

## **Práctica 3**

# **CONSTRUCCIÓN DE UN VOCODER CON CELP Y COMPARACIÓN CON LP**



**Fecha de Entrega : 06/11/2020**

# Índice

- Descripción
- Objetivo
- Resolución
- Conclusiones

## Descripción

La práctica trata del segundo tema de la asignatura, “voz y audio”.

Esta vez, vamos a utilizar un vocoder Celp, estudiado en clase sobre, para analizar la señal de voz utilizada en la práctica anterior. Estudiaremos distintos casos para comprobar cómo funciona este vocoder dependiendo del valor de unos parámetros y finalmente vamos a comparar los resultados con los experimentados en la práctica anterior correspondientes al uso de un vocoder lpc.

En este informe se reflejan los comentarios sobre el código llevado a cabo en los scripts de matlab y se adjuntan imágenes para comprender y analizar los resultados.

## Objetivo

Diseñar un vocoder celp y realizar tanto el análisis como la síntesis de una frase corta. Extraer conclusiones y comparar resultados con un vocoder lpc.

## Resolución

Vamos a utilizar la frase que empleamos en la práctica anterior, [frasecorta.wav](#).

Dividimos esta parte de resolución en tres subapartados. En el primero vamos a realizar los puntos a), b) y c) de la práctica, donde se pide llevar a cabo el estudio de varios casos de análisis celp variando los siguientes parámetros:

- Codebooks (**cb**)
- Orden del predictor lineal (**M**)
- Pesos del error (**c**)

En el segundo subapartado vamos a estudiar que sucede al variar la velocidad de transmisión de los casos, comprobando la calidad del vocoder.

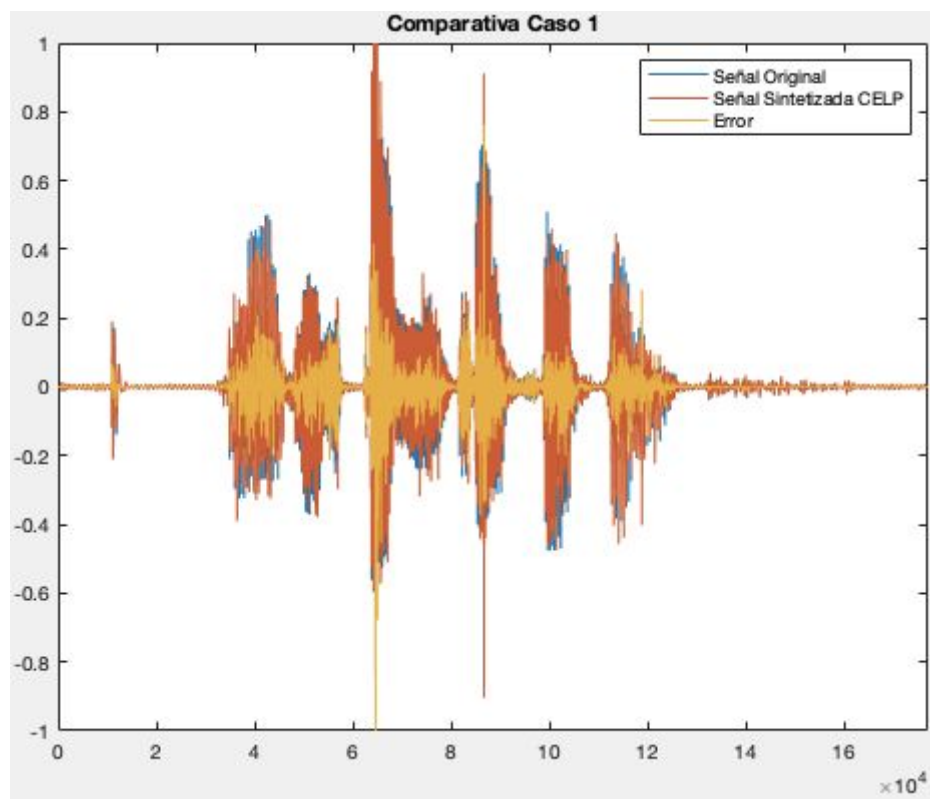
Por último, se va a comparar el vocoder celp con el vocoder lpc que utilizamos en la práctica anterior.

## 1. Variación de parámetros

### 1.1 Variamos los codebooks

#### Caso 1 - codebook = 40 veces 1024

$N=160$ ;  $M=10$ ;  $L=40$ ;  $c=0.8$ ;  $Pidx=[16 \ 160]$ ;

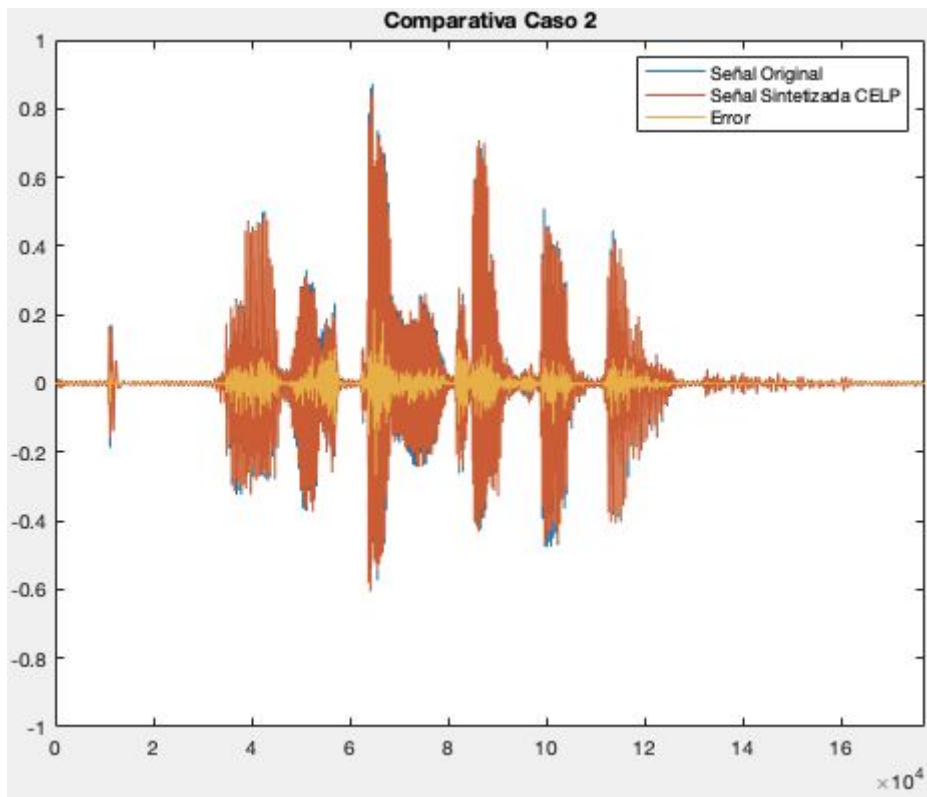


En este caso existe un error que no podemos considerar despreciable. Al escuchar la señal se escucha bastante peor que la original.

#### Caso 2 - codebook = 20 veces 1024

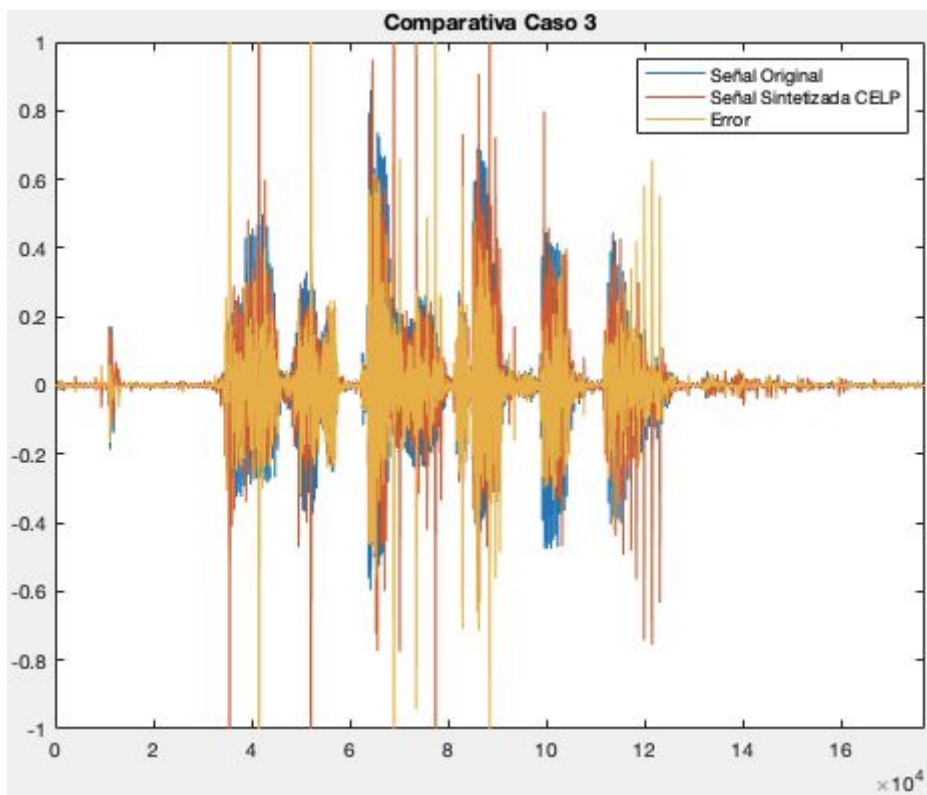
$N=160$ ;  $M=10$ ;  $L=40$ ;  $c=0.8$ ;  $Pidx=[16 \ 160]$ ;

Al reducir el valor del cb, veremos en la gráfica como la señal sintetizada es prácticamente igual a la señal original y por lo tanto el error en este caso es mínimo. Esto nos dice que cuanto menor sea el valor de la  $L$ , es decir, el tamaño de cada subframe, mejor se va a llevar a cabo la sintetización de la señal original. Al escuchar tanto **f<sub>n</sub>** como **x<sub>hat</sub>**(sintetizada) apenas se aprecia diferencia, como era de esperar.



### Caso 3 - codebook = 80 veces 1024

N=160; M=10; L=40; c=0.8; Pidx=[16 160];



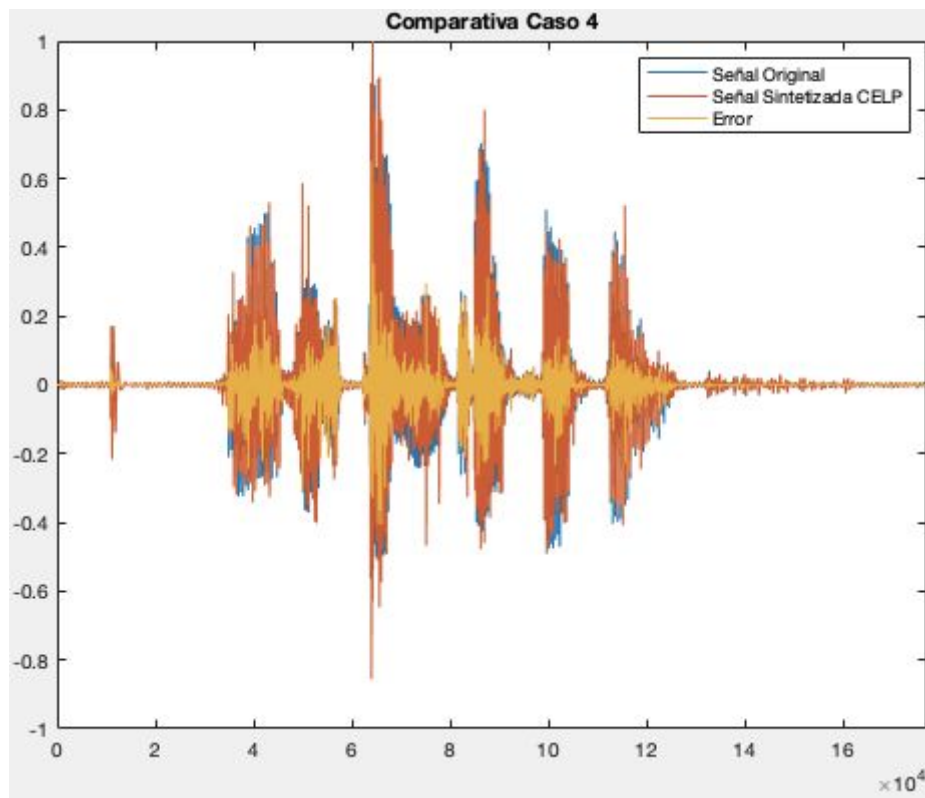
Como podemos apreciar en la imagen, el error es bastante grande, debido al aumento del codebook hasta  $80 \times 1024$ . Al escuchar las señales se aprecia perfectamente este error.

## 1.2 Variamos los el orden del predictor lineal

Ahora vamos a variar M. Utilizamos el caso 1, en el que M era 10 como punto de partida, y cambiamos su valor a 14, 20 y 24. Cabe mencionar que en esta parte del estudio vamos a mantener el valor del cb constante como  $40 \times 1024$

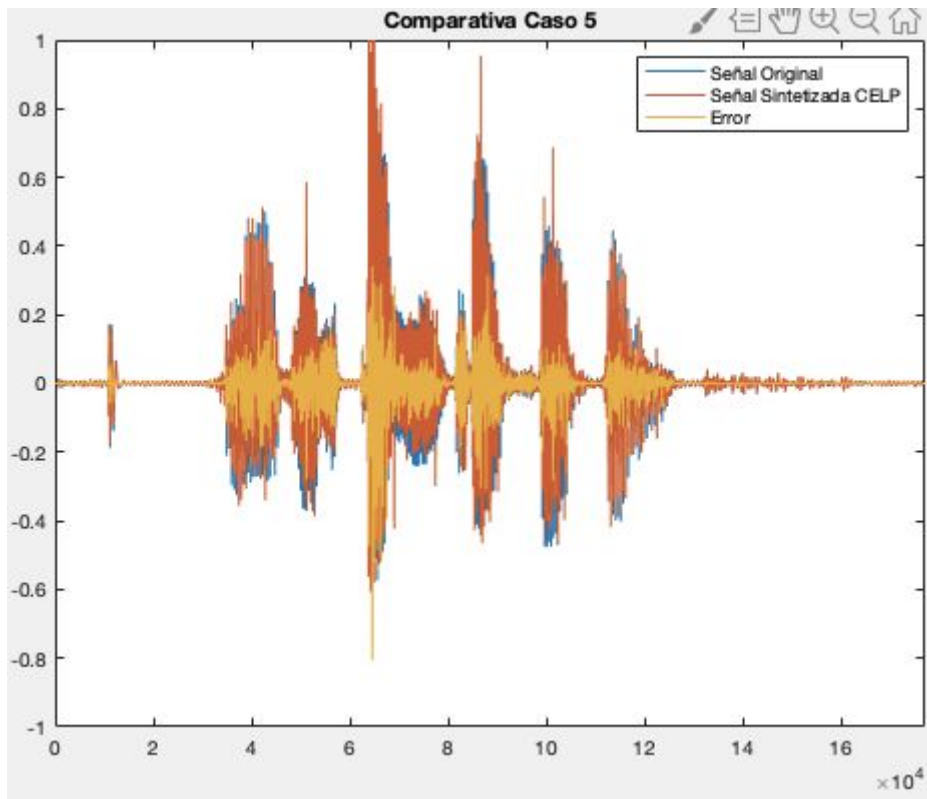
### Caso 4 - M = 14

N=160; M=14; L=40; c=0.8; Pidx=[16 160];



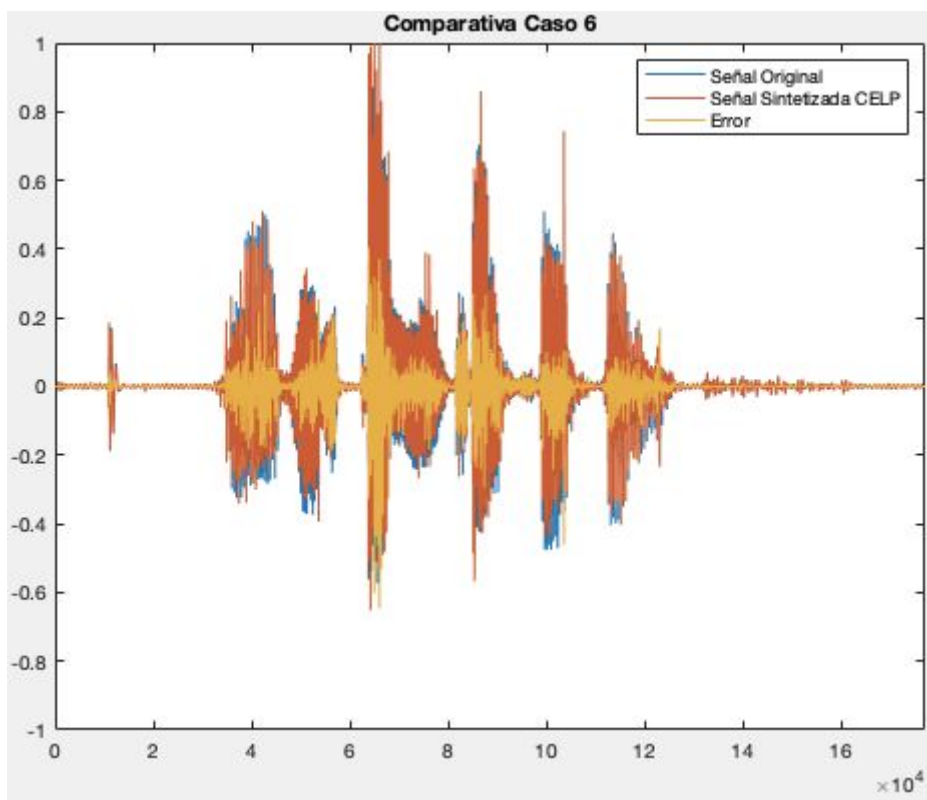
### Caso 5 - M = 20

N=160; M=20; L=40; c=0.8; Pidx=[16 160];



### Caso 6 - $M = 24$

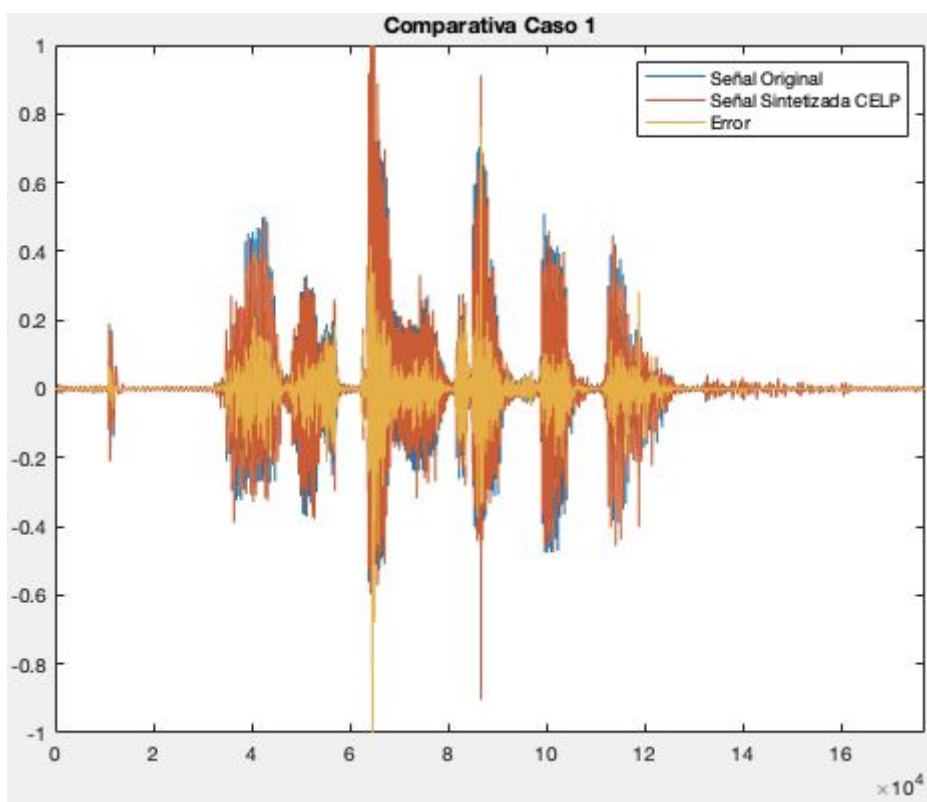
$N=160$ ;  $M=24$ ;  $L=40$ ;  $c=0.8$ ;  $Pidx=[16 \ 160]$ ;



Se aprecia que el mejor caso es cuando  $M=14$ . Hay menor error. Realmente lo que sucede es que en teoría el mejor caso es cuando  $M$  es mayor, es decir, es 20. Si hacemos la media de los errores con **mean** sale eso, pero es tan pequeña la diferencia entre casos, que no nos merece la pena seleccionar un  $M$  mayor cuando el beneficio va a ser apenas notable, y por eso 14 sería el mejor valor.

### 1.3 Variamos los pesos del error

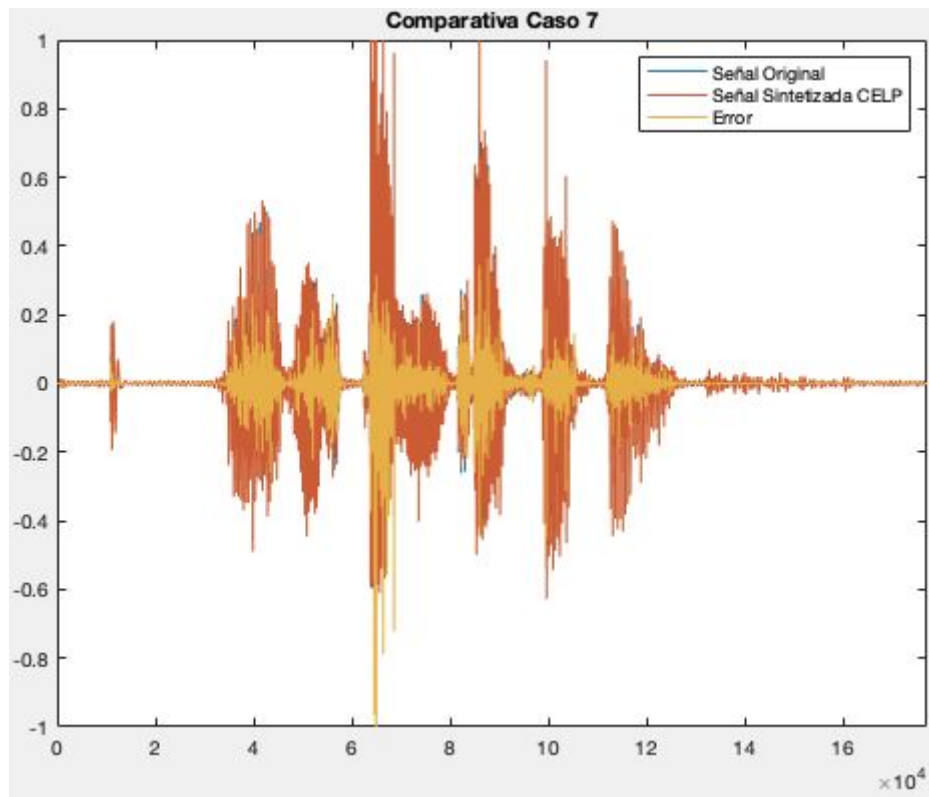
Ahora vamos a variar  $c$ . Utilizamos el caso 1, en el que  $c$  era 0.8 como punto de partida, y cambiamos su valor a 1 y 0.4. Cabe mencionar que en esta parte del estudio vamos a mantener el valor del  $cb$  constante como  $40 \times 1024$  y el valor de  $M$  como 10.



#### **Caso 7 - $c = 1$**

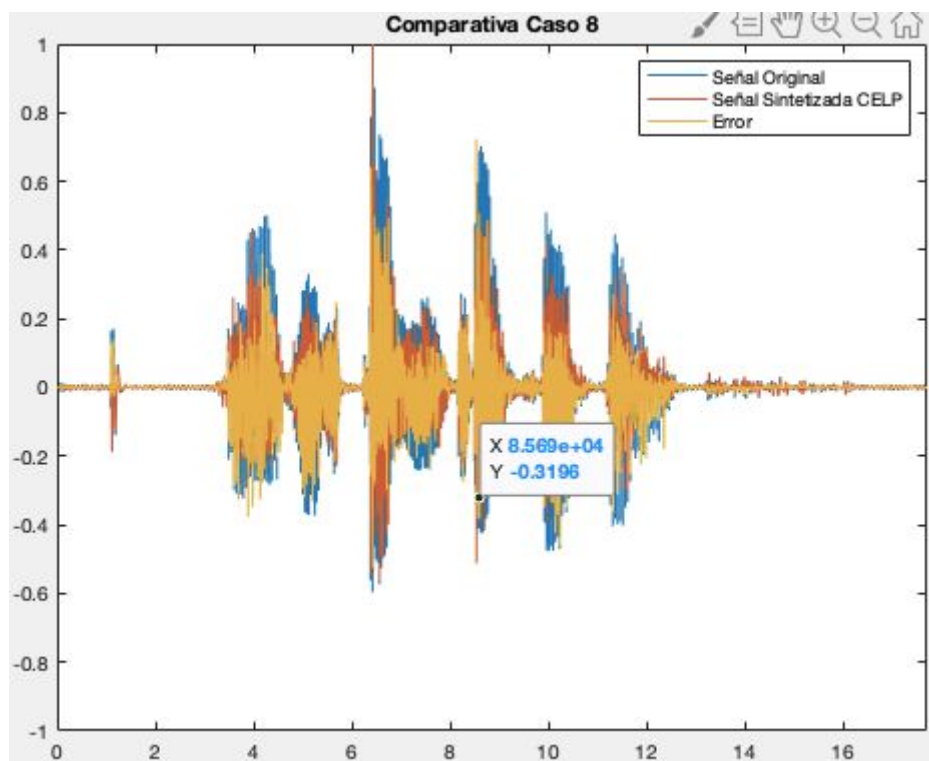
$N=160$ ;  $M=10$ ;  $L=40$ ;  $c=1$ ;  $Pidx=[16 \ 160]$ ;





### Caso 8 - $c = 0.4$

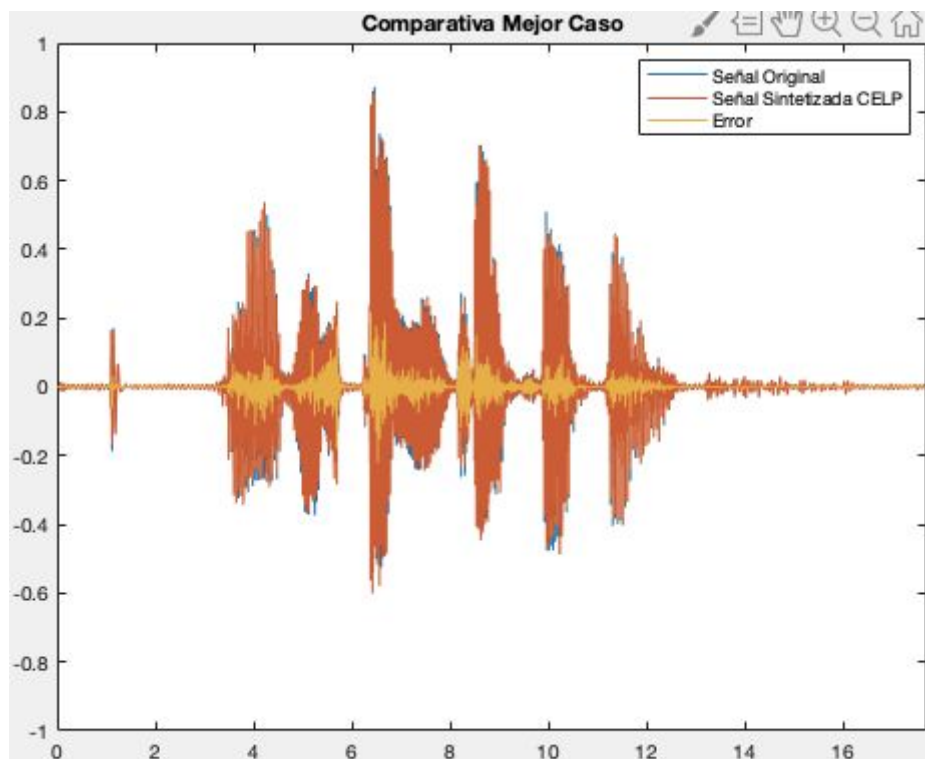
$N=160$ ;  $M=10$ ;  $L=40$ ;  $c=1$ ;  $\text{Pid}x=[16 \ 160]$ ;



Como vemos, el mejor caso es con  $c = 1$ . Podemos concluir por lo tanto que cuanto mayor sea el valor del peso, menos error va a aparecer y mejor va a resultar el análisis. Se corrobora escuchando las distintas señales.

#### 1.4 EXTRA. MEJOR CASO

Mejor caso -  $N=160$ ;  $M=20$ ;  $L=20$ ;  $c=0.8$ ;  $Pidx=[16 \ 160]$ ;



Con estos parámetros encontramos el mejor caso, y al escucharlo corroboramos que en efecto la señal sintetizada se escucha prácticamente igual que la original.

## 2. Variación en la velocidad de transmisión

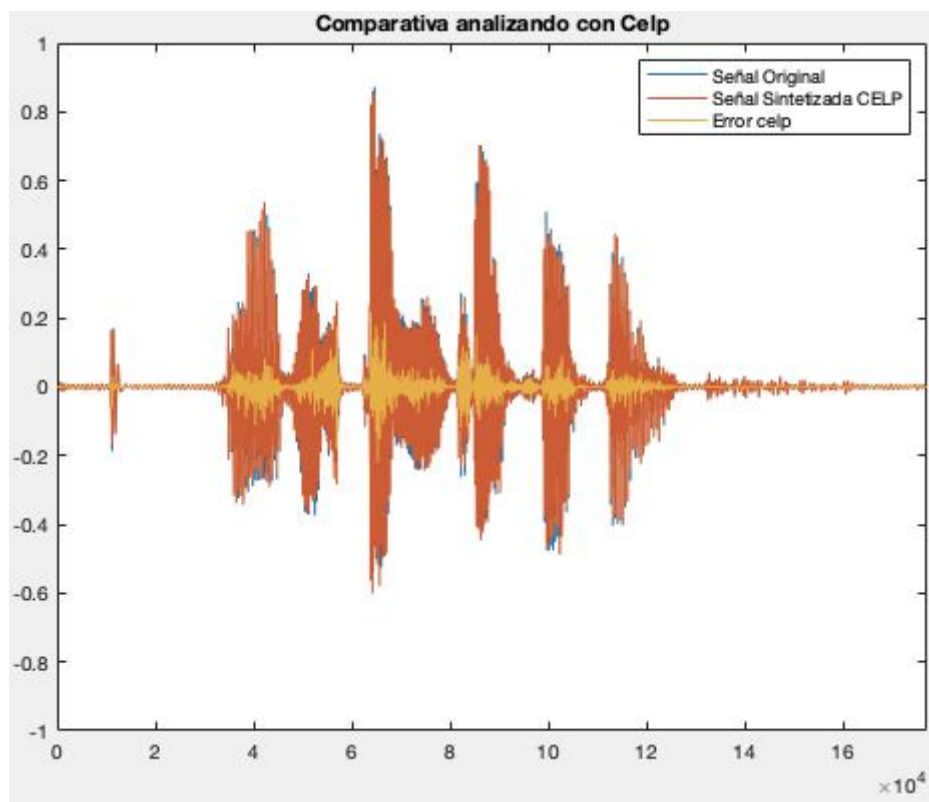
En este apartado vamos a usar las funciones **celp16k.m** y **celp9600.m** para cambiar la velocidad de transmisión de un caso.

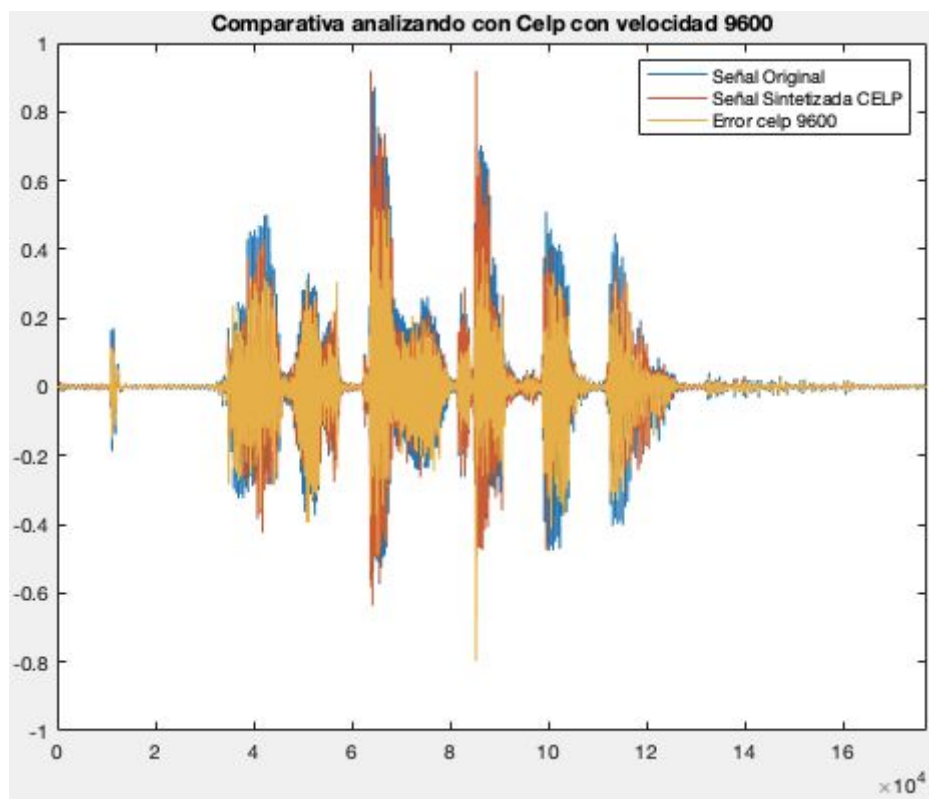
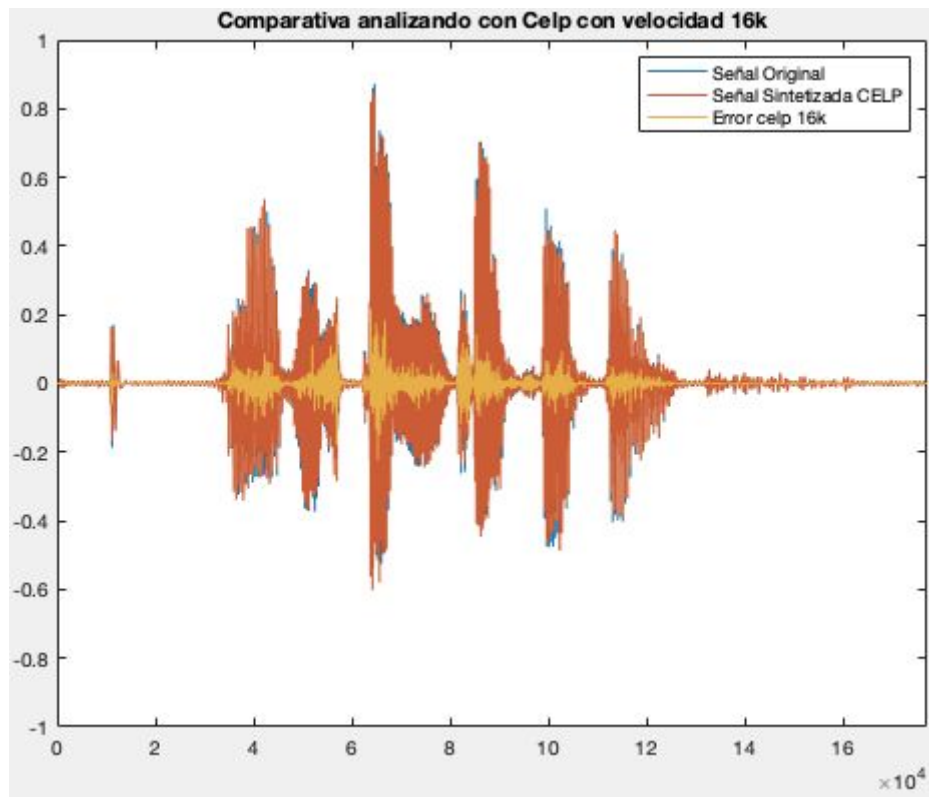
El caso elegido es el correspondiente al mejor caso del apartado anterior:

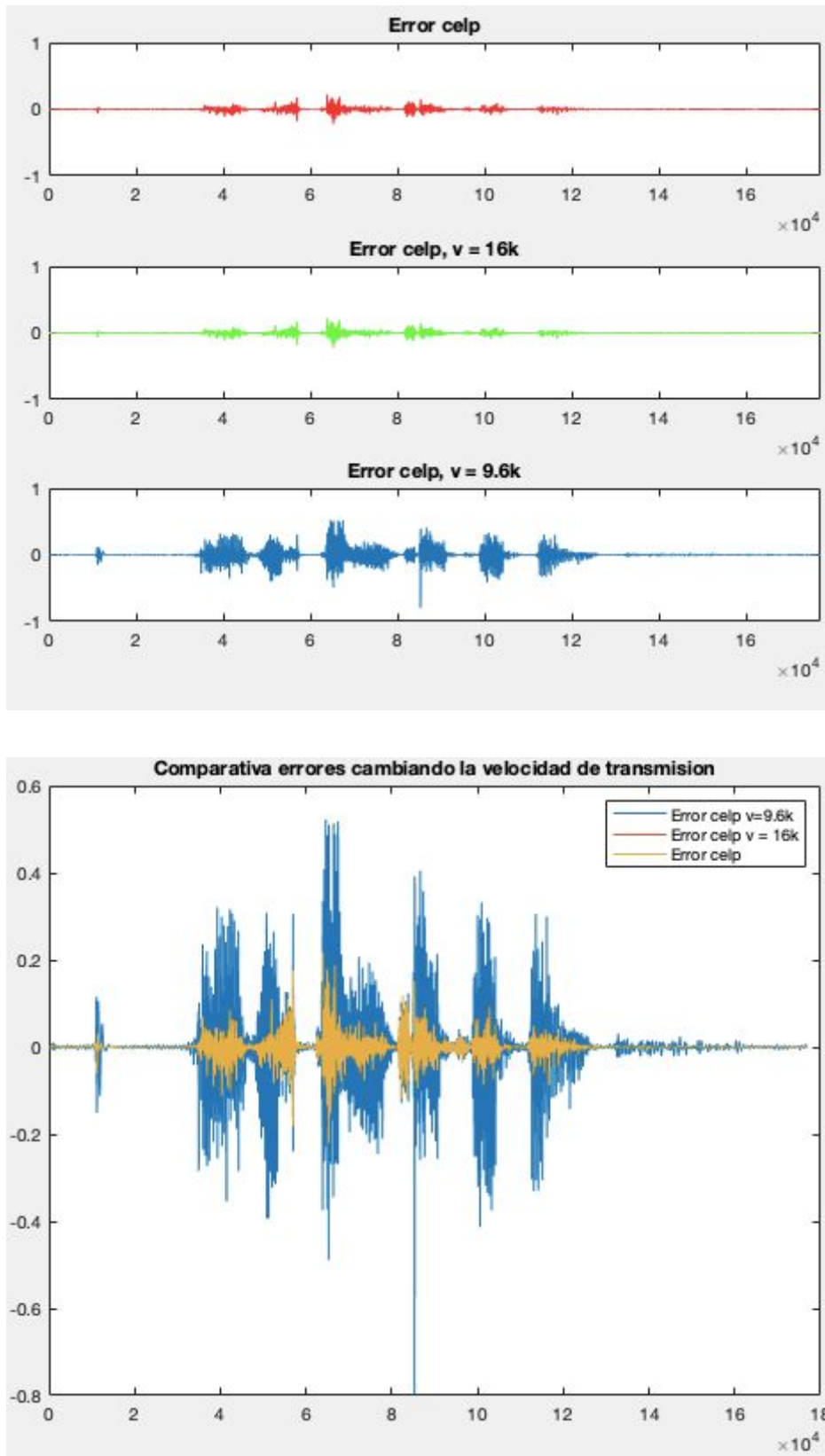
$cb = 20 \times 1024$

$N=160$ ;  $M=20$ ;  $L=20$ ;  $c=0.8$ ;  $Pidx=[16 \ 160]$

Vamos a mostrar el resultado del análisis con cada uno de los vocoders de distinta velocidad, y luego la comparación entre los tres errores.







Se puede apreciar claramente en todas las imágenes que cuando bajamos la velocidad a 9600 bps, se produce un error bastante grande en la síntesis de la señal, y que sin embargo entre los dos otros vocoders no se aprecia diferencia.

### 3. Celp vs LPC

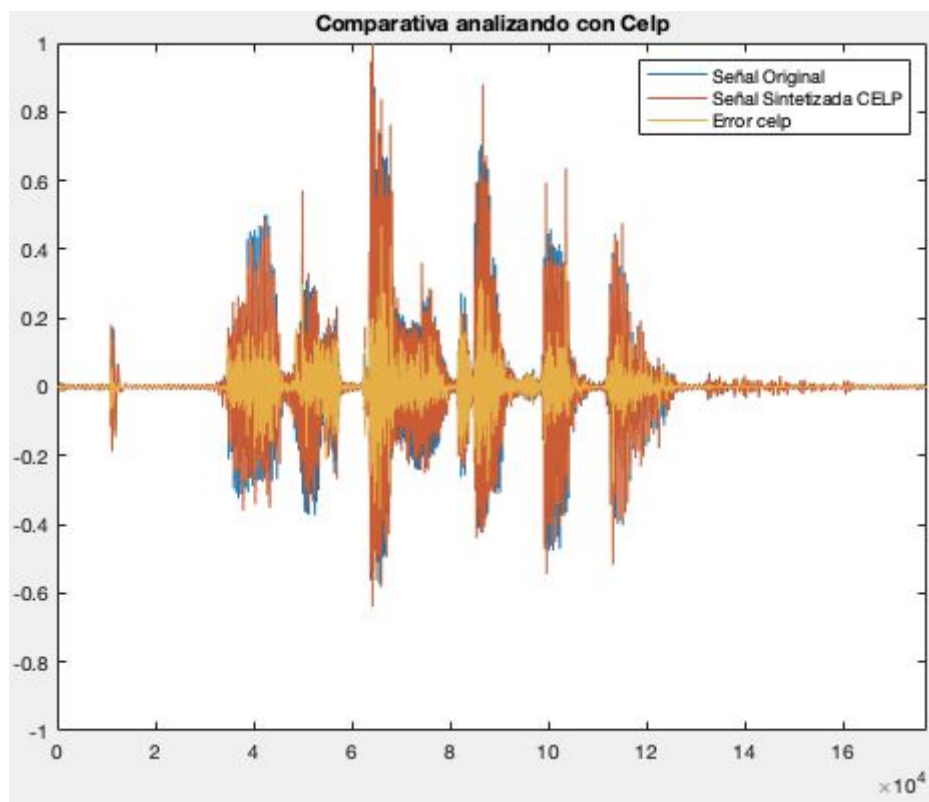
En este apartado vamos a reciclar el vocoder que utilizamos en la práctica 2 para analizar la frase corta con ambos vocoders: celp y lpc.

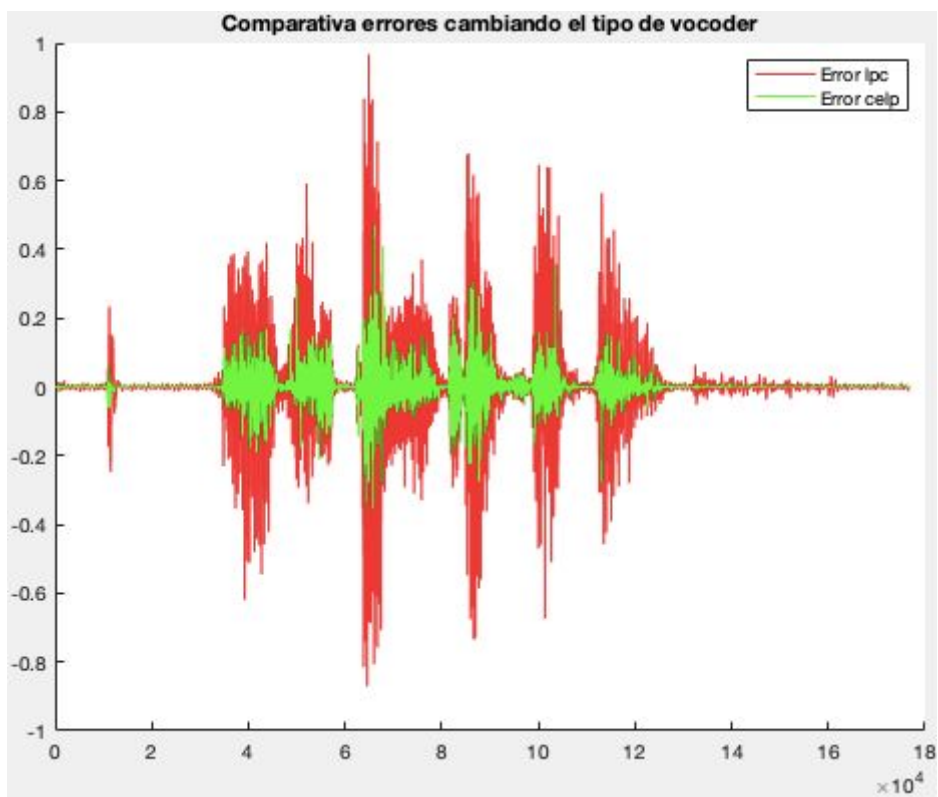
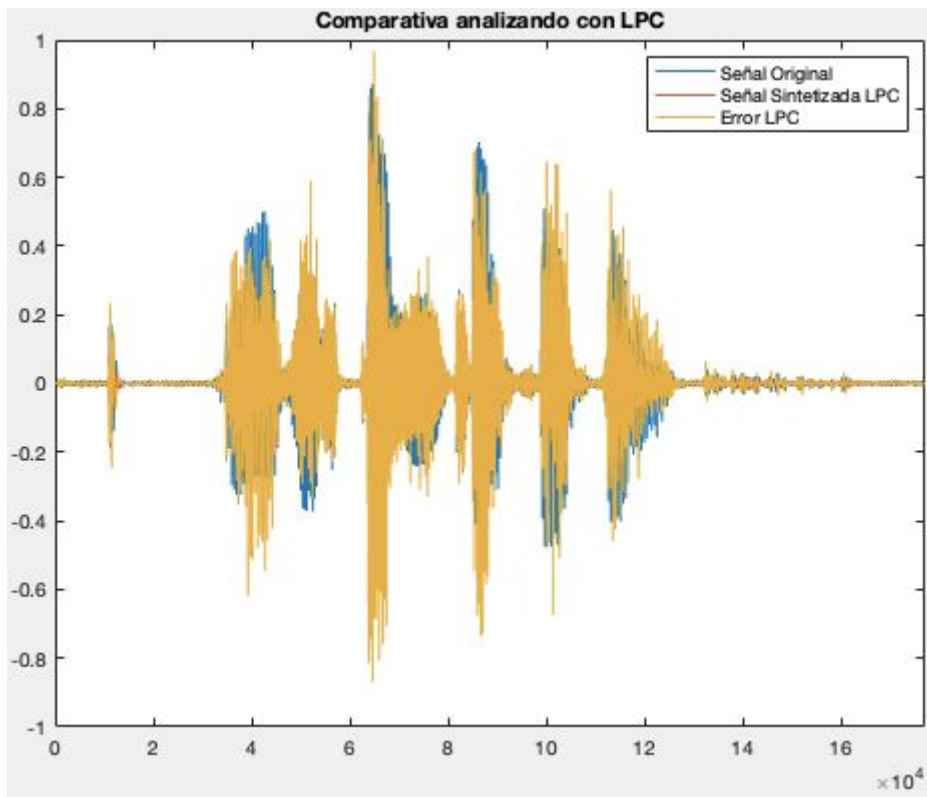
El caso **celp** elegido es:

Caso:  $N=160$ ;  $M=12$ ;  $L=40$ ;  $c=0.8$ ;  $\text{Pid}x=[16 \ 160]$ ;  $\text{cb} = 40 \times 1024$

Para **lpc** :  $N=256$ ;  $M = 12$ ;

El resultado es el siguiente:





Como podemos apreciar, el vocoder LP sintetiza peor nuestra señal de voz, obteniendo un error mayor. Es lo que tiene que pasar ya que celp suele ser mejor.

## Conclusiones

Me ha parecido una práctica muy interesante para acabar de entender el funcionamiento del vocoder LP, y comprender también cómo funciona el celp.

Ha sido una práctica sencilla ya que en todas las imágenes se comprueba fácilmente cual es el resultado, y en todos los casos sucede lo esperado y estudiado en clase.