

Proyecto Fin de Carrera

**Desarrollo de una
aplicación software para el
análisis de características
fundamentales de la voz**

Autor: David Brezmes Alonso
Tutor: Norberto Corral Blanco

Índice

- 1. Introducción
- 2. Entorno de trabajo
- 3. Estrategias para abordar el problema del reconocimiento del habla
- 4. Generación y Clasificación de los sonidos.
- 5. División vocal-consonante
- 6. Características espectrales de la voz
- 7. Parametrización

Introducción

- Marco del trabajo
 - Avances tecnología de la información
 - Análisis onda sonora
 - Síntesis del lenguaje
- Objetivos del proyecto Amper
 - Contribución a las descripciones lingüísticas y dialectales
 - Estudio entonación. Su papel en la fonética y fonología
- Objetivos de la herramienta desarrollada
 - Obtención de las características principales de la voz

Localización del proyecto Amper



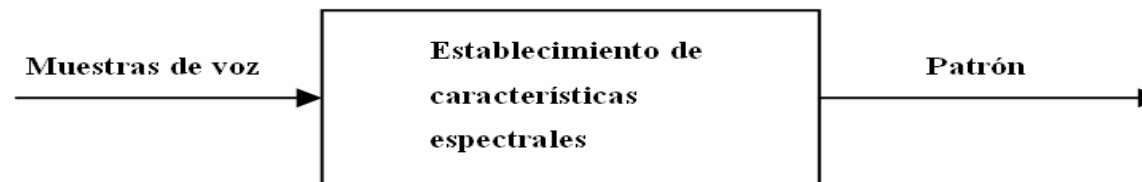
Entorno de trabajo

- **MATLab:**

- Elemento de desarrollo actual
- **Amplia gama de librerías** en el área de tratamiento de señales
- **Entorno de trabajo amigable**
- Lenguaje de alto nivel para cálculo técnico.
- Entorno de desarrollo para la gestión de código, archivos y datos.
- Herramientas interactivas para exploración, diseño y resolución de problemas iterativos.
- **Funciones matemáticas para álgebra lineal, estadística, análisis de Fourier, filtraje, optimización e integración numérica.**
- Funciones gráficas bidimensionales y tridimensionales para visualización de datos.
- Herramientas para crear interfaces gráficas de usuario personalizadas.
- Funciones para integrar los algoritmos basados en MATLAB en aplicaciones y lenguajes externos, tales como C/C++, FORTRAN, Java, COM y Microsoft Excel.

Estrategias fundamentales para abordar el problema del reconocimiento del habla

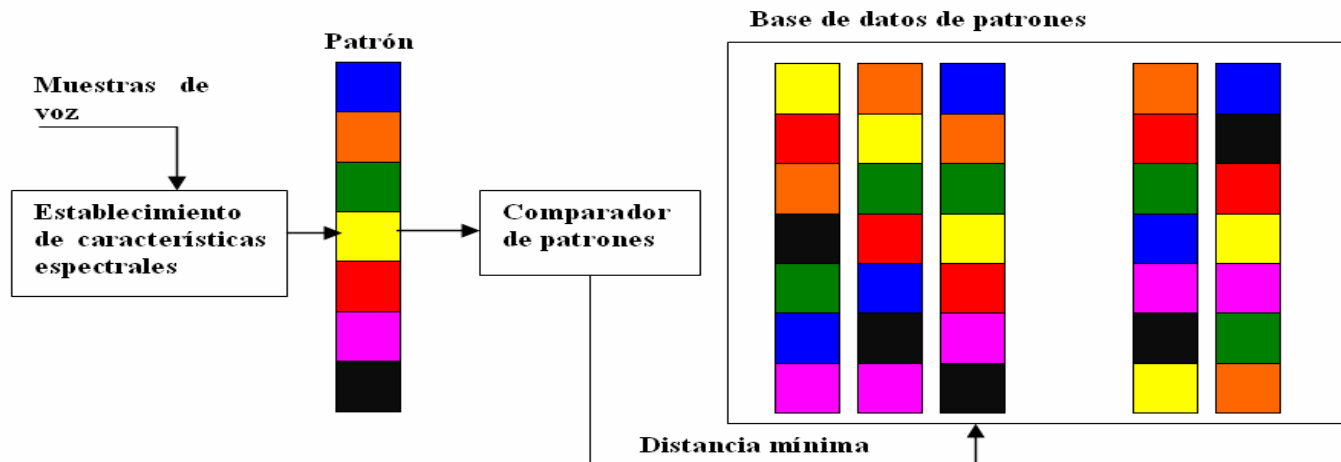
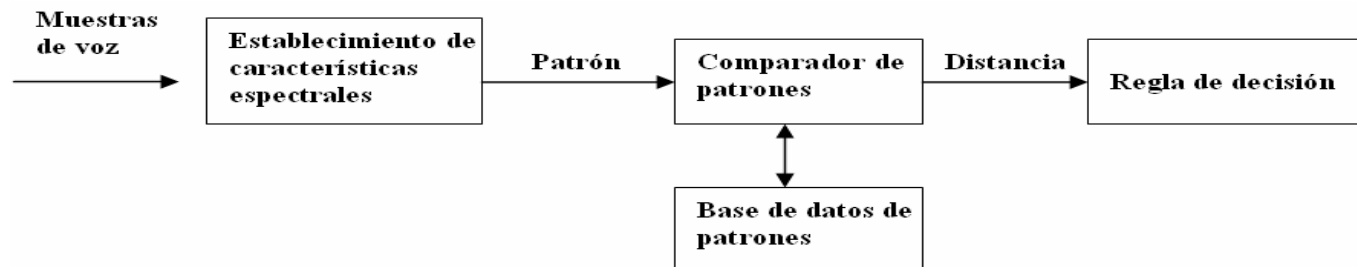
- Traspaso de la señal al dominio de la frecuencia



- Tres estrategias básicas:
 - Reconocimiento del habla empleando técnicas de comparación de patrones (I)
 - Reconocimiento del habla empleando modelos automáticos paramétricos (II)
 - Reconocimiento del habla empleando técnicas de obtención de características fonéticas (III)

Estrategias fundamentales para abordar el problema del reconocimiento del habla (I)

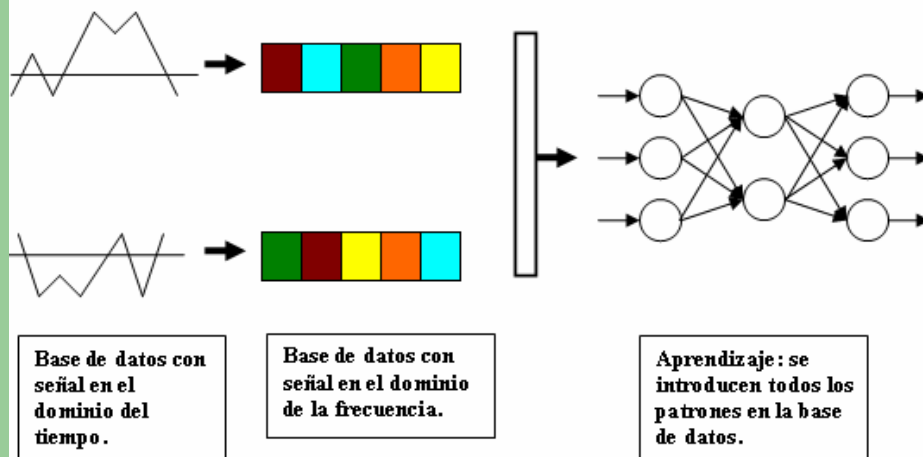
- Reconocimiento del habla empleando técnicas de comparación de patrones (I)



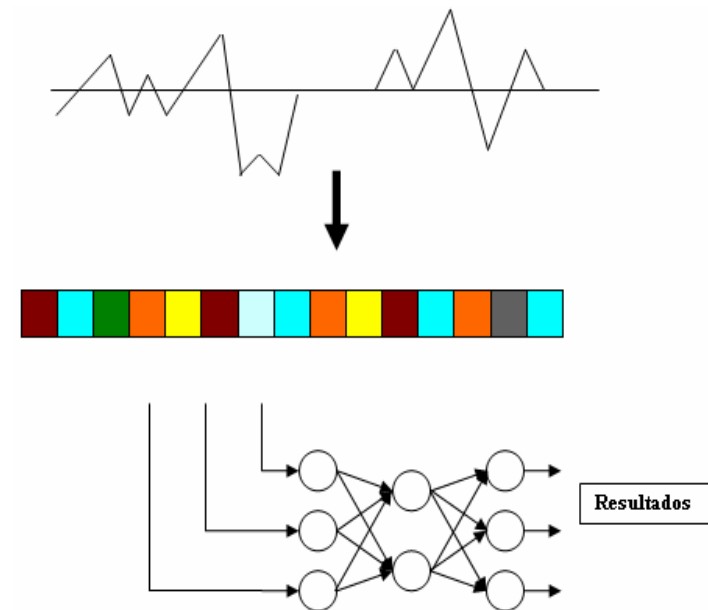
Estrategias fundamentales para abordar el problema del reconocimiento del habla (II)

- Reconocimiento del habla empleando modelos automáticos paramétricos (II)

- Fase de aprendizaje

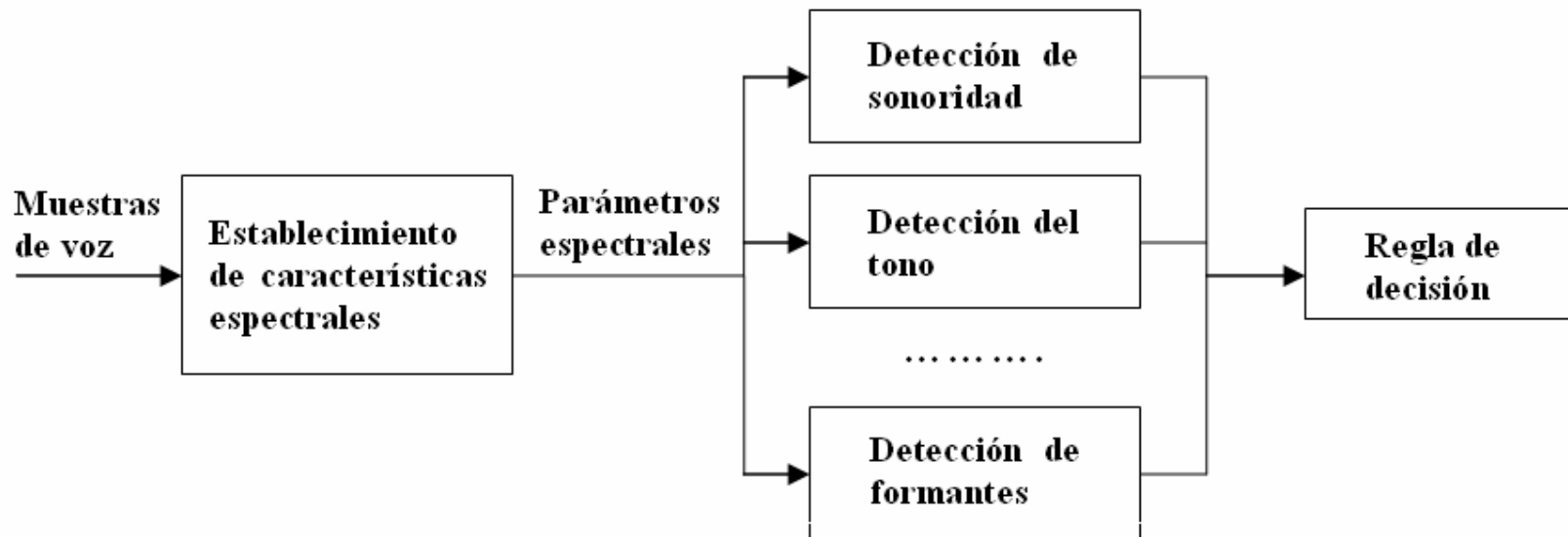


- Etapa de reconocimiento



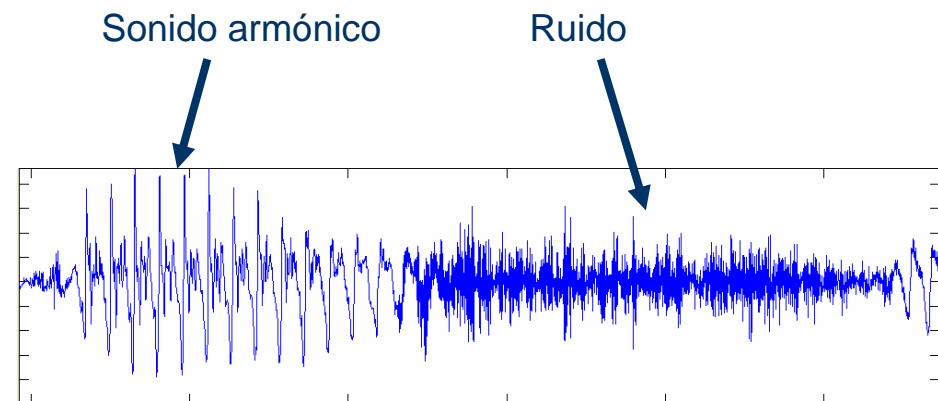
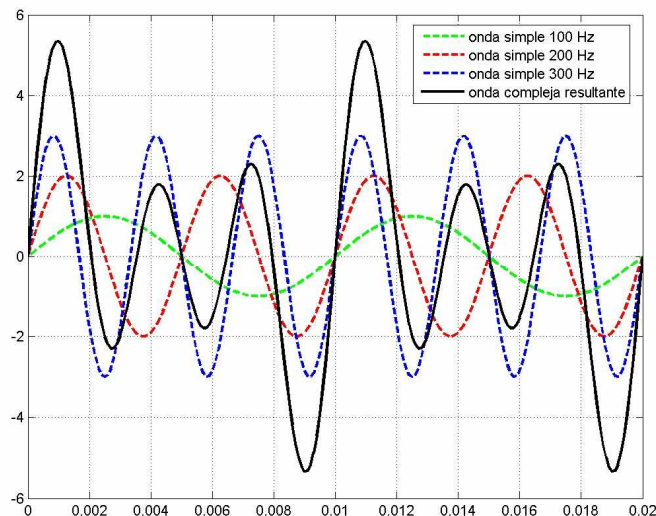
Estrategias fundamentales para abordar el problema del reconocimiento del habla (III)

- Reconocimiento del habla empleando técnicas de obtención de características fonéticas (III)



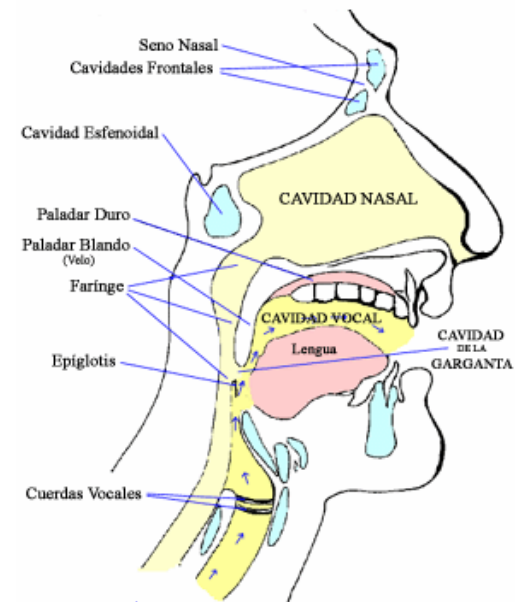
Generación del sonido (I)

- **Sonido:** Descodificación que efectúa nuestro cerebro de las vibraciones percibidas a través de los órganos de audición. Estas vibraciones se transmiten a través de cualquier medio elástico
- **Onda sonora:** Creación de un movimiento vibratorio en un cuerpo.
 - M.A.S.
 - La onda compuesta
 - Fourier
 - Ruido



Generación del sonido (II)

- La producción de los sonidos
 - Generador (fuente)
 - Articulador (moldeador)
 - Radiador (emisor)
- Glotis: Apertura/cierre cuerdas vocales: F0
- Resonancias: Formantes

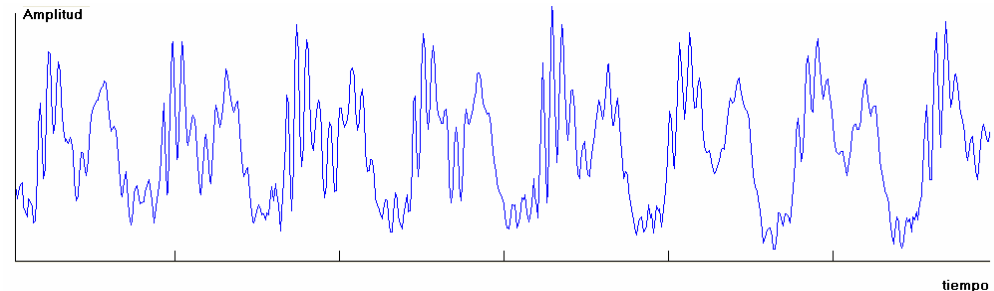


División vocal-consonante por motivos:

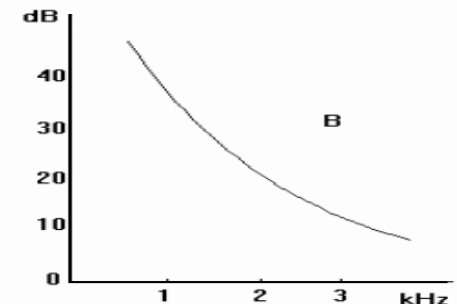
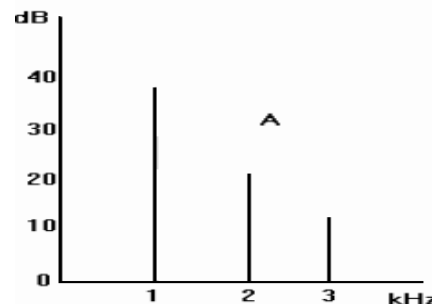
- Auditivos: Nivel de audibilidad
- Fisicoacústico: Estabilidad de los componentes acústicos de la vocal frente a la inestabilidad en la consonante. Formantes,...
- Funcional: No hay sílabas sin vocales
- Fisiológico: Straka (1963). La energía articulatoria puede marcar el límite entre la vocal y la consonante

Características espectrales de la voz

- Señal original a representación parametrizada
- Poca fiabilidad material acústico. Utilización representaciones gráficas
- Oscilograma



- Espectros



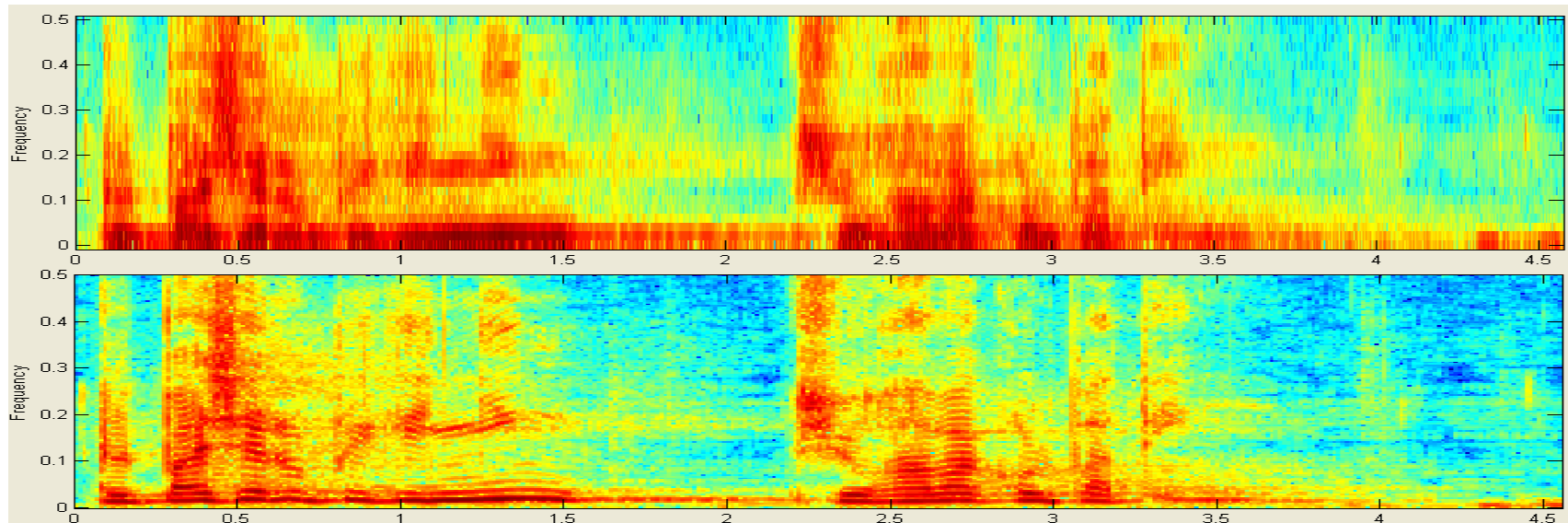
Características espectrales de la voz

- Espectrograma

- Cohen: Distribuciones tiempo-frecuencia

$$p(t, \omega) = \frac{1}{4\pi^2} \iiint e^{-j\theta t} e^{-j\tau\omega} e^{-j\mu t} \phi(\theta, \tau) x\left(\mu + \frac{1}{2}\tau\right) x^*\left(\mu - \frac{1}{2}\tau\right) d\mu d\tau d\theta$$

$$\phi(\theta, \tau) = \int h^*\left(\mu - \frac{1}{2}\tau\right) e^{-j\theta\mu} h\left(\mu + \frac{1}{2}\tau\right) d\mu,$$



Parametrización (I)

- Sonoridad: Vibración o no de las cuerdas vocales
- Nivel de Intensidad acústica: Parámetro que mide la forma en que es percibida la intensidad acústica (dB)
- Principio-fin de voz:
 - Segmentación de la voz: Tasa de cruces por cero
- Formantes: Análisis LPC

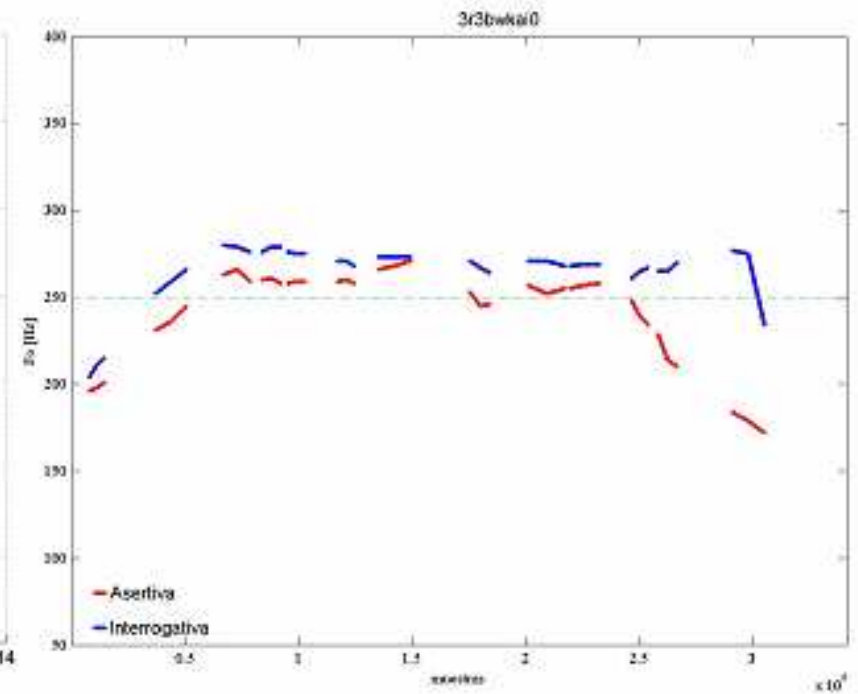
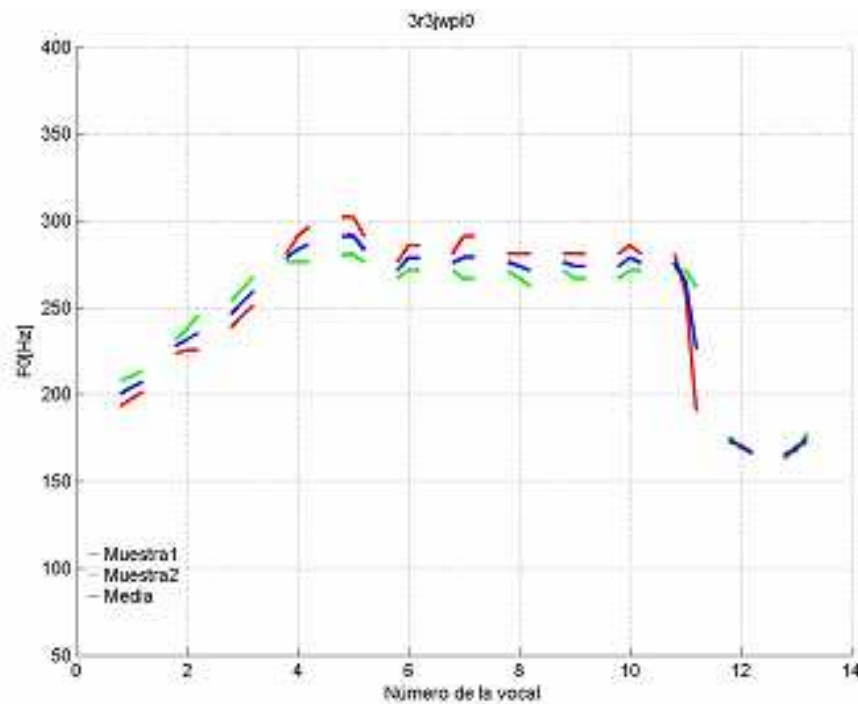
Parametrización (II)

- Frecuencia fundamental:
 - Método de la autocorrelación: Ventana de hamming con solapamiento.
 1. Filtrado paso bajo.
 2. Center Clipping.
 3. Cálculo de la energía.
 4. Autocorrelación.
 5. Límites de búsqueda.
 6. Buscamos máximos y su posición.
 7. Fijamos umbral de silencio.
 8. Alisado.

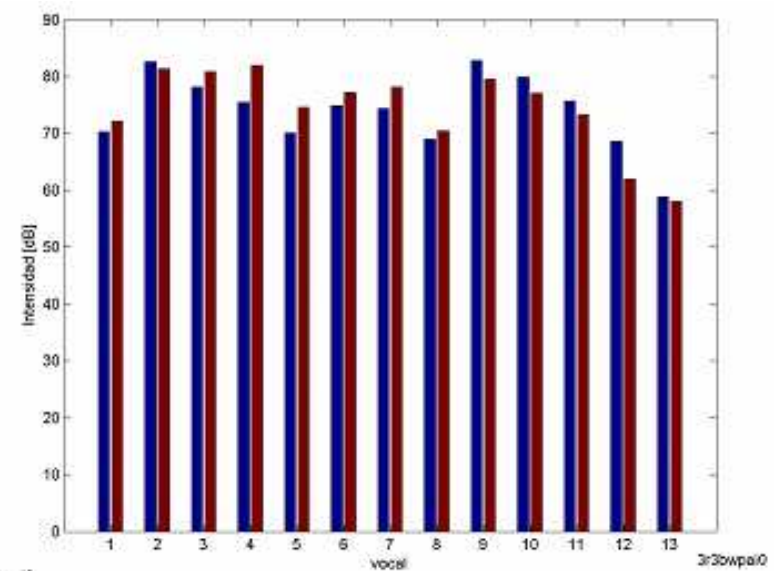
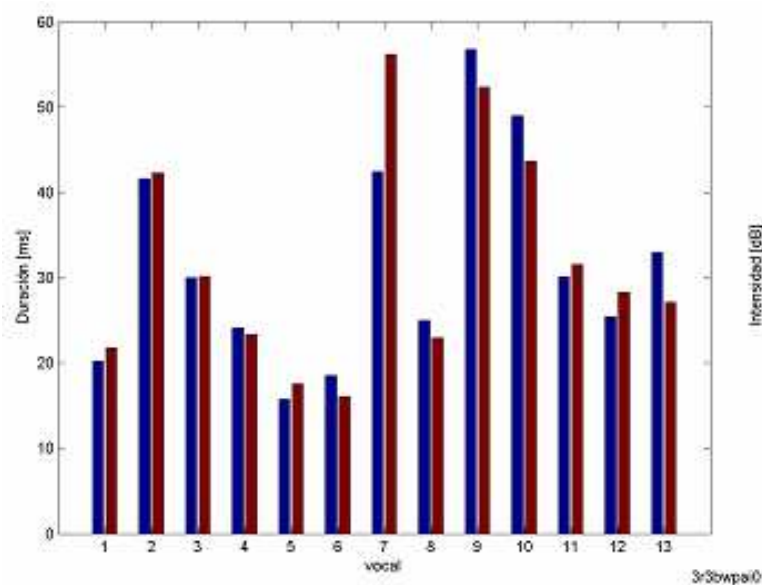
Características del programa

- Programa realizado a medida para un determinado cliente
- Diseñado para minimizar la curva de aprendizaje
- Toda la aplicación se ha realizado de manera modular: Actualización de forma sencilla
- Preparado para analizar in situ las grabaciones de los informantes
- Elevada velocidad de trabajo: Rutinas optimizadas para trabajar con gran cantidad de información

Salidas de la aplicación



Salidas de la aplicación



Salidas del programa

Microsoft Excel - 3e3bwka1

Archivo Edición Ver Insertar Formato Herramientas Datos Ventana 2

19-Apr-2007

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
	zona		duration [ms]		I1	I2	I3 [dB]		fo1	fo2	fo3 [Hz]		Izq	Central	Dcha
1															
2															
3	1		76		90	101	100		208	208	203		586	1191	1795□□□□
4	2		45		102	102	100		203	205	205		5273	5632	5991□□□□
5	3		50		89	100	100		208	242	239		7578	7975	8372□□□□
6	4		40		92	99	100		239	250	250		9014	9336	9657□□□□
7	5		38		86	99	91		235	235	229		11395	11698	12000□□□□
8	6		57		92	101	99		211	208	198		13285	13739	14192□□□□
9	7		52		98	100	99		208	216	225		15440	15855	16271□□□□
10	8		50		94	96	94		186	186	186		19635	20032	20429□□□□
11	9		33		99	99	99		182	182	180		21525	21789	22054□□□□
12	10		28		97	99	100		193	211	216		23112	23339	23566□□□□
13	11		26		99	101	101		239	239	222		25455	25663	25871□□□□
14	12		28		90	99	99		213	205	186		28290	28517	28744□□□□
15	13		66		96	97	94		184	178	170		30936	31465	31994□□□□
16	3e3bwka1.txt														
17	19-Apr-2007														
18															
19															
20															
21															
22															
23															

3e3bwka1/

Listo

Publicaciones relacionadas

- “Análisis y representación de la entonación. Replanteamiento metodológico en el marco del proyecto Amper”.
 - Workshop Internacional III Jornadas Científicas del Proyecto Amper.

Proyecto Amper

América Latina



Europa