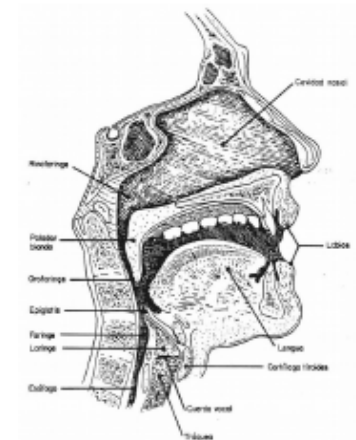


Tema 1. Introducción al Procesamiento del Habla

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]*





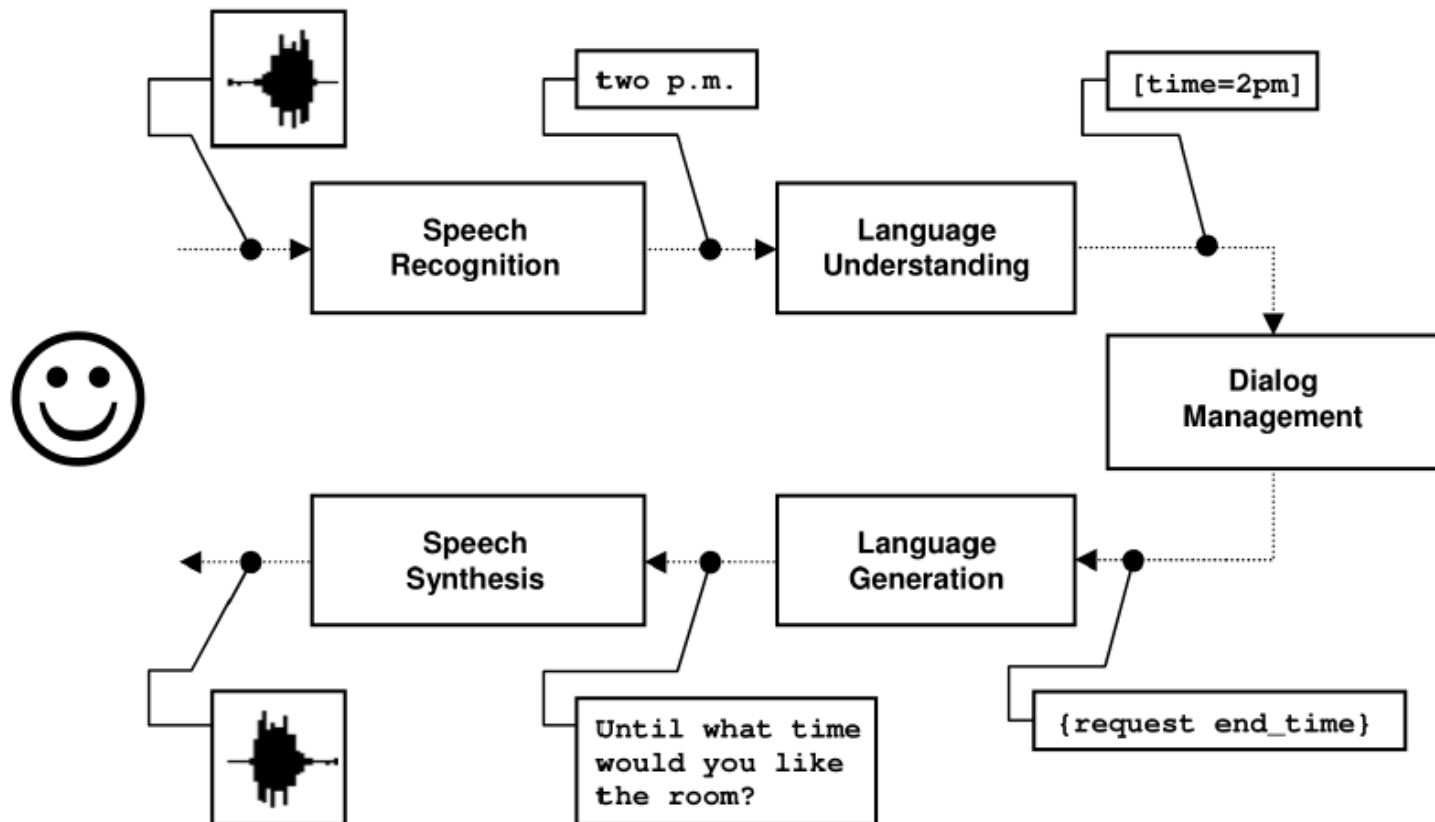
Tema 1. Introducción al Procesamiento del Habla

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.”

Tema 1. Introducción al Procesamiento del Habla

Sistemas de Diálogo Hablado





Tema 1. Introducción al Procesamiento del Habla

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”*



Tema 1. Introducción al Procesamiento del Habla

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.



Tema 1. Introducción al Procesamiento del Habla

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.



Tema 1. Introducción al Procesamiento del Habla

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.
- ...



Tema 1. Introducción al Procesamiento del Habla

Temas

- 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.



Tema 1. Introducción al Procesamiento del Habla

- **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>

Tema 1. Introducción al Procesamiento del Habla

Acústica

¿Qué es el sonido?





Tema 1. Introducción al Procesamiento del Habla

Sonido

- Fluctuaciones de presión en el aire causadas por fuentes como instrumentos musicales, bocinas de autos, voces, etc.
 - 1) Las ondas se propagan por el aire y llegan al oído.
 - 2) El sistema auditivo las traduce a impulsos neuronales.
 - 3) El cerebro los interpreta como sonido.

Tema 1. Introducción al Procesamiento del Habla

Propagación del Sonido: Analogías



Analogía más adecuada:



Tema 1. Introducción al Procesamiento del Habla





Tema 1. Introducción al Procesamiento del Habla

Sonidos

- Sonidos periódicos
 - Ondas periódicas simples
 - Ondas periódicas complejas (o compuestas)
- Sonidos aperiódicos
 - Ruido blanco, sonido transitorio.

Tema 1. Introducción al Procesamiento del Habla

Ondas Periódicas Simples

- Ondas senoidales.
- **Ciclo**: Cada repetición del patrón senoidal.
- **Período** (T): Duración del ciclo.
- **Frecuencia** (f): Cantidad de ciclos por segundo.
 - En otras palabras: Cantidad de veces que el patrón senoidal se repite por segundo.
 - Se mide en hertz (Hz).
 - $1 \text{ Hz} = 1 \text{ ciclo por segundo}$.
 - $f = 1/T$
- Ej: $T=0.01\text{s}$, $f=100\text{Hz}$

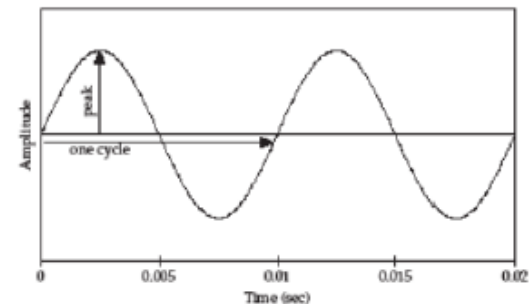


Figure 1.2 A 100 Hz sine wave with the duration of one cycle (the period) and the peak amplitude labeled.

Tema 1. Introducción al Procesamiento del Habla

Ondas Periódicas Simples

- **Amplitud:** Desviación máxima de fluctuación de la presión por sobre la presión atmosférica normal.
- **Fase:** Timing de la forma de onda relativo a algún punto de referencia.

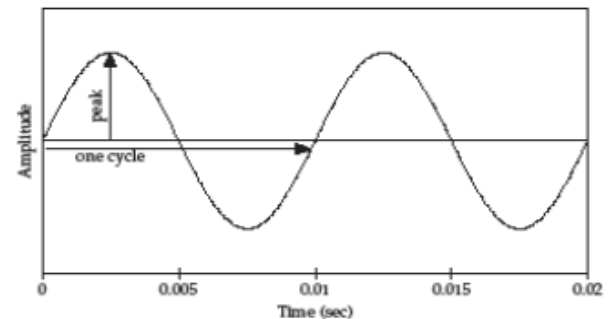


Figure 1.3 A 100 Hz sine wave with the duration of one cycle (the period) and the peak amplitude labeled.

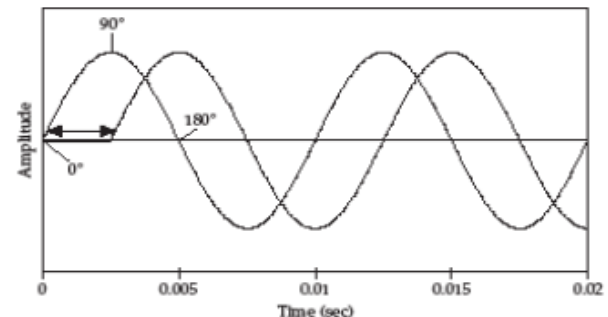


Figure 1.4 Two sine waves with identical frequency and amplitude, but 90° out of phase.

Tema 1. Introducción al Procesamiento del Habla

Ondas Periódicas Complejas

- **Ondas periódicas complejas:** Ondas cíclicas formadas por múltiples ondas senoidales.
 - Ej: onda formada por senoidales de 100 y 1000Hz.

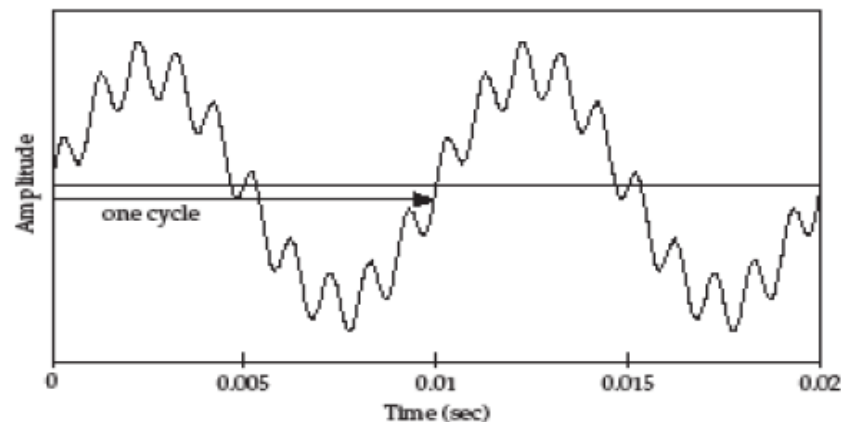
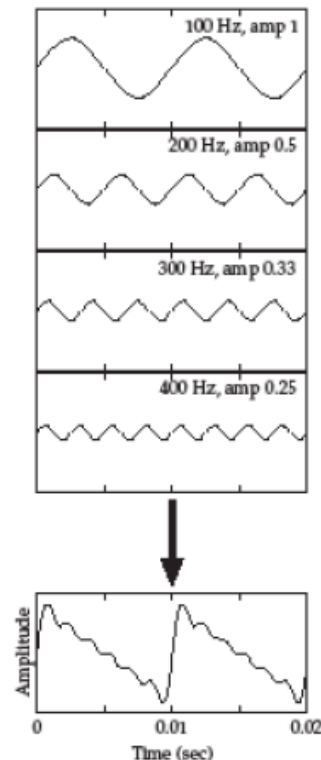


Figure 1.5 A complex periodic wave composed of a 100 Hz sine wave and a 1,000 Hz sine wave. One cycle of the fundamental frequency (F_0) is labeled.

Tema 1. Introducción al Procesamiento del Habla

Ondas Periódicas Complejas



Ejemplo: onda compleja formada por 4 senoidales

Figure 1.6 A complex periodic wave that approximates the "sawtooth" wave shape, and the four lowest sine waves of the set that were combined to produce the complex wave.



Tema 1. Introducción al Procesamiento del Habla

Sonidos Aperiódicos

No tienen un patrón que se repita regularmente.

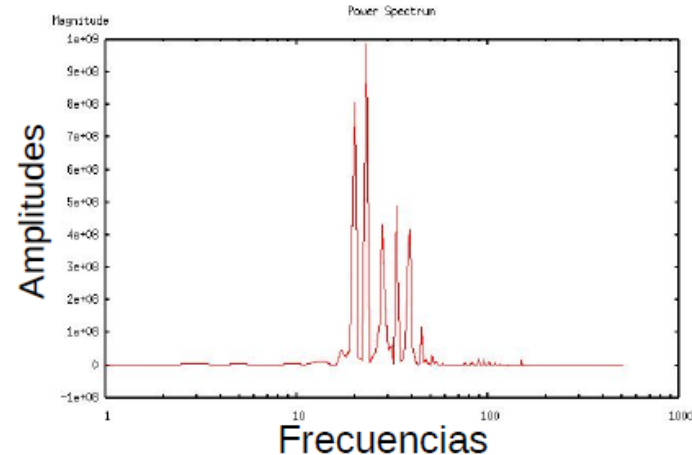
Ejemplos:

- **Ruido blanco**: Fluctuación aleatoria de presión.
 - Espectro plano: igual amplitud para todas las frecuencias.
 - Estática de radio, viento entre los árboles, [s], [f].
- **Ondas transitorias**: Fluctuaciones súbitas de presión que no se sostienen ni se repiten.
 - Portazos, disparos, mouse clicks, [p], [t].

Tema 1. Introducción al Procesamiento del Habla

Análisis de Fourier

- Cualquier función puede descomponerse en una suma de ondas senoidales.
- **Espectro energético:**

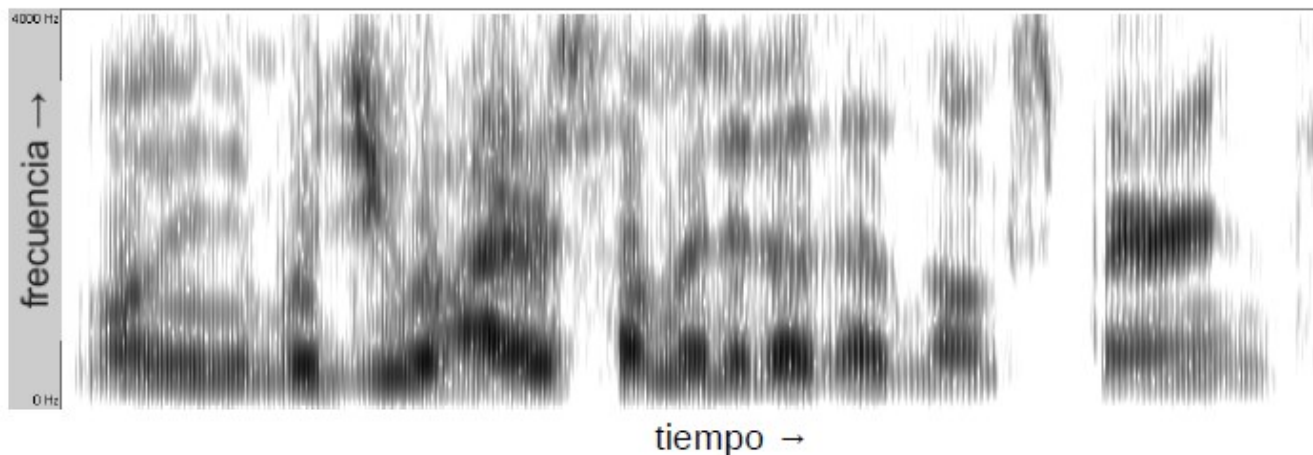


- Fast Fourier Transform (FFT)
 - Algoritmo eficiente para computar el espectro de una señal a partir de un número de muestras.

Tema 1. Introducción al Procesamiento del Habla

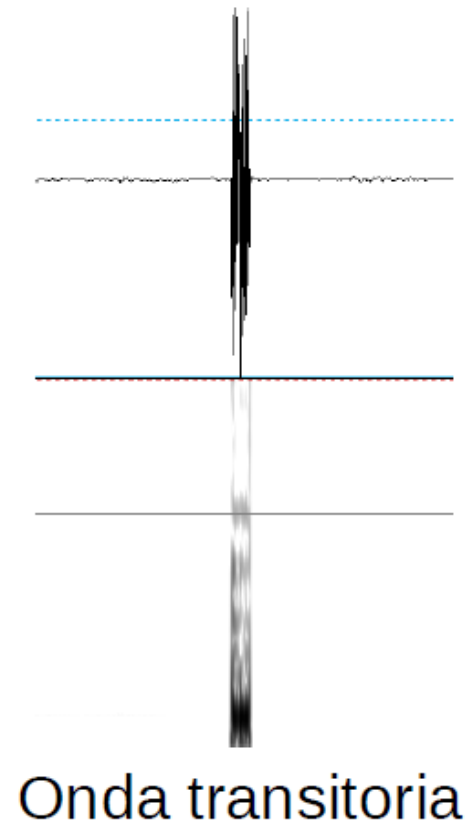
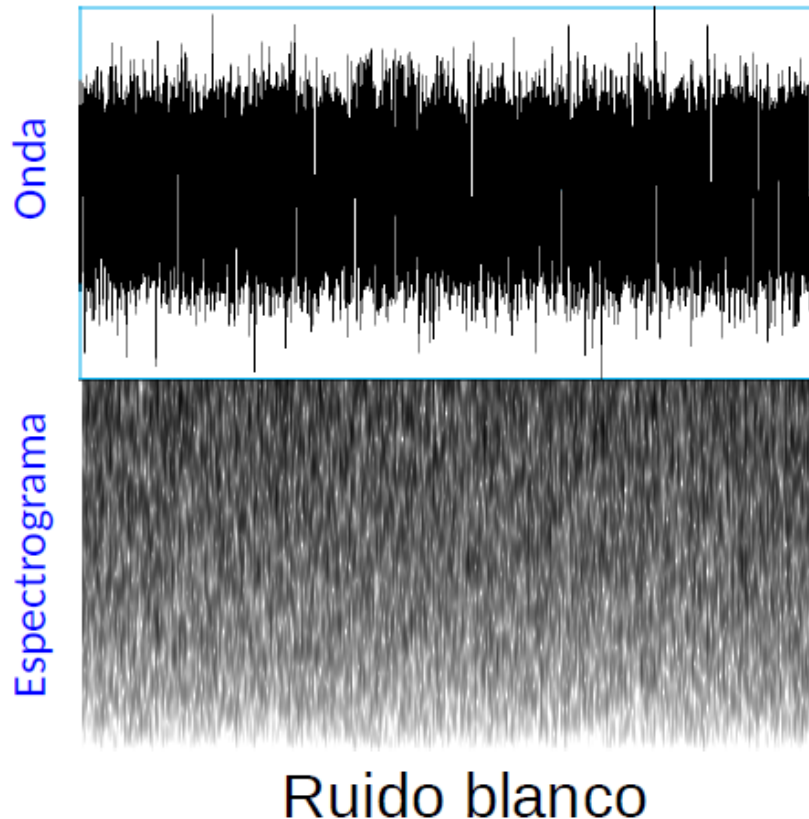
Visualizaciones

- Al agregar una dimensión temporal al espectro energético, obtenemos un **espectrograma**.



Tema 1. Introducción al Procesamiento del Habla

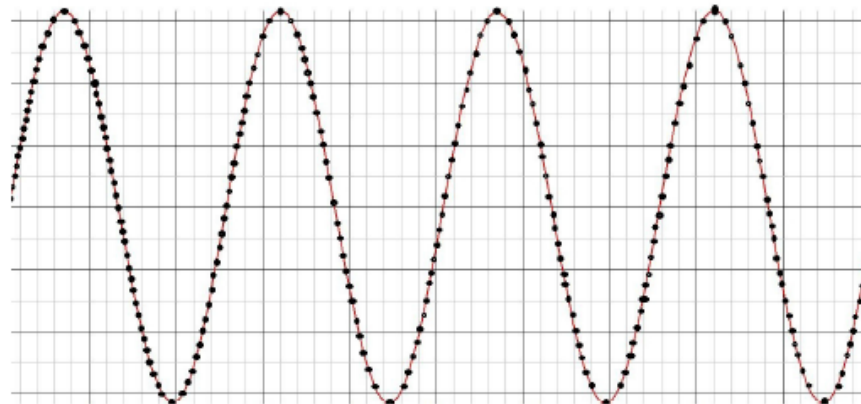
Sonidos Aperiódicos



Tema 1. Introducción al Procesamiento del Habla

Procesamiento Digital de Señales

- **Señal analógica (continua)**: La línea de tiempo tiene valores de amplitud con precisión **infinita** en **todos** los puntos.
- **Señal digital (discreta)**: La línea de tiempo tiene sólo una **secuencia** de valores de amplitud con precisión **finita**.





Procesamiento Digital de Señales

- Un **micrófono** convierte oscilaciones de presión en el aire (sonido) en oscilaciones de voltaje.
 - Los dispositivos analógicos (discos de vinilo, cassettes) las guardan como señales continuas.
 - Los dispositivos digitales (computadoras, CDs) las convierten y guardan como señales discretas.
- Conversión Analógica-Digital (Digitalización)
 - 1) **Muestreo**: Discretización del tiempo.
 - Tasa de muestreo. Ej: 8 kHz (teléfono), 44.1 kHz (CD), 16 kHz.
 - 2) **Cuantización**: Discretización de la amplitud.
 - Precisión de las muestras. Ej: 16 bits.



Tema 1. Introducción al Procesamiento del Habla

Acústica – Resumen

- Sonidos periódicos y aperiódicos.
 - Ondas periódicas simples y complejas.
 - Ruido blanco; ondas transitorias.
 - Análisis de Fourier. FFT. Espectrograma.
- Procesamiento digital de señales.
 - Tasa de muestreo; precisión.