

Conversión de texto a voz

Iván López Espejo

Generación de difonemas

► Generamos todos los difonemas posibles mediante la combinación de fonemas en castellano:

► *(a,b,d,e,f,g,i,j,k,l,m,n,ñ,o,p,r,R (como rr),s,t,u,w,x (como ch),y (como ll),z)*

```
► % Código para la generación de todos los difonemas posibles.  
► % Vector de fonemas.  
► letras = ['a','b','d','e','f','g','i','j','k','l','m','n','ñ','o','p','r','R','s','t','u','w','x','y','z','S'];  
► % Generación de difonemas.  
► matrizDifonemas = [];  
► ind = 1;  
► for j = 1:1:length(letras)  
►     for k = 1:1:length(letras)  
►         if j ~= k  
►             matrizDifonemas(ind,1) = 'e';  
►             matrizDifonemas(ind,2) = letras(j);  
►             matrizDifonemas(ind,3) = letras(k);  
►             matrizDifonemas(ind,4) = 'a';  
►             ind = ind + 1;  
►         end  
►     end  
► end  
► % Muestra el resultado por pantalla.  
► char(matrizDifonemas)
```



Generación de difonemas

- ▶ Les agregamos al comienzo y al final una vocal (e y a en nuestro caso). Ejemplo:
- ▶ Para el difonema */fe/*

▶ **efea**

- ▶ Grabamos los difonemas con las vocales con ayuda de un metrónomo que marca cada **1s** y con una frecuencia de muestreo de **44100Hz**.
- ▶ Sólo tendremos que extraer todos los difonemas sabiendo que cada uno de ellos se encuentra en un bloque de **44100 muestras**.

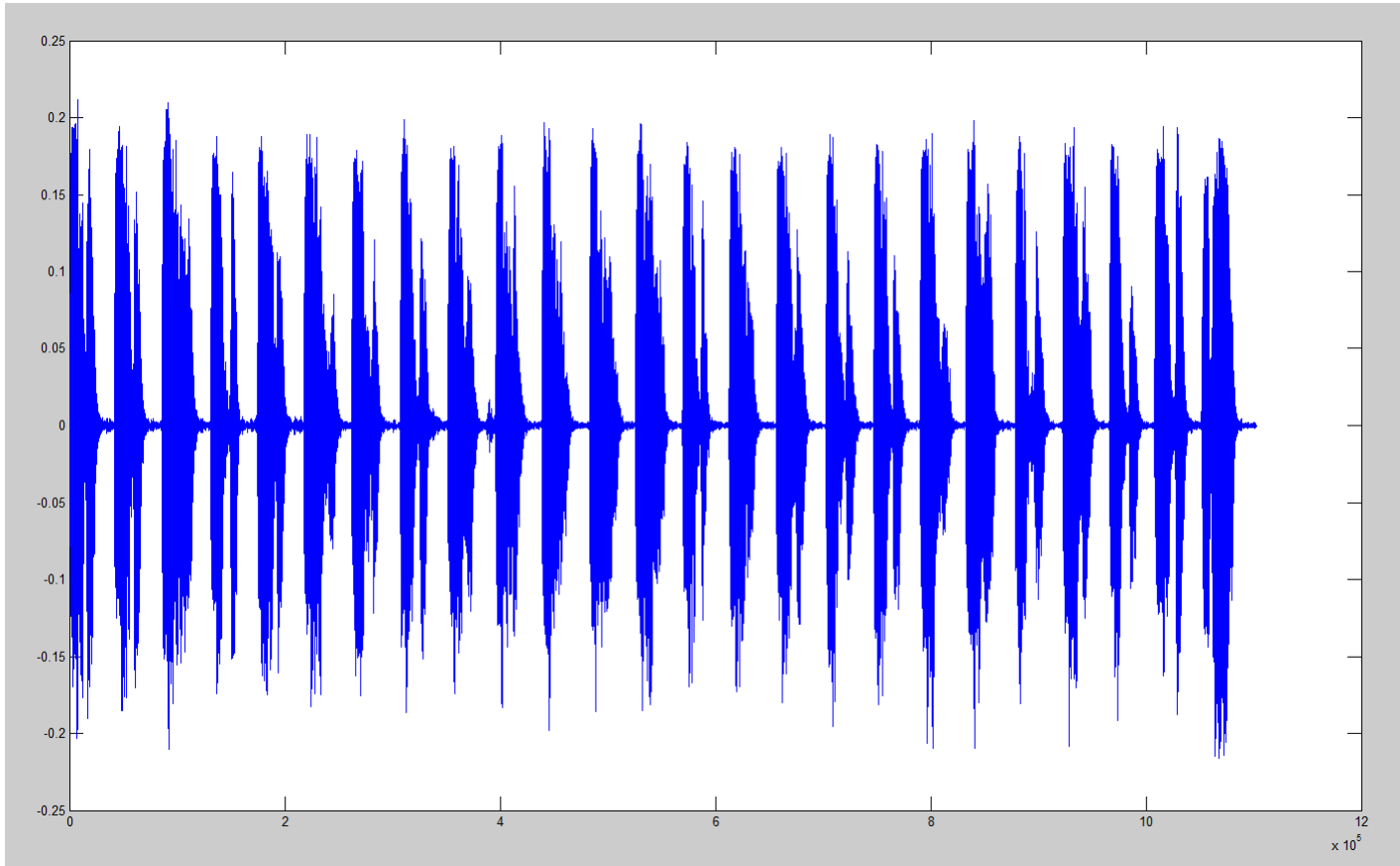


Extracción de difonemas

- ▶ Diferenciamos entre cuatro tipos de difonemas, según el primer y segundo fonema que lo componga:
 - ▶ Tipo 1 - Vocal a consonante (9625/14750)
 - ▶ Tipo 2 - Consonante a vocal (9500/15125)
 - ▶ Tipo 3 - Vocal a vocal (8500/16000)
 - ▶ Tipo 4 - Consonante a consonante (10000/15875)
- ▶ Observando varios ejemplos de cada tipo, establecemos en promedio el desplazamiento de muestras respecto del comienzo del bloque de 1s para cada uno de los tipos, a fin de extraer de forma automatizada el difonema, desde la parte estable del primero a la parte estable del segundo pasando por la transición.

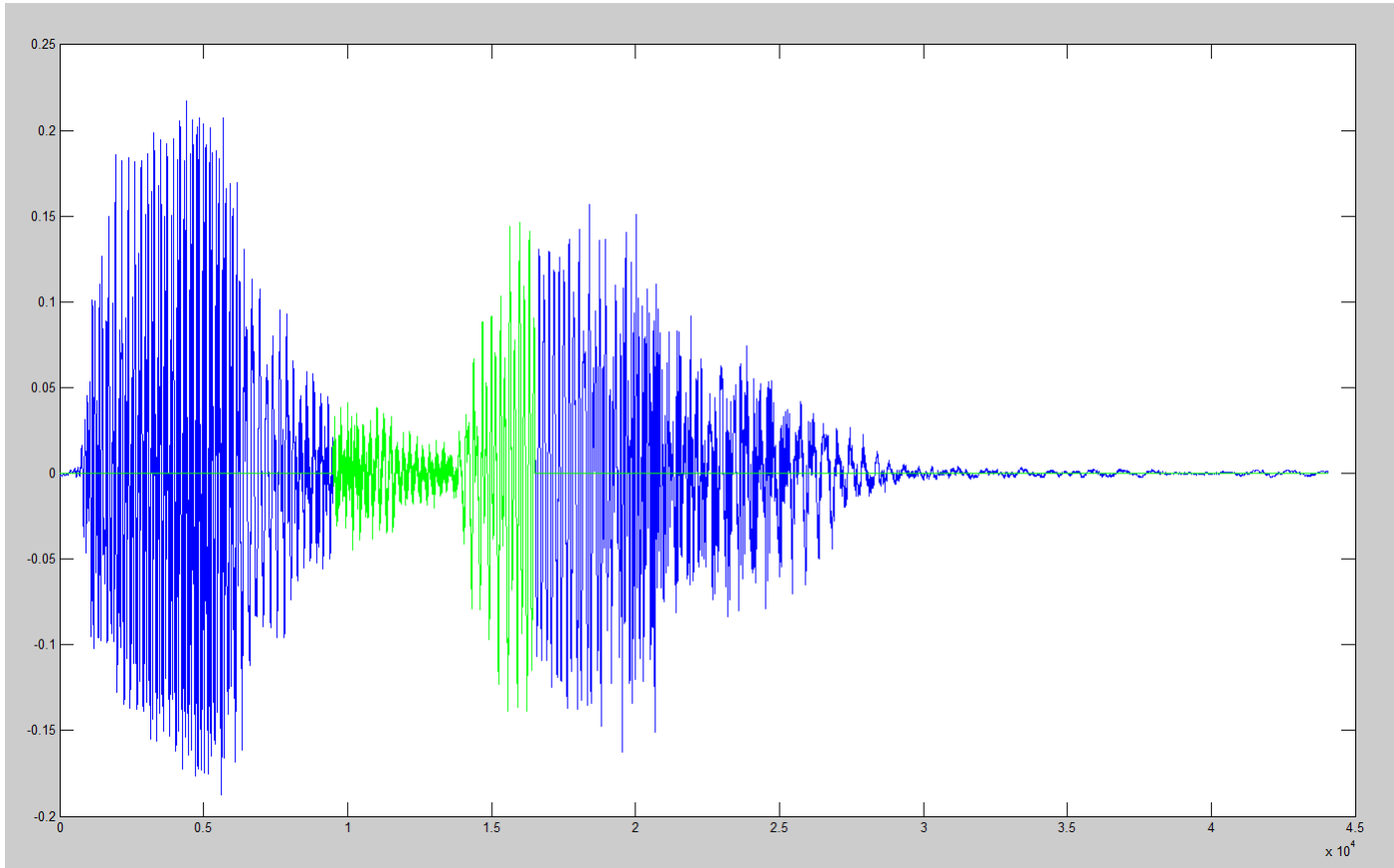
Extracción de difonemas

- ▶ Secuencia de difonemas que extraer



Extracción de difonemas

► Ejemplo: **efia**



Código para la segmentación

```
▶ function [y] = segDif(audio,numDif)
▶ difonemas = wavread(audio);
▶ difonemas = difonemas(:,1);
▶ for j = 1:1:numDif
▶     nombre = input('Introduzca el nombre del archivo tal y como será exportado: ');
▶     tipo = input('Introduzca el tipo de difonema: ');
▶     if tipo > 4 || tipo < 1
▶         disp('Error en el tipo de difonema. Las posibilidades son las siguientes: ');
▶         disp('1 -> Vocal/Consonante, 2 -> Consonante/Vocal, 3 -> Vocal/Vocal, 4 ->
Consonante/Consonante')
▶     else
▶         if tipo == 1
▶             wavwrite(difonemas((j-1)*44100+9625:(j-1)*44100+14750),44100,32,nombre);
▶         end
▶         if tipo == 2
▶             wavwrite(difonemas((j-1)*44100+9500:(j-1)*44100+15125),44100,32,nombre);
▶         end
▶         if tipo == 3
▶             wavwrite(difonemas((j-1)*44100+8500:(j-1)*44100+16000),44100,32,nombre);
▶         end
▶         if tipo == 4
▶             wavwrite(difonemas((j-1)*44100+10000:(j-1)*44100+15875),44100,32,nombre);
▶         end
▶     end
▶ end
▶ end
```



Conversión de texto a voz

- ▶ La siguiente función, muy limitada (no permite la repetición de caracteres consecutivos y sólo admite caracteres fonéticos y no texto escrito normal) comprueba la segmentación:

```
▶ function vector = text2voice(texto)
▶ texto = [' ' texto ' '];
▶ difonema = [];
▶ vector = [];
▶ for j = 1:1:length(texto)-1
▶     difonema = texto(j:j+1);
▶     if texto(j) == ' '
▶         difonema = ['SIL' texto(j+1)];
▶     end
▶     if texto(j+1) == ' '
▶         difonema = [texto(j) 'SIL'];
▶     end
▶     vector = [vector; wavread(difonema)];
▶ end
▶ soundsc(vector, 44100);
```



Ejemplos

- ▶ “Calculadora.”
 - ▶ “Explícamelo, por favor.”
 - ▶ “Impedancia equivalente.”
- ▶ “Esto es una prueba de una frase más larga.”

