



# ***WAVELETS EN EL RECONOCIMIENTO DE VOZ***

***Ismael Chávez Álvarez***

***José Antonio Camarena Ibarrola***

***Facultad de Ingeniería Eléctrica  
Universidad Michoacana de San  
Nicolás de Hidalgo***

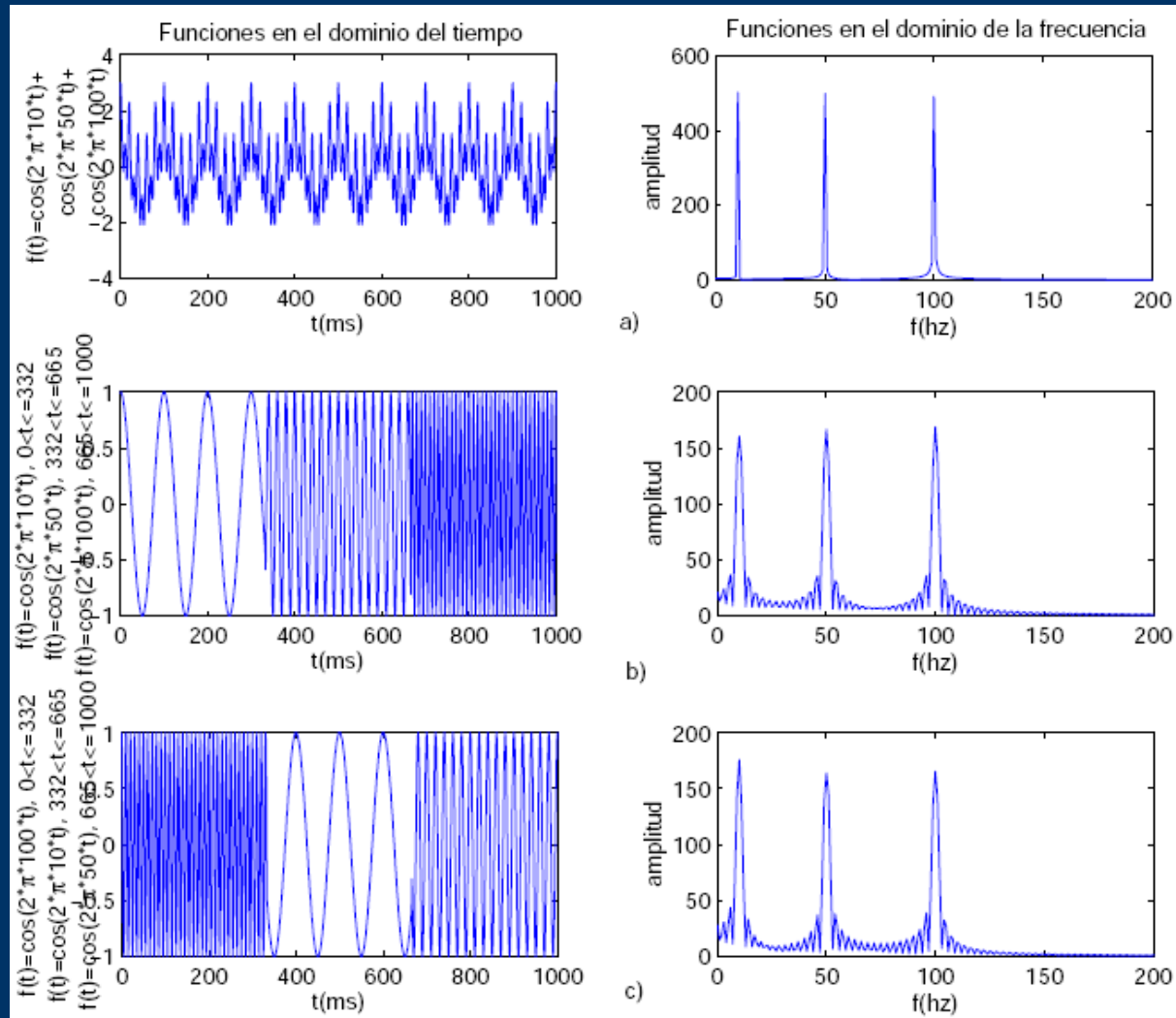
***XIII ROPEC Noviembre 2011***

---

---

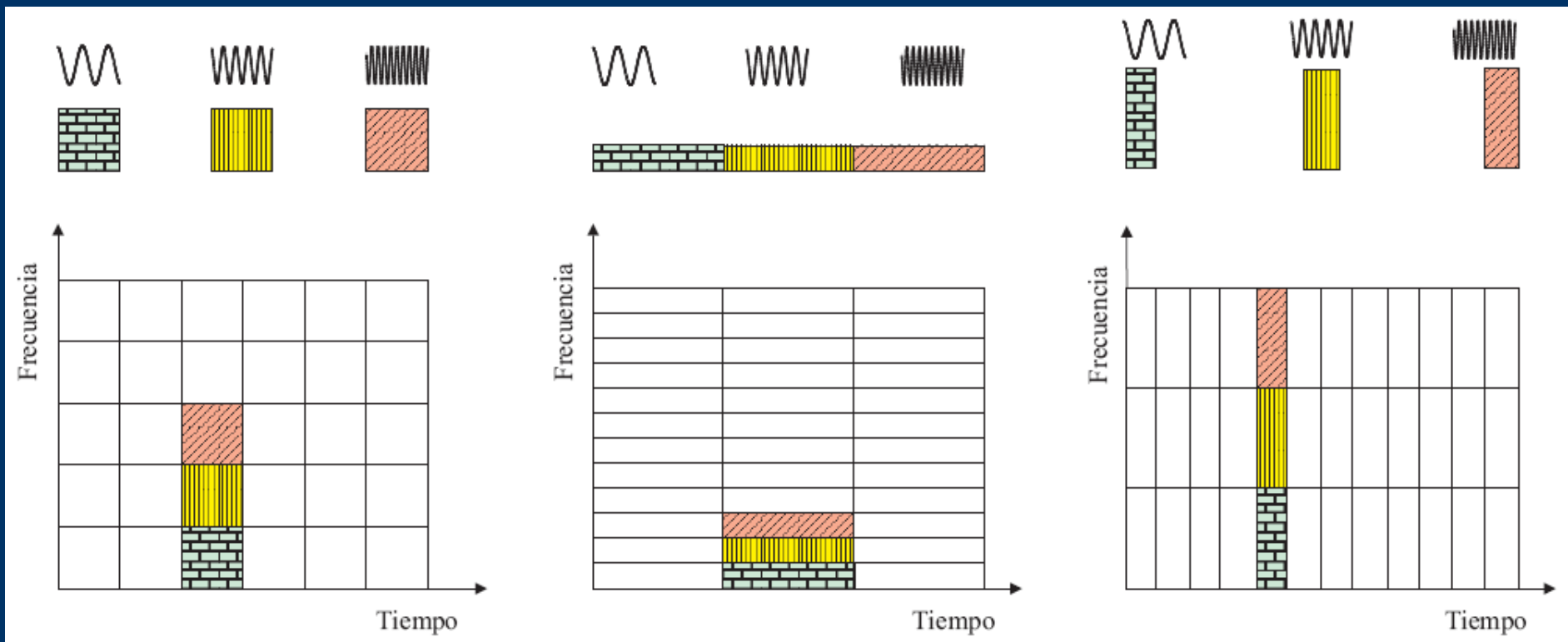
## Procesamiento digital

- Transformada de Fourier.



## Procesamiento digital

- Resolución del análisis de Fourier (constante a lo largo de un estudio).

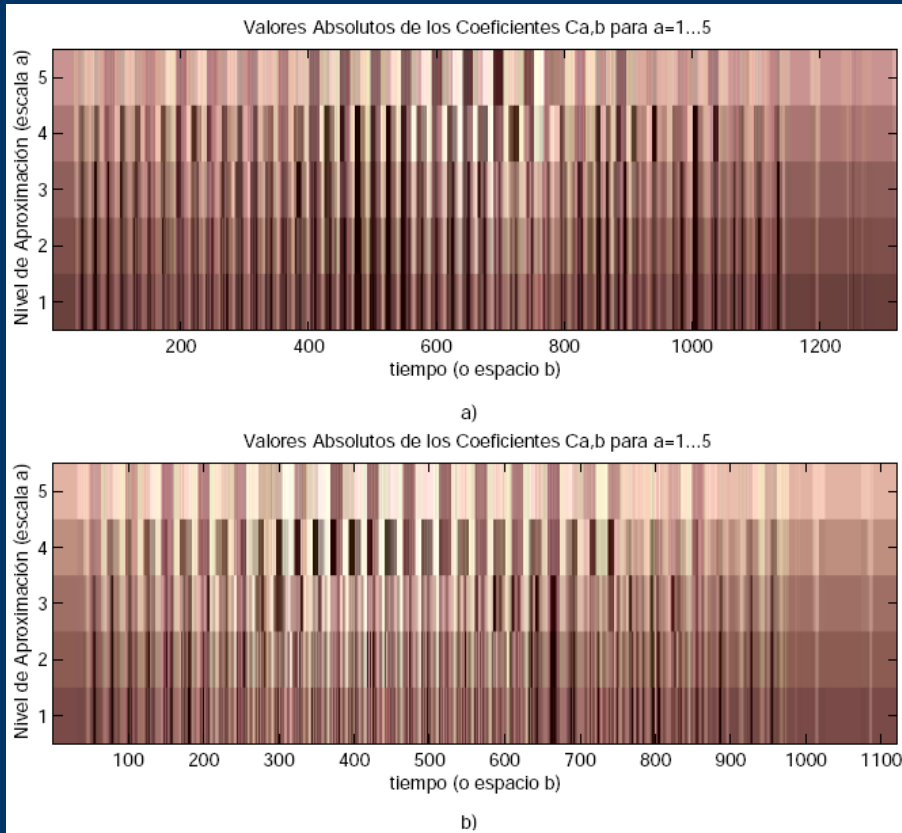


## *Procesamiento digital*

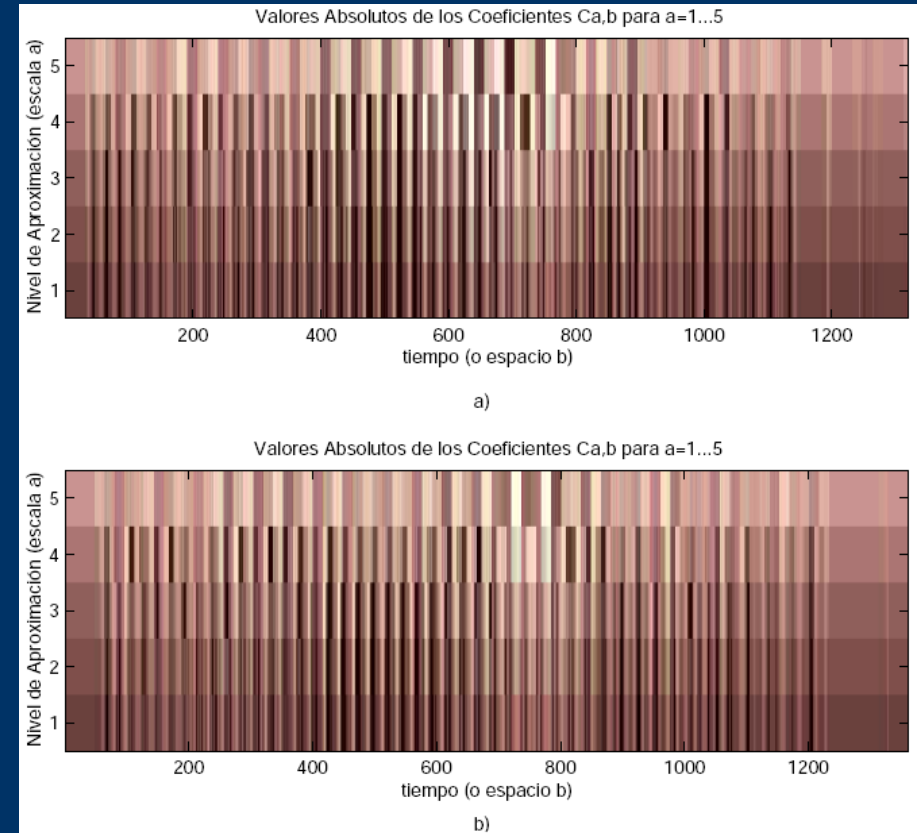
- Transformada Wavelet como técnica complementaria para la extracción de características de señales, propuesta para ser usada por sistemas de reconocimiento.
- Permite estudiar las señales en un marco acorde con su escala.

## Procesamiento digital

- Transformada Wavelet.



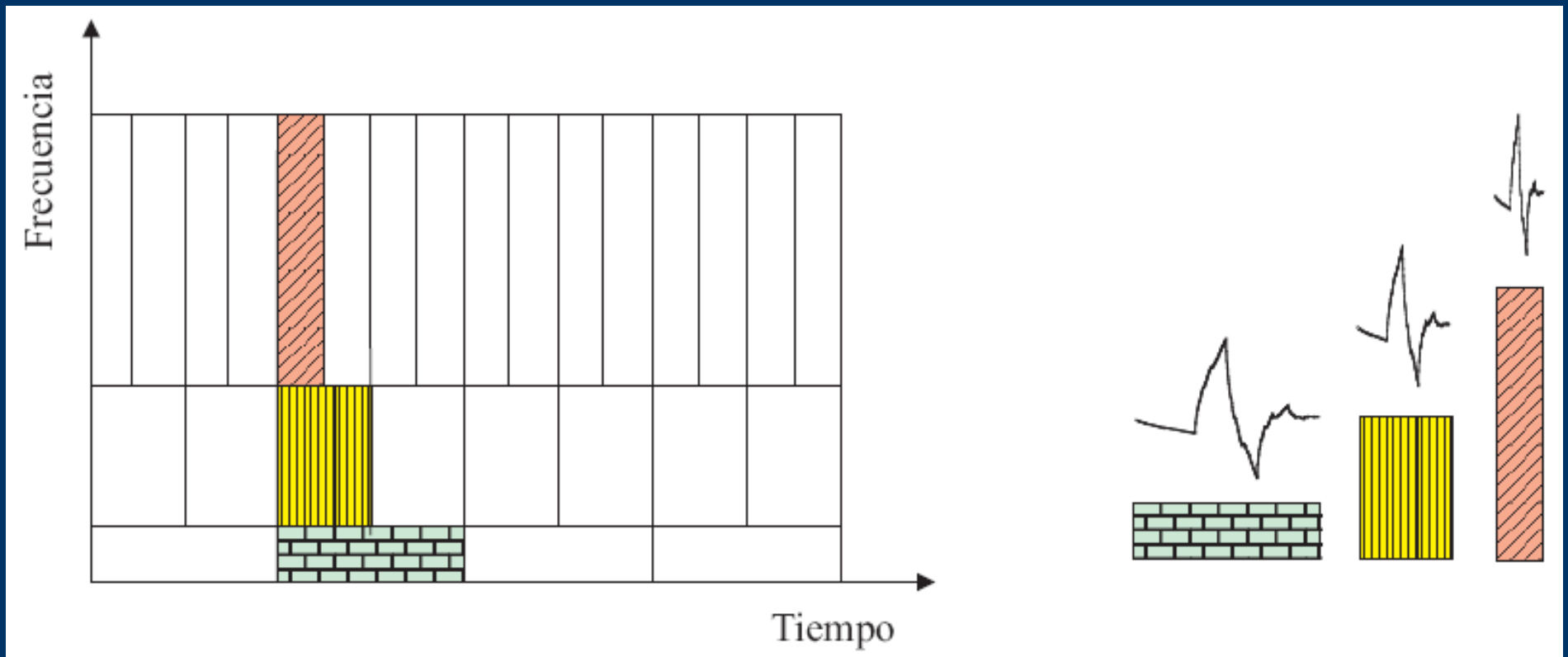
- a) Palabra “uno”
- b) Palabra “cero”



- a) Palabra “uno”
- b) Palabra “uno”

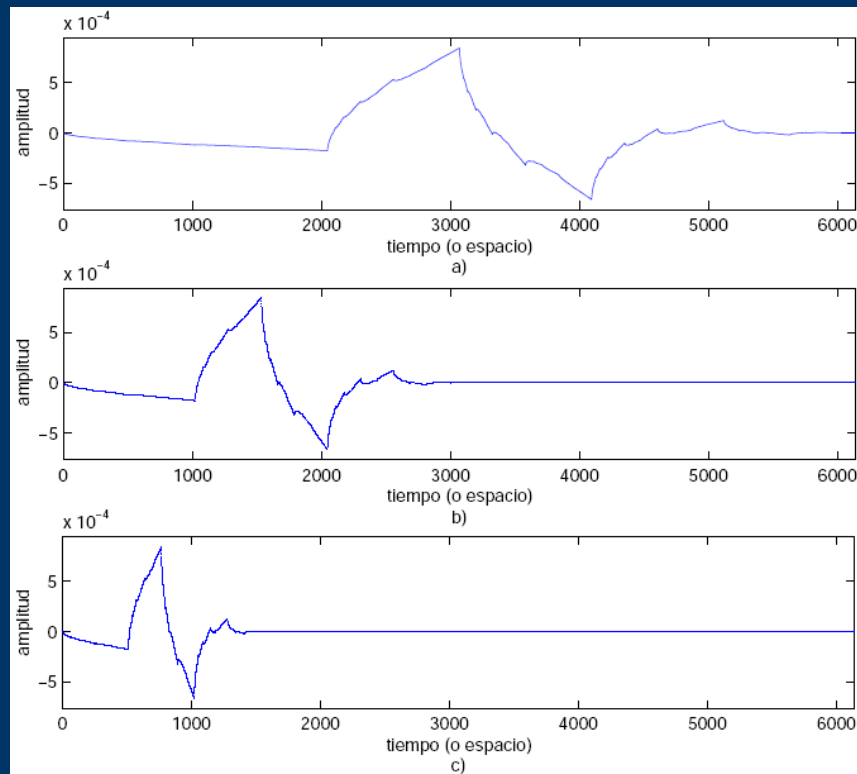
## Procesamiento digital

- Resolución del análisis de Wavelets (variable a lo largo de un estudio).

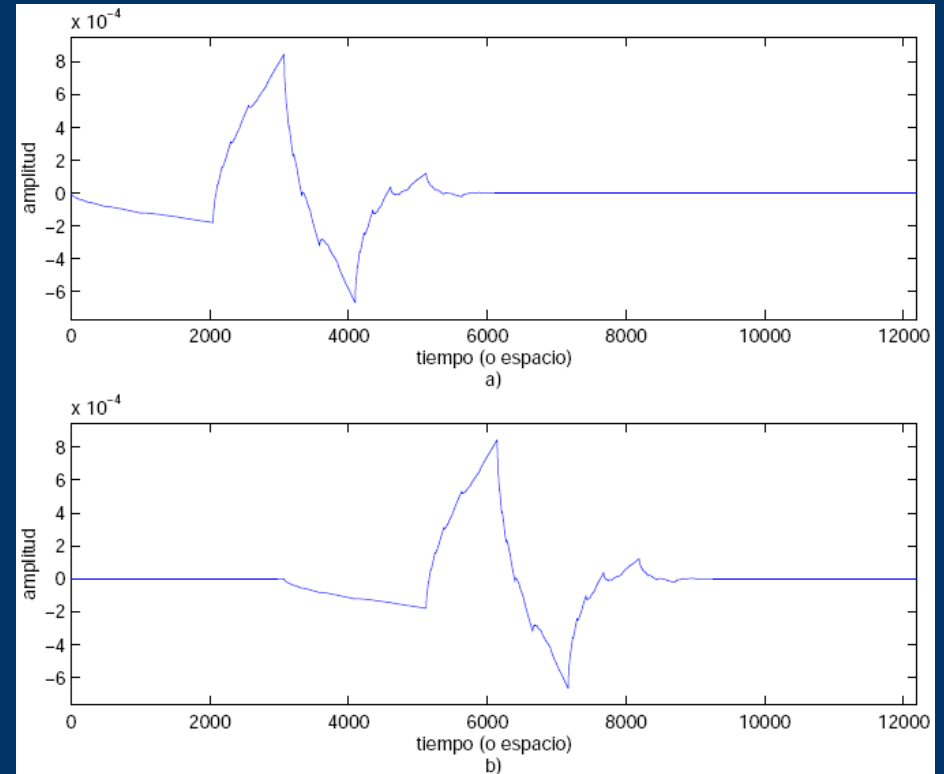


## Procesamiento digital

- La transformada Wavelet se apoya en el escalamiento y la traslación de una función base.



- Escalamiento



- Traslación

## Procesamiento digital

- Las funciones base en el análisis de Wavelets se definen a partir de los coeficientes para filtros digitales.
- El filtro pasabajas de análisis para el wavelet de Haar:

$$[L] = \left[ \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right]$$

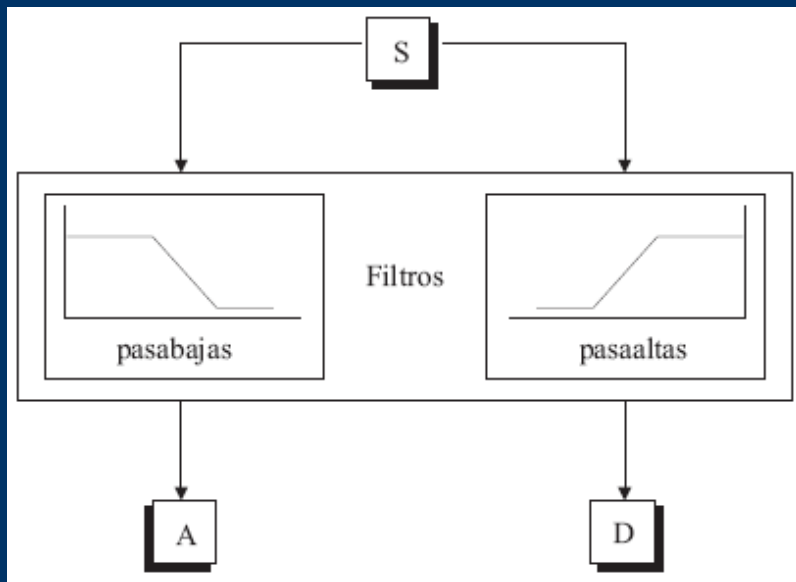
- El filtro pasabajas de análisis para el wavelet de Daubechies:

$$[L] = \left[ \frac{1 + \sqrt{3}}{4\sqrt{2}}, \frac{3 + \sqrt{3}}{4\sqrt{2}}, \frac{3 - \sqrt{3}}{4\sqrt{2}}, \frac{1 - \sqrt{3}}{4\sqrt{2}} \right]$$

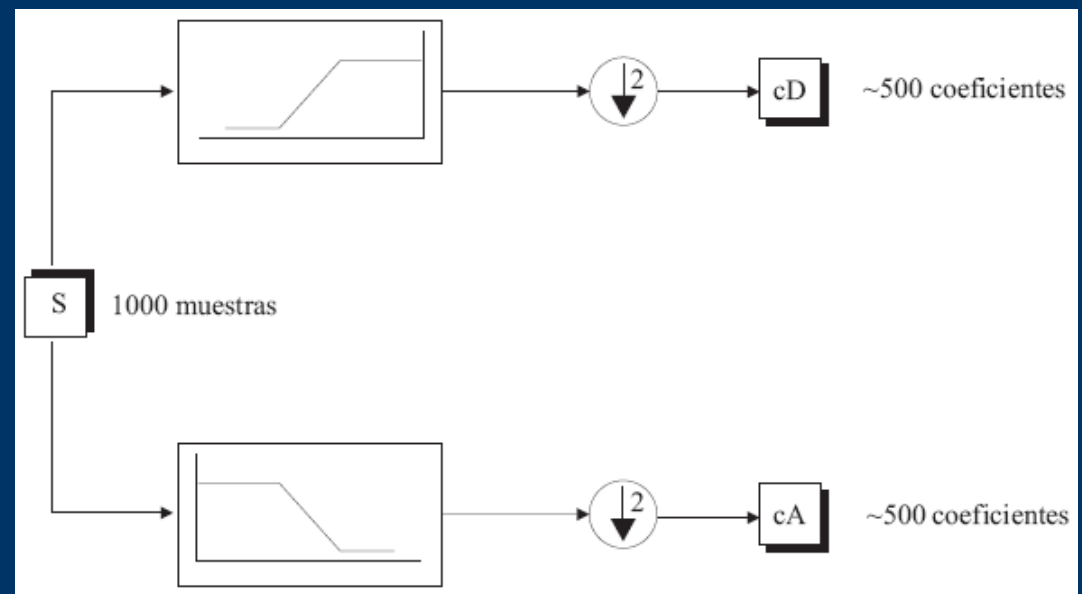


## Procesamiento digital

- La caracterización de la señal de voz se efectúa mediante los coeficientes de Aproximación (cA) en una escala que resultan del paso de la señal a través de un filtro pasabajas y un posterior submuestreo por dos.



- Filtrado simple



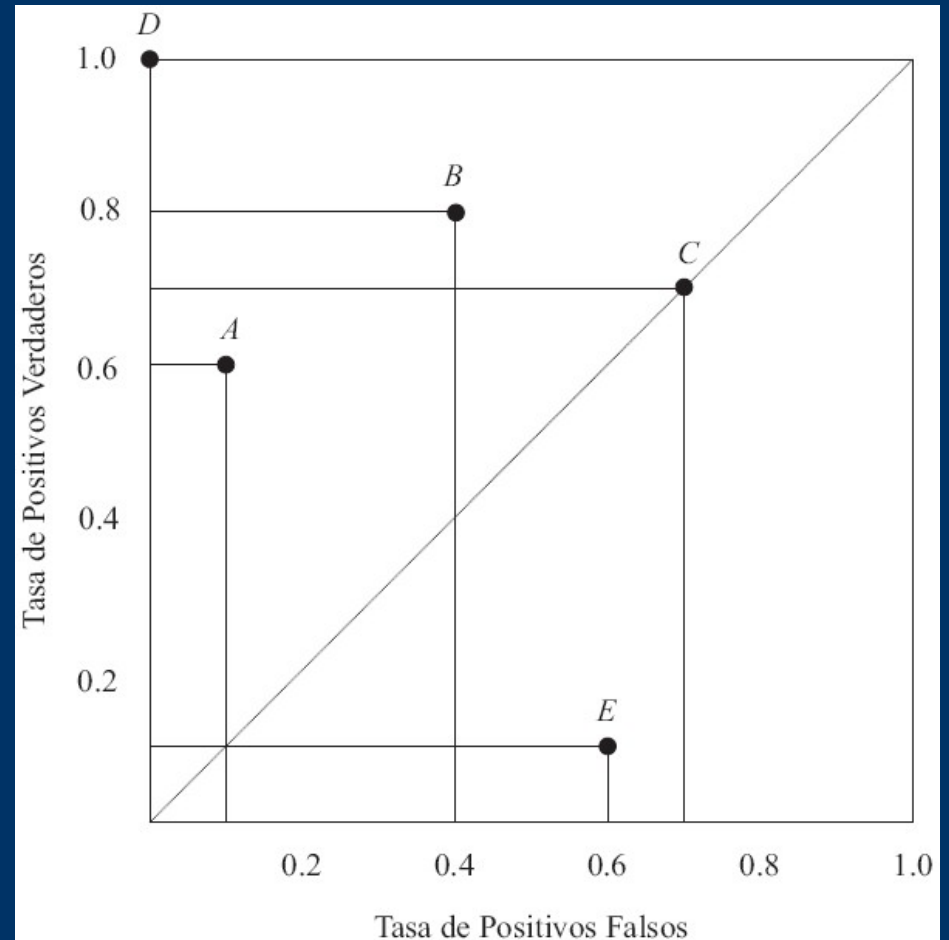
- Filtrado con submuestreo

## Experimentación

- El espacio ROC se forma con puntos en las intersecciones de dos métricas.

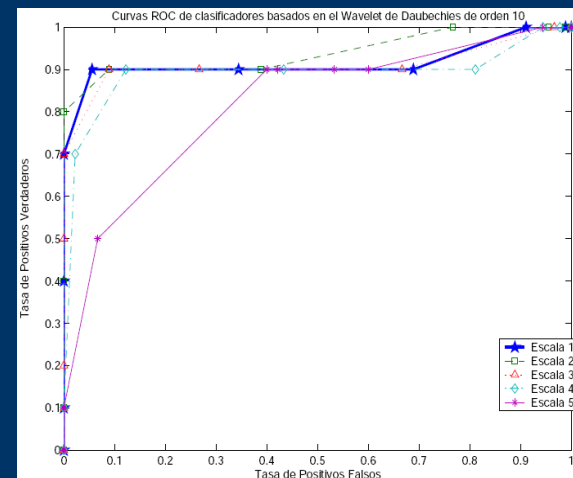
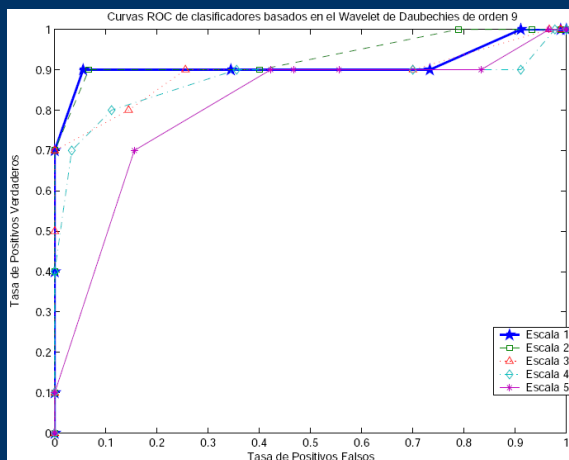
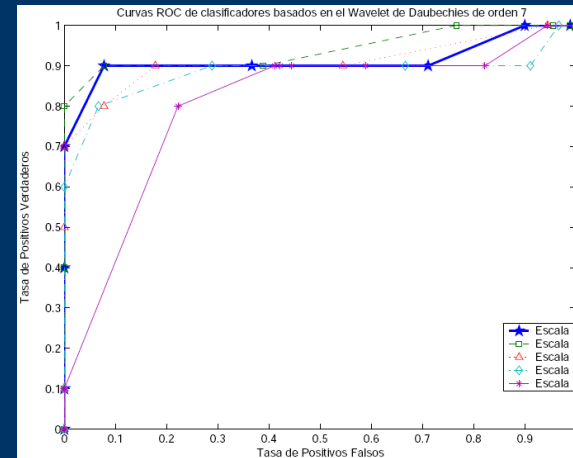
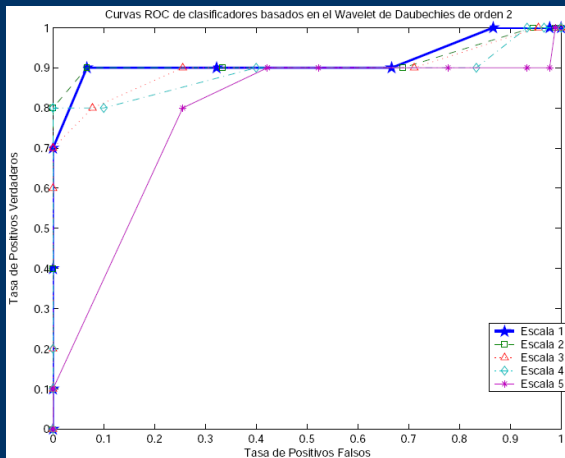
$$\text{tasa de positivos verdaderos} \approx \frac{\text{Positivos clasificados correctamente}}{\text{Positivos en total}}$$

$$\text{tasa de positivos falsos} \approx \frac{\text{Negativos clasificados incorrectamente}}{\text{Negativos en total}}$$

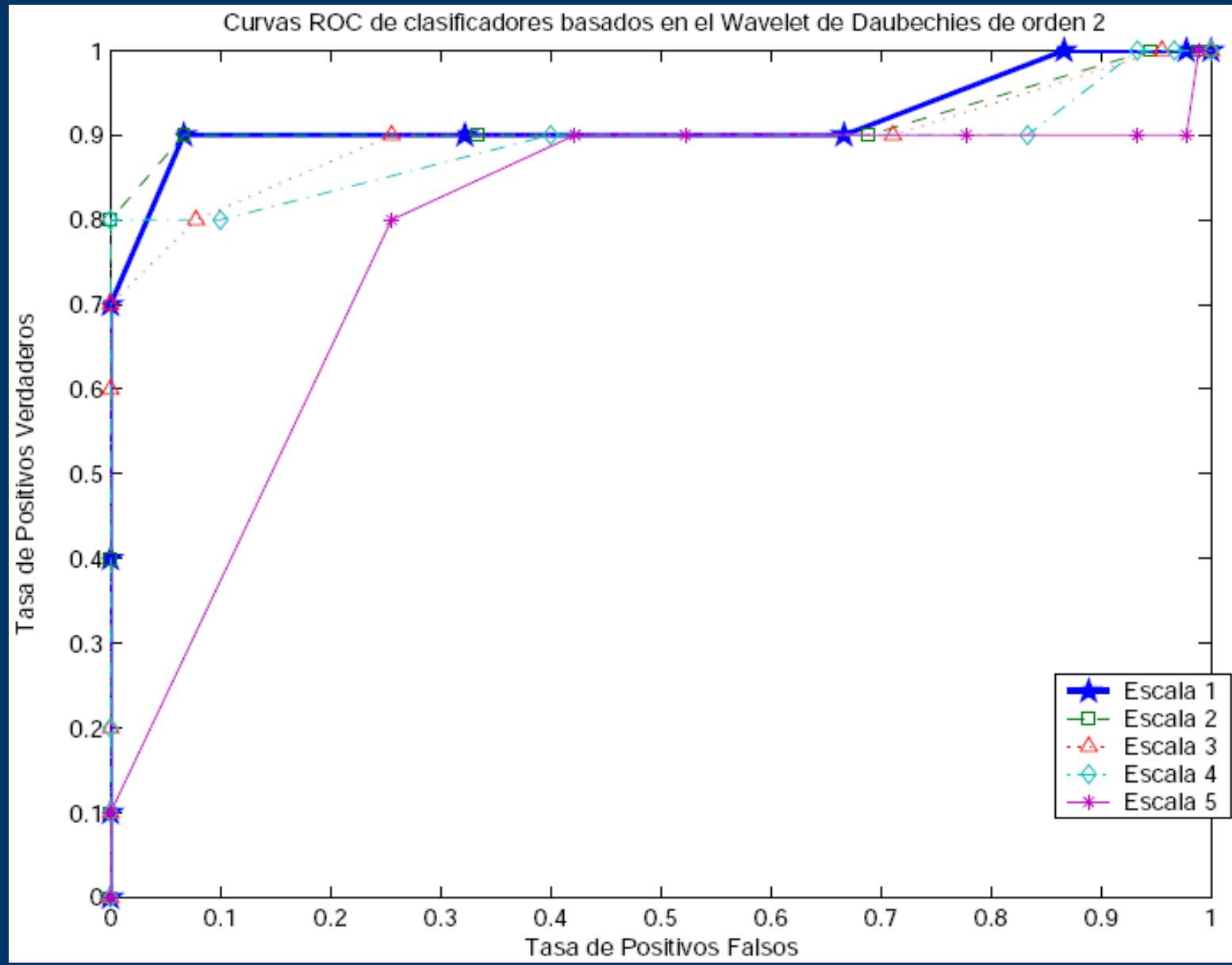


## Resultados

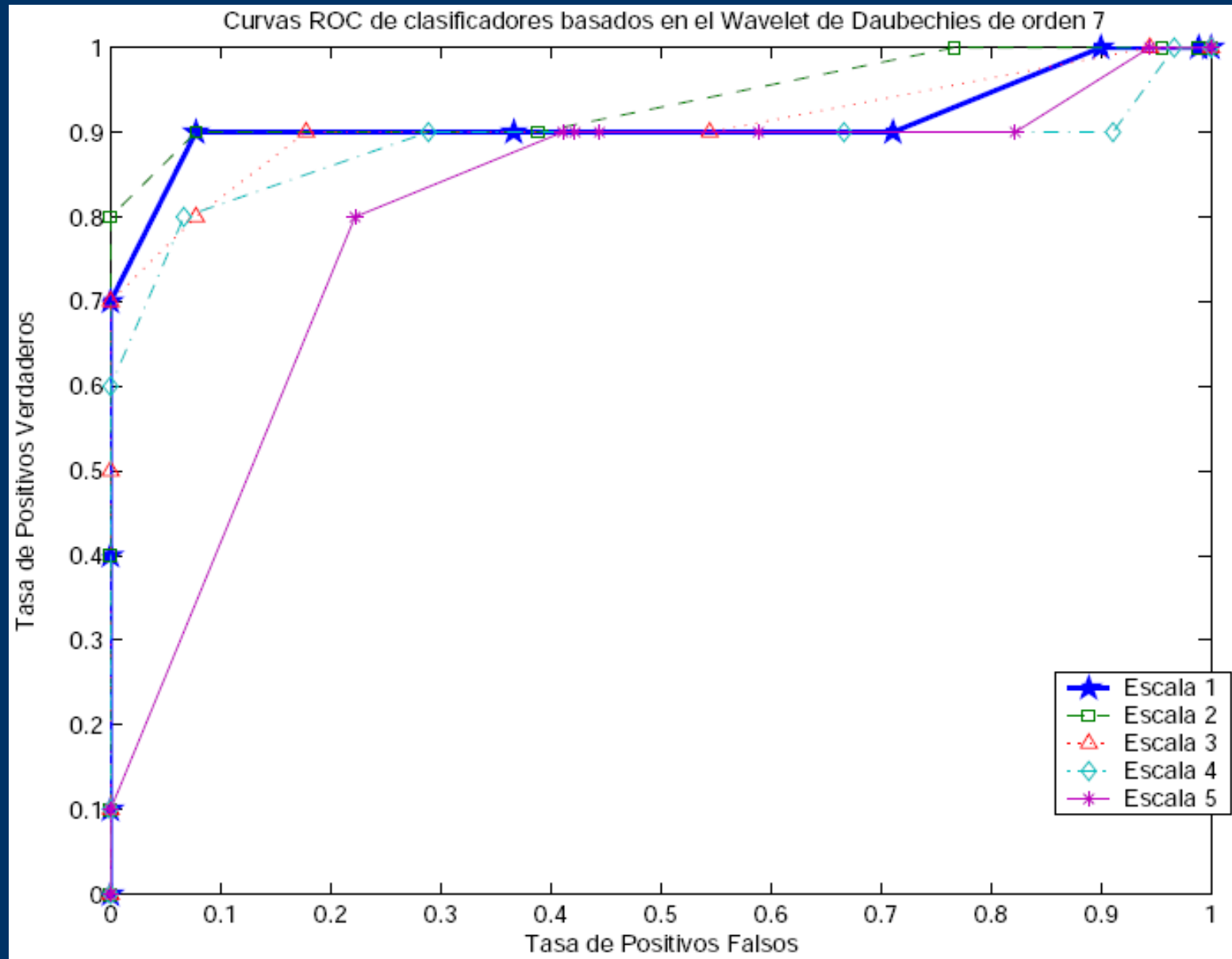
- Las curvas ROC para todos los clasificadores delimitan superficies mayores a 0.5, lo cual denota sensibilidad por parte del clasificador.



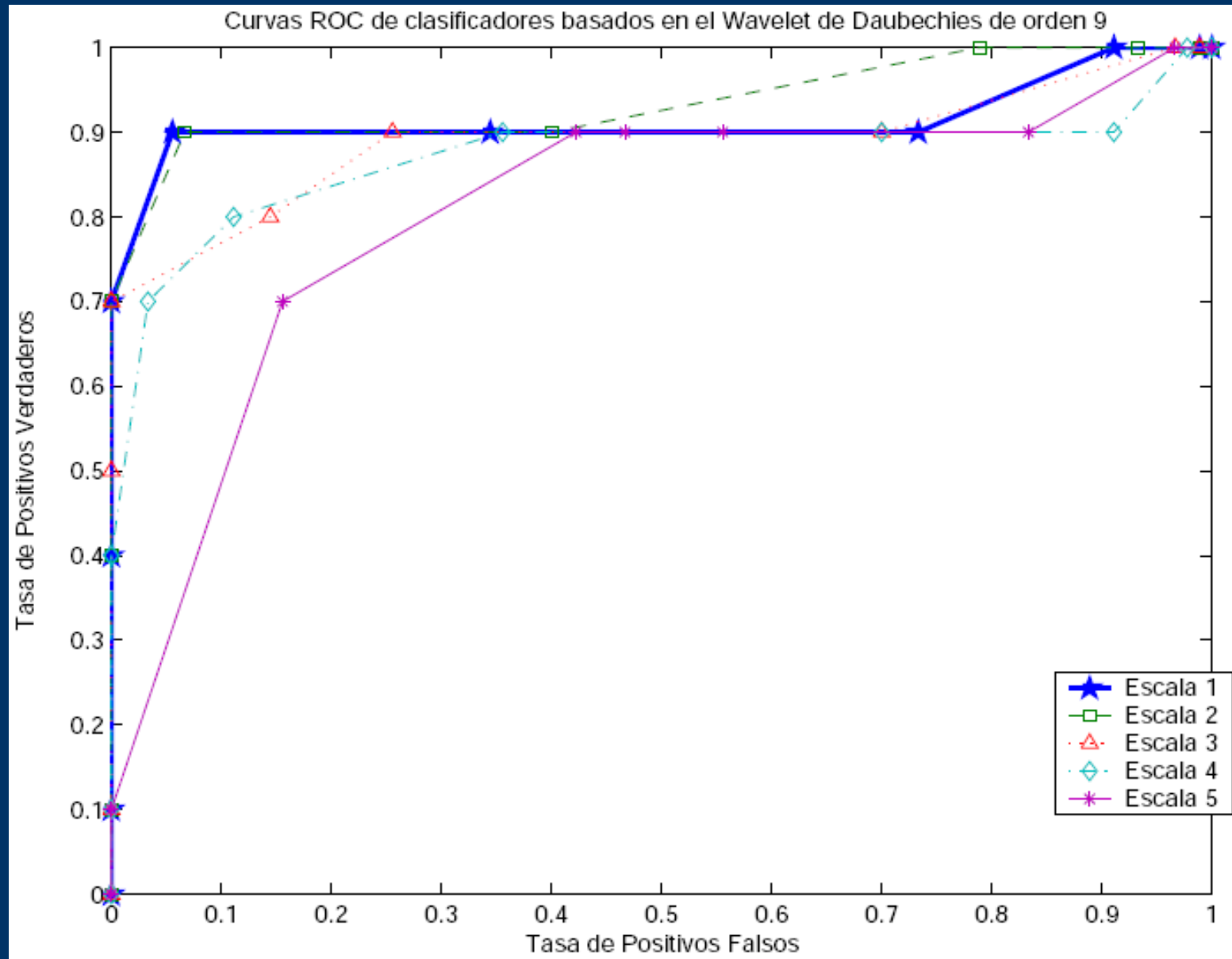
# Resultados



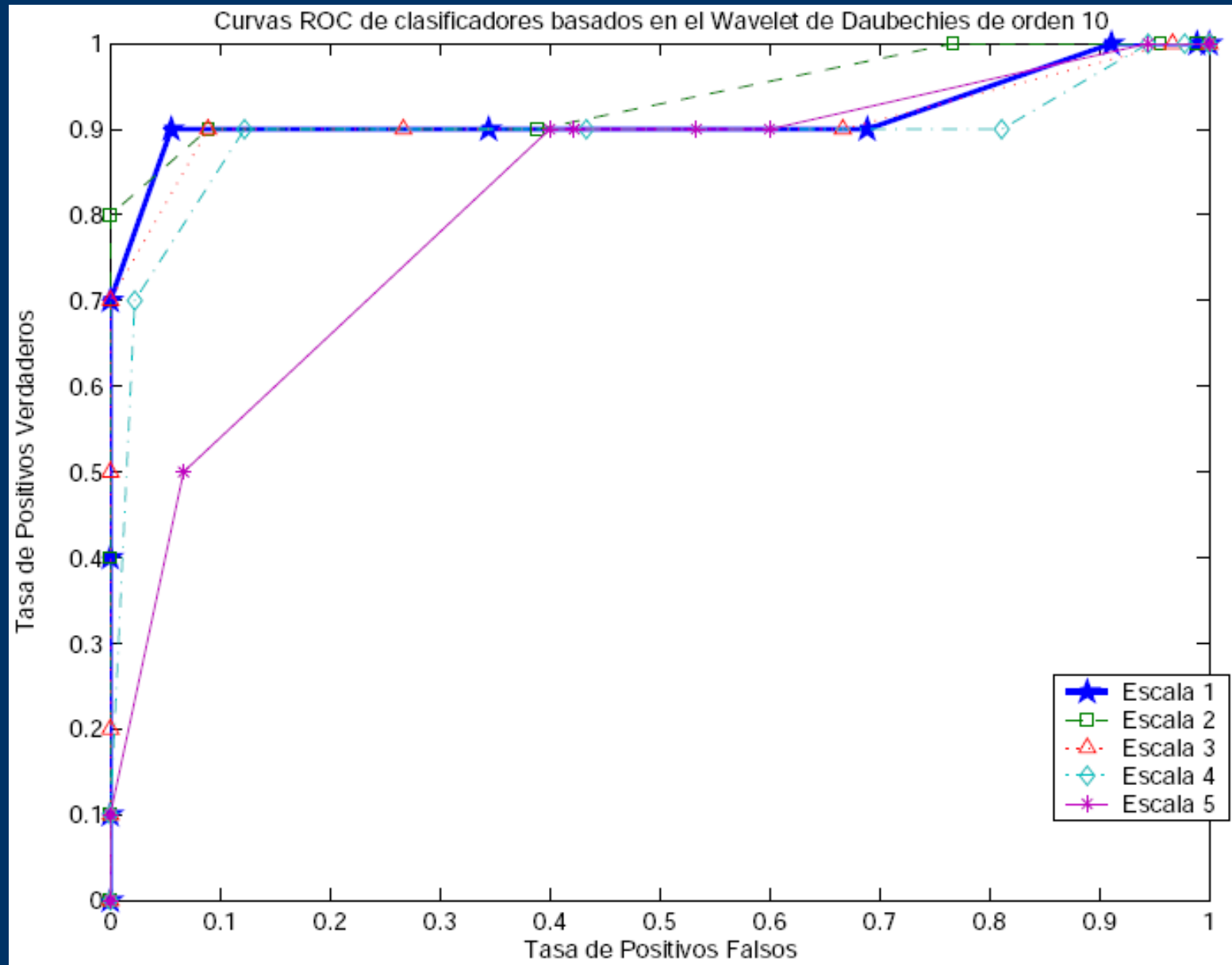
# Resultados



# Resultados

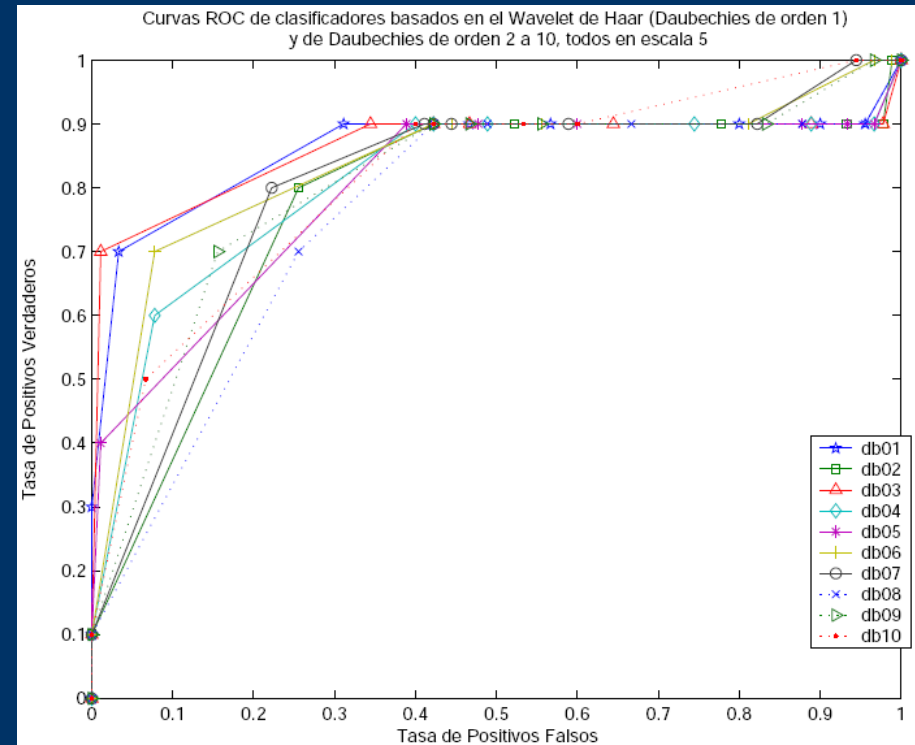
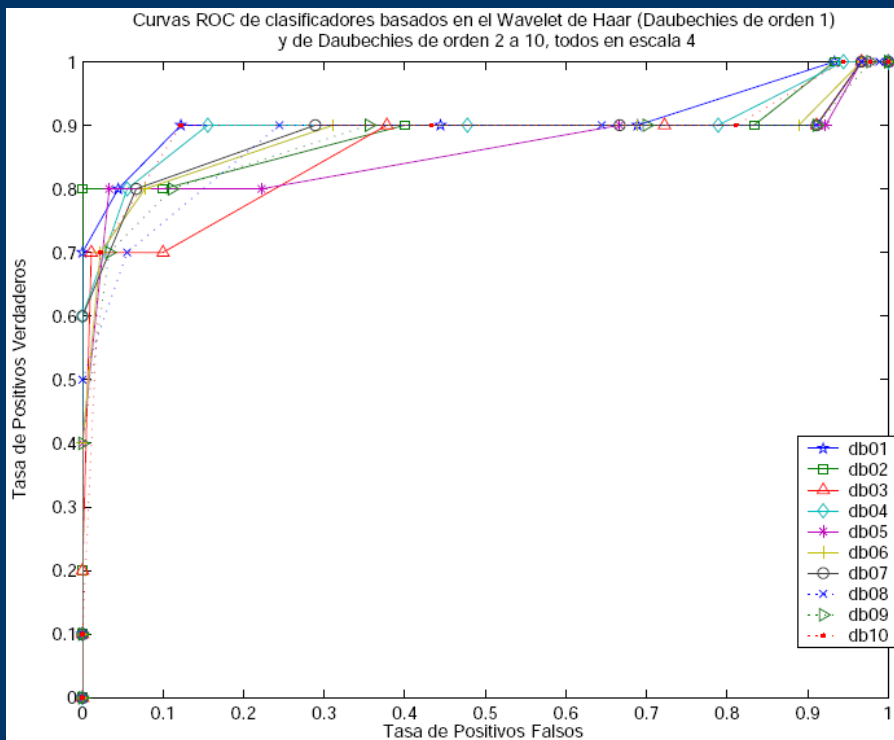


# Resultados



## Resultados

- En un aspecto general el rendimiento de un clasificador a otro decae cuando el Wavelet empleado se sustituye por uno de escala superior, aún cuando algunos tipos de Wavelets podrían incluso hacer mejor al clasificador al elevar su escala.



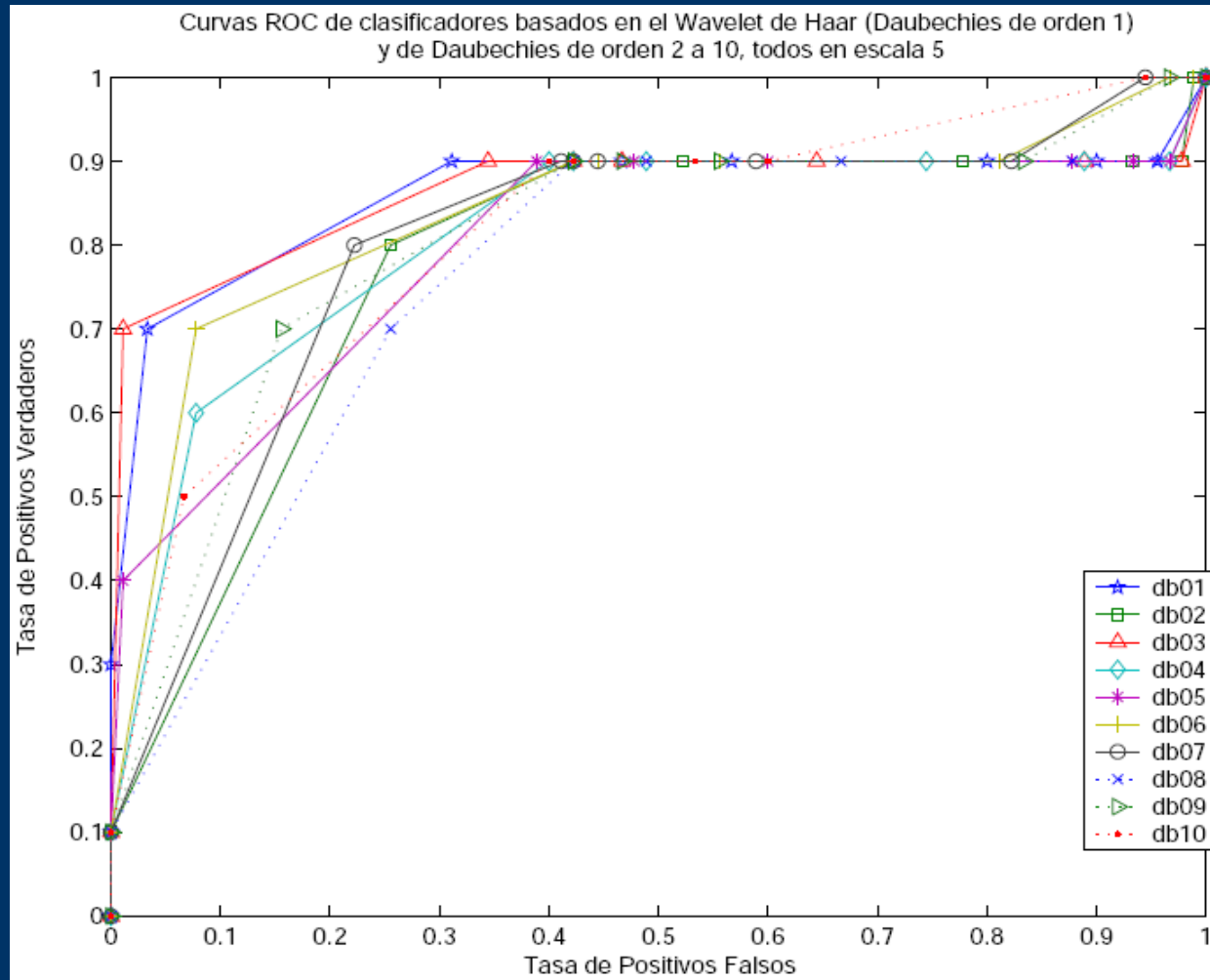






## Resultados

- El sistema con mejor comportamiento, de entre los construidos con Wavelets en escala 5, es el que emplea el Wavelet de Daubechies de tercer orden.





## Conclusiones

- El acercamiento a la esquina superior izquierda del espacio ROC de las curvas para todos los 50 clasificadores considerados en este proyecto (10 diferentes Wavelets, cada uno aplicado en 5 niveles de escala) demuestran que cualquiera de ellos funciona de manera aceptable.
- De entre los Wavelets explorados, el que permite la construcción del mejor sistema de reconocimiento de palabras aisladas con una transformación de escala 5 es el de Daubechies de orden 3 (db03).
- De entre los Wavelets explorados, el que permite la construcción del mejor sistema de reconocimiento de palabras aisladas con una transformación de escala 4 es el de Haar(db01 ).

*Gracias*

