

4. EXPERIMENTOS.

La finalidad de los experimentos es conseguir flujos de voz sintetizada con una tasa de bit menor que la de los flujos de voz usados como fuente, ya que de otra manera no conseguiríamos ninguna ventaja al comprimir (sufriríamos pérdida de calidad y aumento de tasa). Los flujos de voz propuestos que sirven de entrada a los codificadores son de tipo PCM (8000 muestras/s y 8 bits/muestra) con una tasa total de 64000 bits/s.

Antes de seguir tenemos que recalcar la dificultad de comparar codificadores teniendo en cuenta que presentan diferentes principios de análisis y síntesis, ya que no se centran en los mismos aspectos: unos tratan de reflejar la forma de onda de la señal y otros plantean un modelo de producción de voz que proporciona resultados apreciables a pesar de la disparidad entre la forma de onda sintetizada respecto a la original.

Para valorar cada codificador tendremos en cuenta varias medidas: tasa de salida, medida señal-ruido SNR, tiempo de computación y medida de percepción subjetiva MOS. Para fijar la tasas de salida de cada uno de los codificadores se ha partido de una cuantización de parámetros (asignación de bits) con un alto de número de bits, decreciendo éste hasta notar un salto cualitativo en la percepción de la voz. Cabe destacar que el codificador DCT presenta una tasa variable, ya que el número de parámetros seleccionados depende de cuántos coeficientes de la transformada superen el umbral. Esto depende exclusivamente de las características del fichero fuente. Vamos a puntualizar algunos aspectos de estas medidas:

1. Tasa de salida. Es una consecuencia directa de la asignación de bits de cada codificador.
2. SNR. Se trata de evaluar el parecido de la voz original y la sintetizada, aunque no es una medida definitiva para todos los codificadores. Previamente a la comparación se han igualado en media y energía las señales sintetizadas con la original.
3. Tiempo de computación. Es una medida del tiempo de procesamiento de análisis y síntesis para cada codificador. Debido a que depende de la duración del flujo de entrada, esta medida se va a presentar como un múltiplo del tiempo de computación del mínimo de los tiempos de

Comparación entre codificadores de voz

codificación en cada experimento. Debe considerarse en términos de orden de magnitud, ya que el tiempo de ejecución no se mantiene entre realizaciones sucesivas.

4. MOS. Es una medida de calidad subjetiva en la que los flujos de voz sintetizada se valorarán en una escala de cinco puntos, con la mayor puntuación asociada a la mejor calidad.

En los experimentos se ha usado la siguiente tabla de asignación de bits:

<i>ADPCM</i>	
Ganancia (bits)	7
Error (bits)	7

<i>Codificador Armónico</i>	
Amplitud (bits)	10
Frecuencia (bits)	16
Fase (bits)	6

<i>CELP</i>	
Coeficientes LPC (bits)	12
Ganancia predicción corta (bits)	9
Ganancia predicción larga (bits)	5
Periodo de pitch (bits)	7
Índice del libro de códigos (bits)	9

<i>DCT</i>	
Coeficientes DCT (bits)	12
Índices (bits)	6

<i>LPC</i>	
Coeficientes LPC (bits)	10
Ganancia predicción corta (bits)	5
Periodo de pitch (bits)	9

4.1. EXPERIMENTO 01.

Datos de entrada.

Nombre del fichero	sonido01.wav
Tipo de voz	hombre
Transcripción	"¿Dígame?"
Duración	0.34 segundos
Tasa de entrada	64000 bits/segundo

Datos de salida.

<i>Codificador\Datos</i>	Tasa de salida (bits/segundo)	Ratio de compresión (%)	SNR (dB)
ADPCM	56350	11.95	12.34
Armónico	32000	50.00	-1.75
CELP	7500	88.28	-2.45
DCT	61700	3.52	35.44
LPC	4133	93.54	-6.23

<i>Codificador\Datos</i>	Tiempo de computación	Percepción subjetiva (MOS)
ADPCM	83	5
Armónico	3	3
CELP	38	1
DCT	1	4
LPC	4	3

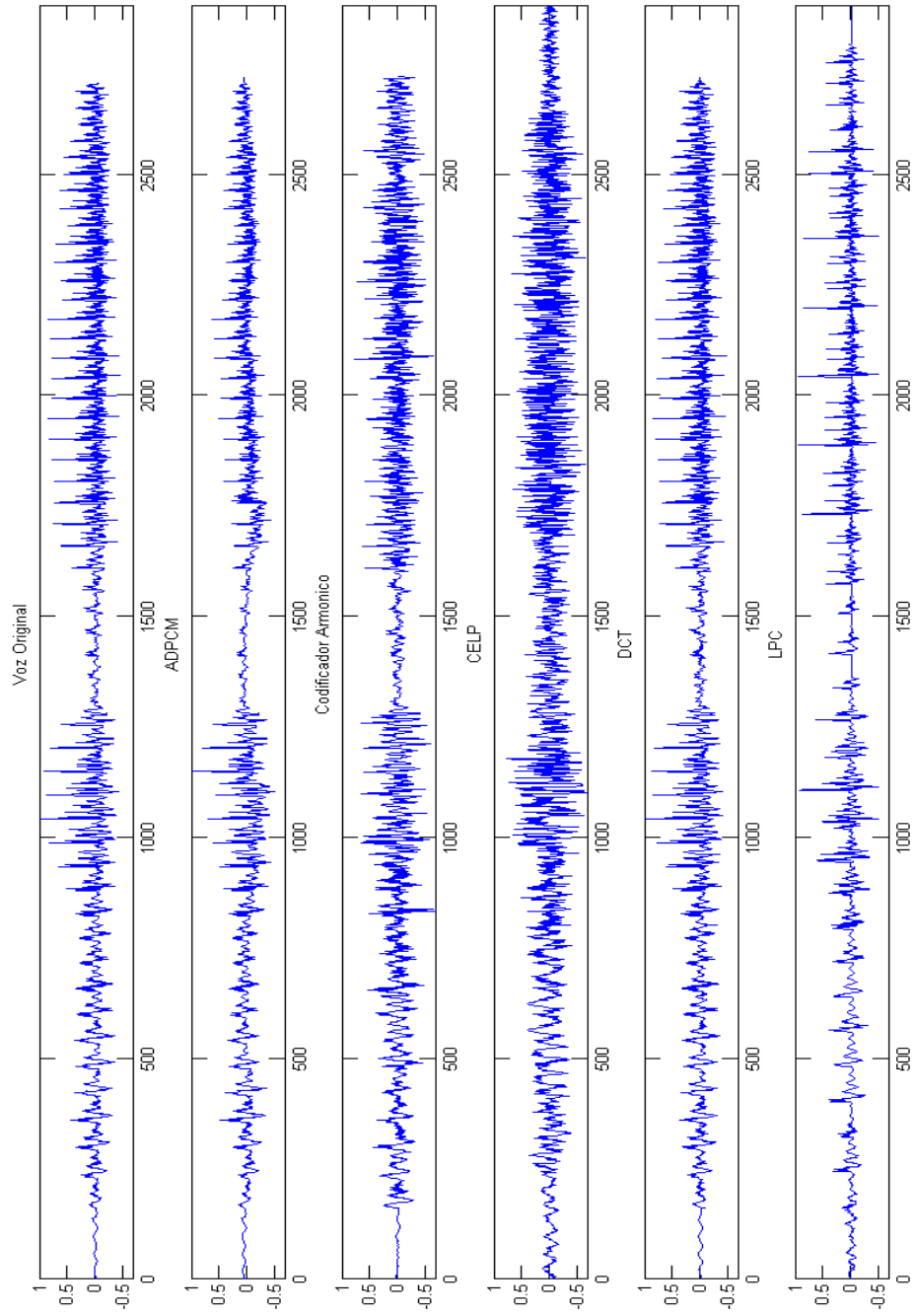


Figura 57: Experimento 01

4.2. EXPERIMENTO 02.

Datos de entrada.

Nombre del fichero	sonido02.wav
Tipo de voz	hombre
Transcripción	"Adiós"
Duración	0.72 segundos
Tasa de entrada	64000 bits/segundo

Datos de salida.

<i>Codificador\Datos</i>	Tasa de salida (bits/segundo)	Ratio de compresión (%)	SNR (dB)
ADPCM	56350	11.95	22.05
Armónico	32000	50.00	2.66
CELP	7500	88.28	-2.14
DCT	24200	62.17	31.82
LPC	4133	93.54	-7.03

<i>Codificador\Datos</i>	Tiempo de computación	Percepción subjetiva (MOS)
ADPCM	90	5
Armónico	3	3
CELP	39	1
DCT	1	3
LPC	4	2

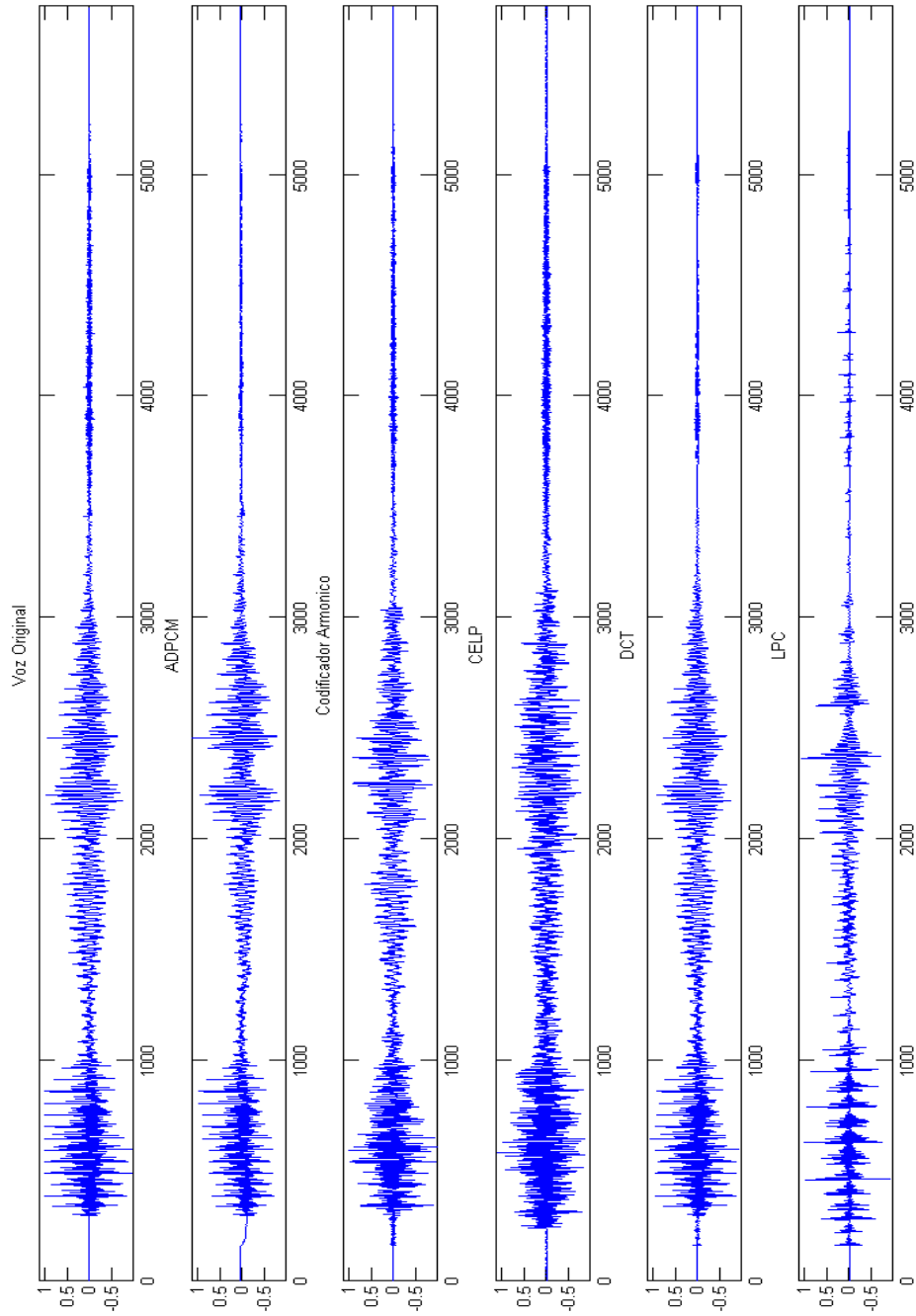


Figura 58: Experimento 02

4.3. EXPERIMENTO 03

Datos de entrada.

Nombre del fichero	sonido03.wav
Tipo de voz	hombre
Transcripción	"¡Jesús!"
Duración	0.91 segundos
Tasa de entrada	64000 bits/segundo

Datos de salida.

<i>Codificador\Datos</i>	Tasa de salida (bits/segundo)	Ratio de compresión (%)	SNR (dB)
ADPCM	56350	11.95	17.23
Armónico	32000	50.00	0.23
CELP	7500	88.28	-1.01
DCT	19700	69.11	28.88
LPC	4133	93.54	-6.18

<i>Codificador\Datos</i>	Tiempo de computación	Percepción subjetiva (MOS)
ADPCM	116	5
Armónico	4	2
CELP	50	1
DCT	1	4
LPC	5	2

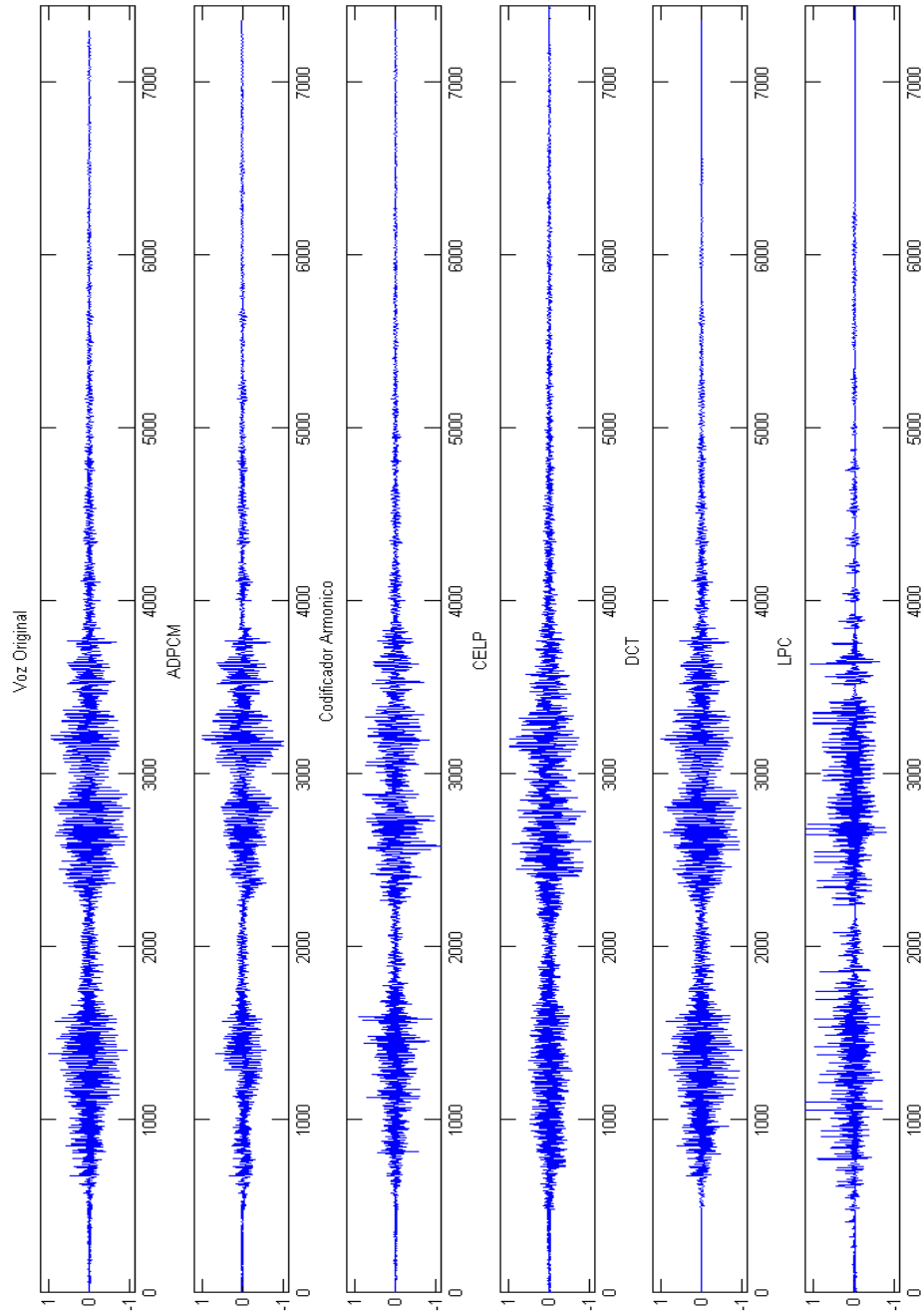


Figura 59: Experimento 03

4.4. EXPERIMENTO 04

Datos de entrada.

Nombre del fichero	sonido04.wav
Tipo de voz	hombre
Transcripción	"Necesito un aumento de sueldo."
Duración	1.47 segundos
Tasa de entrada	64000 bits/segundo

Datos de salida.

<i>Codificador\Datos</i>	Tasa de salida (bits/segundo)	Ratio de compresión (%)	SNR (dB)
ADPCM	56350	11.95	6.75
Armónico	32000	50.00	2.96
CELP	7500	88.28	0.19
DCT	16700	73.90	30.07
LPC	4133	93.54	-5.79

<i>Codificador\Datos</i>	Tiempo de computación	Percepción subjetiva (MOS)
ADPCM	74	4
Armónico	2	2
CELP	33	1
DCT	1	3
LPC	3	2

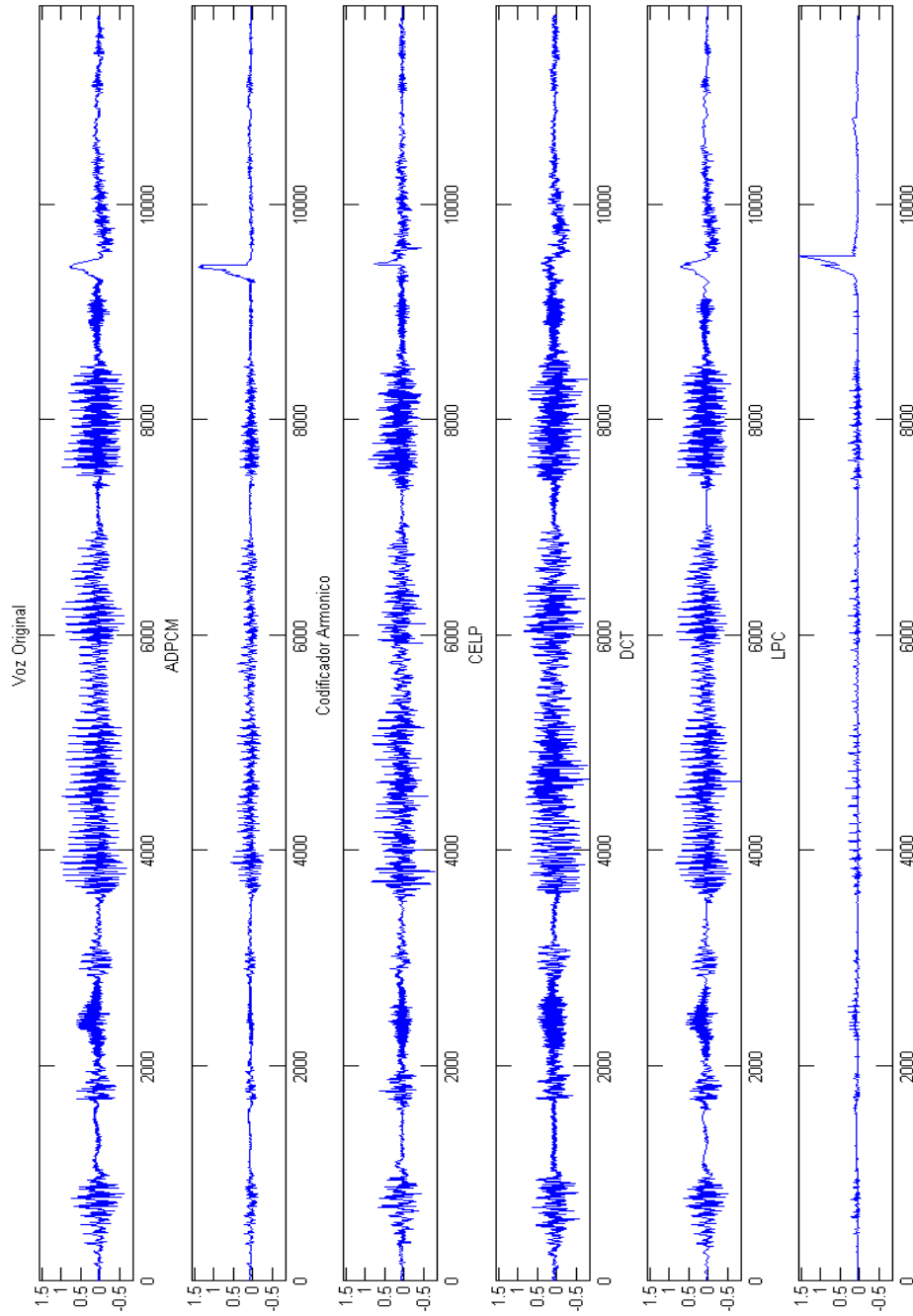


Figura 60: Experimento 04

4.5. EXPERIMENTO 05.

Datos de entrada.

Nombre del fichero	Sonido05.wav
Tipo de voz	hombre
Transcripción	"Caminante no hay camino, se hace camino al andar."
Duración	3.48 segundos
Tasa de entrada	64000 bits/segundo

Datos de salida.

Codificador\Datos	Tasa de salida (bits/segundo)	Ratio de compresión (%)	SNR (dB)
ADPCM	56350	11.95	13.21
Armónico	32000	50.00	3.79
CELP	7500	88.28	-0.51
DCT	21900	65.76	31.93
LPC	4133	93.54	-4.56

Codificador\Datos	Tiempo de computación	Percepción subjetiva (MOS)
ADPCM	46	4
Armónico	1	3
CELP	20	1
DCT	1	4
LPC	2	2

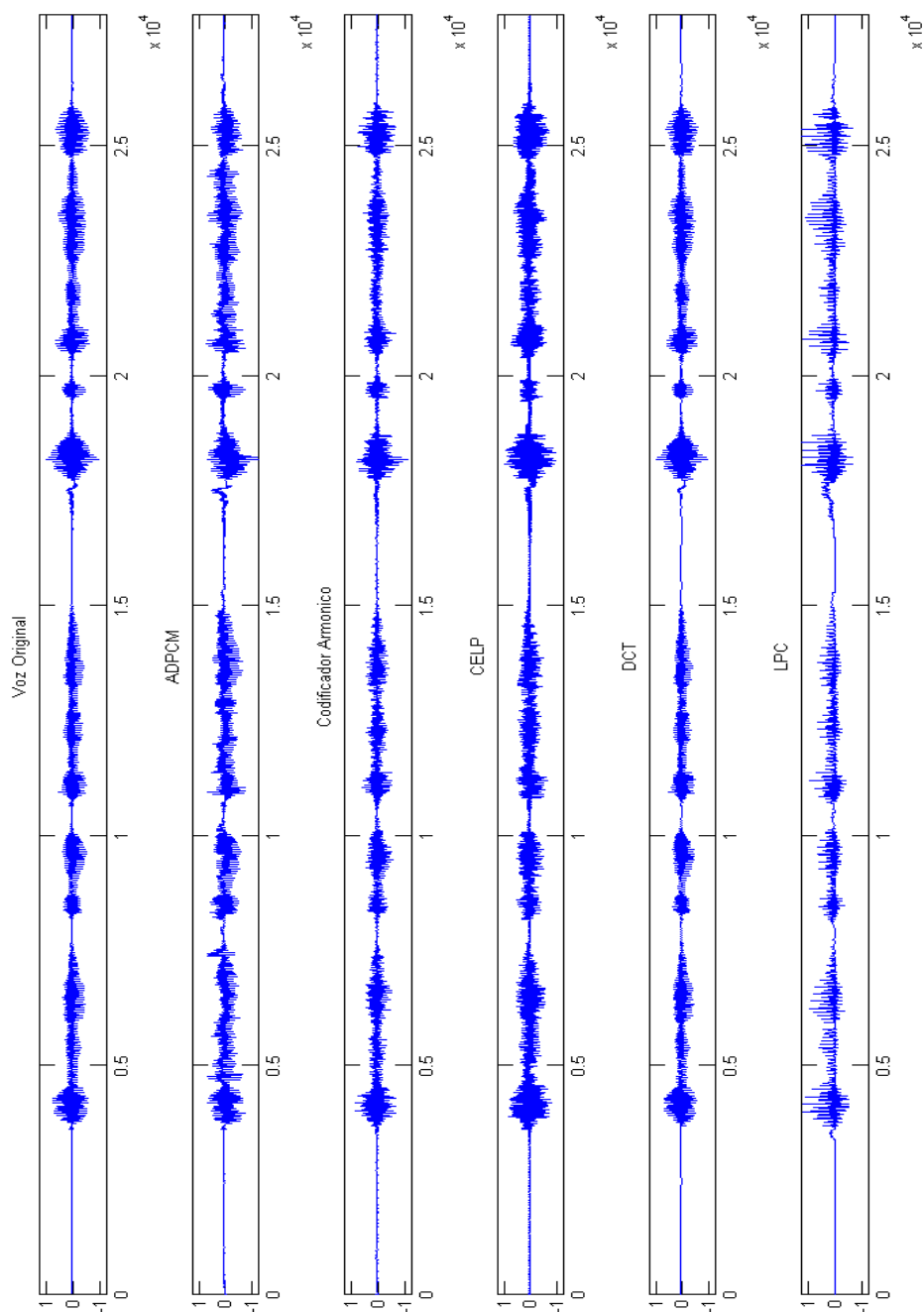


Figura 61: Experimento 05

4.6. EXPERIMENTO 06.

Datos de entrada.

Nombre del fichero	sonido06.wav
Tipo de voz	hombre
Transcripción	"Un vecino de Sevilla ha sido agraciado con el primer premio."
Duración	4.34 segundos
Tasa de entrada	64000 bits/segundo

Datos de salida.

<i>Codificador\Datos</i>	Tasa de salida (bits/segundo)	Ratio de compresión (%)	SNR (dB)
ADPCM	56350	11.95	11.75
Armónico	32000	50.00	4.30
CELP	7500	88.28	3.08
DCT	5700	90.99	25.22
LPC	4133	93.54	-2.79

<i>Codificador\Datos</i>	Tiempo de computación	Percepción subjetiva (MOS)
ADPCM	68	4
Armónico	2	2
CELP	31	1
DCT	1	2
LPC	3	2

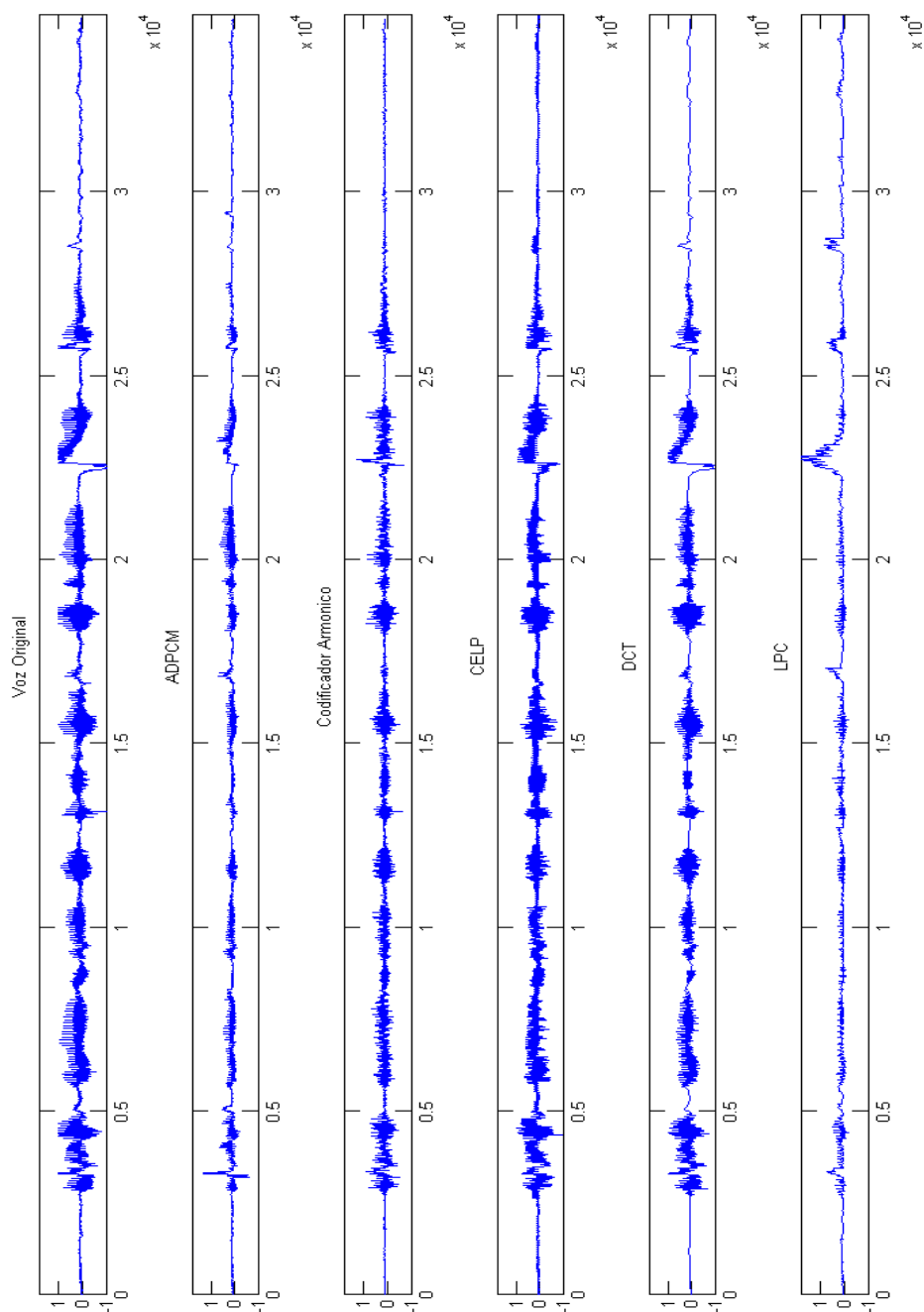


Figura 62: Experimento 06

4.7. EXPERIMENTO 07.

Datos de entrada.

Nombre del fichero	Sonido07.wav
Tipo de voz	hombre
Transcripción	"Me encanta pasear por las tardes."
Duración	4.39 segundos
Tasa de entrada	64000 bits/segundo

Datos de salida.

Codificador\Datos	Tasa de salida (bits/segundo)	Ratio de compresión (%)	SNR (dB)
ADPCM	56350	11.95	9.90
Armónico	32000	50.00	1.05
CELP	7500	88.28	0.90
DCT	6330	90.13	24.92
LPC	4133	93.54	-5.62

Codificador\Datos	Tiempo de computación	Percepción subjetiva (MOS)
ADPCM	66	4
Armónico	2	2
CELP	29	1
DCT	1	3
LPC	3	2

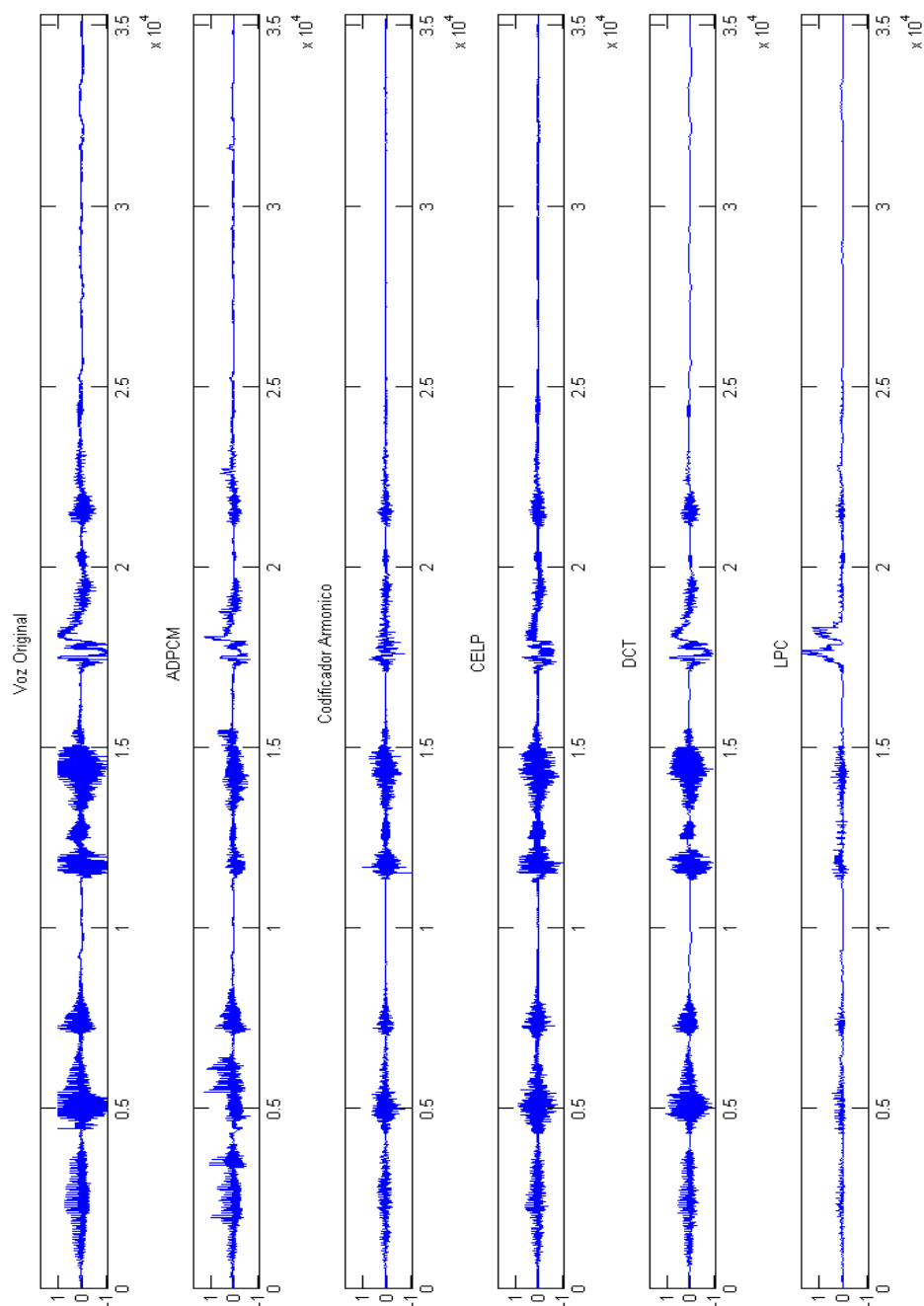


Figura 63: Experimento 07

4.8. EXPERIMENTO 08.

Datos de entrada.

Nombre del fichero	sonido08.wav
Tipo de voz	hombre
Transcripción	"En mis ratos libres juego al dominó."
Duración	2.99 segundos
Tasa de entrada	64000 bits/segundo

Datos de salida.

<i>Codificador\Datos</i>	Tasa de salida (bits/segundo)	Ratio de compresión (%)	SNR (dB)
ADPCM	56350	11.95	9.26
Armónico	32000	50.00	1.73
CELP	7500	88.28	1.14
DCT	8880	86.18	29.67
LPC	4133	93.54	-5.83

<i>Codificador\Datos</i>	Tiempo de computación	Percepción subjetiva (MOS)
ADPCM	76	3
Armónico	2	3
CELP	33	1
DCT	1	3
LPC	3	2

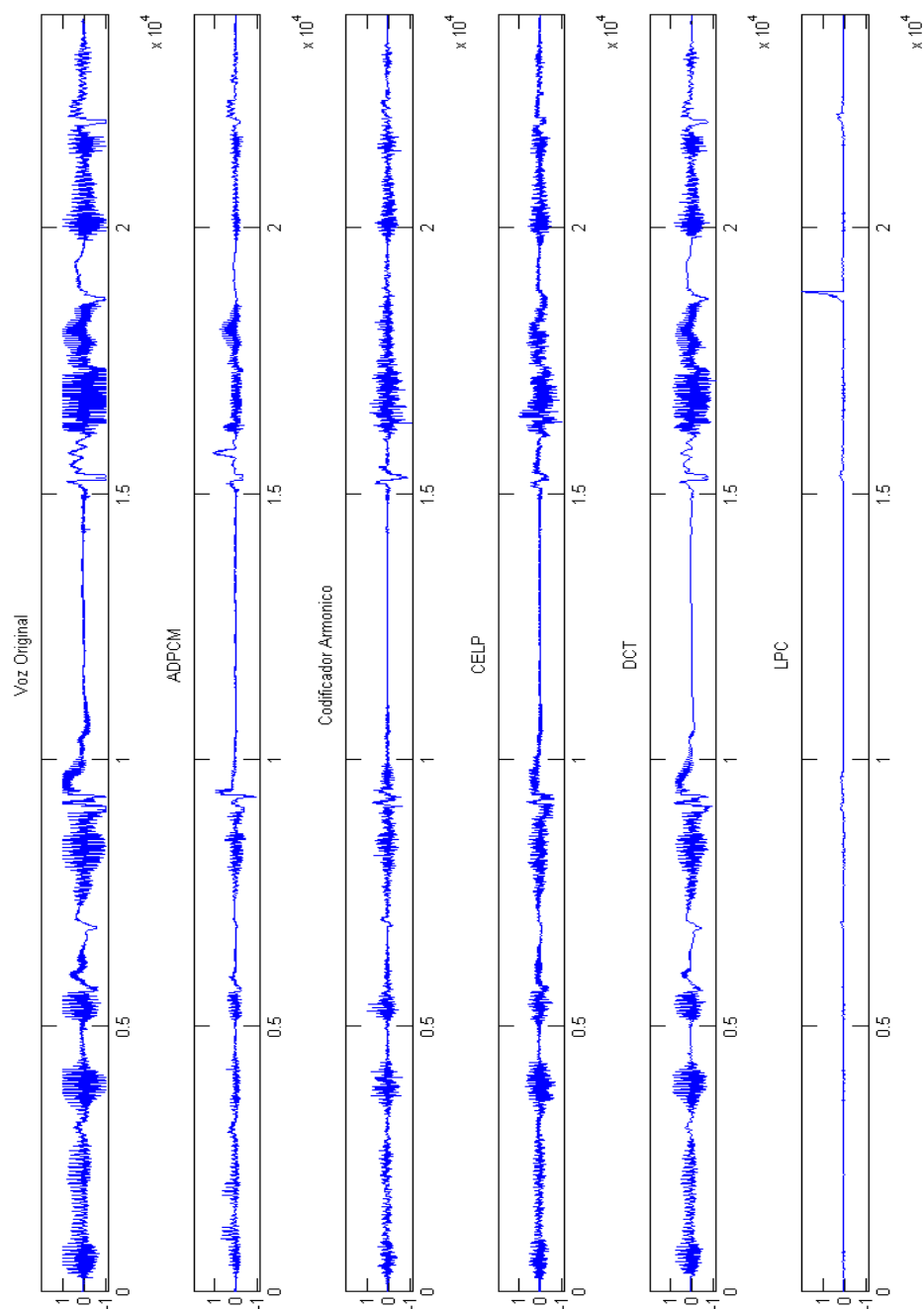


Figura 64: Experimento 08

4.9. EXPERIMENTO 09.

Datos de entrada.

Nombre del fichero	sonido11.wav
Tipo de voz	mujer
Transcripción	"¡Oh, que bonito!"
Duración	2.09 segundos
Tasa de entrada	64000 bits/segundo

Datos de salida.

<i>Codificador\Datos</i>	Tasa de salida (bits/segundo)	Ratio de compresión (%)	SNR (dB)
ADPCM	56350	11.95	23.72
Armónico	32000	50.00	4.14
CELP	7500	88.28	1.20
DCT	15200	76.29	35.21
LPC	4133	93.54	-8.24

<i>Codificador\Datos</i>	Tiempo de computación	Percepción subjetiva (MOS)
ADPCM	88	5
Armónico	3	3
CELP	38	2
DCT	1	4
LPC	3	2

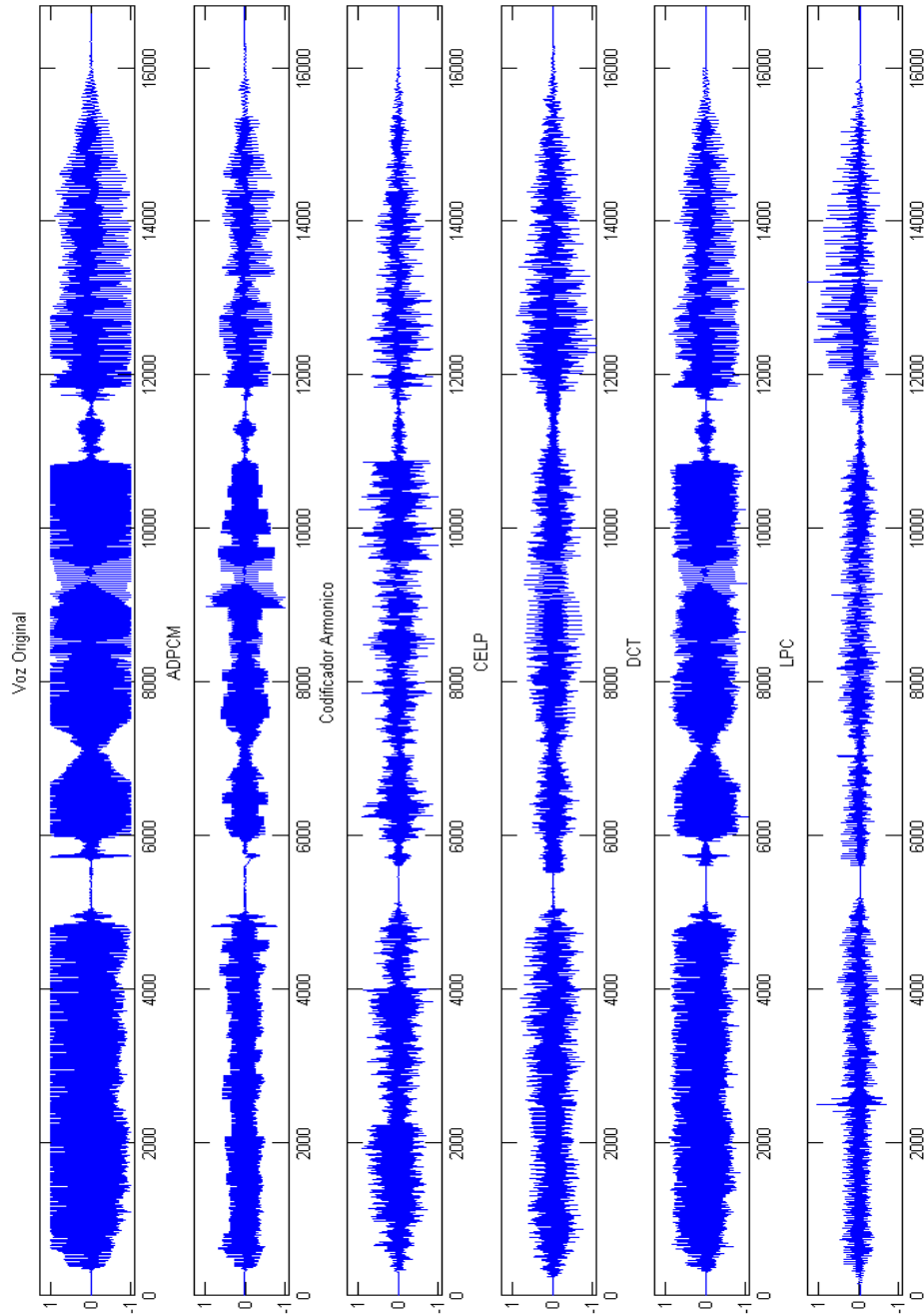


Figura 65: Experimento 09

4.10. EXPERIMENTO 10.

Datos de entrada.

Nombre del fichero	sonido12.wav
Tipo de voz	mujer
Transcripción	"Si quiere, puede dejar un mensaje después de oír la señal, gracias."
Duración	3.53 segundos
Tasa de entrada	64000 bits/segundo

Datos de salida.

<i>Codificador\Datos</i>	Tasa de salida (bits/segundo)	Ratio de compresión (%)	SNR (dB)
ADPCM	56350	11.95	13.75
Armónico	32000	50.00	-0.11
CELP	7500	88.28	-1.92
DCT	14000	78.09	24.74
LPC	4133	93.54	-6.92

<i>Codificador\Datos</i>	Tiempo de computación	Percepción subjetiva (MOS)
ADPCM	69	5
Armónico	2	4
CELP	30	3
DCT	1	5
LPC	3	3

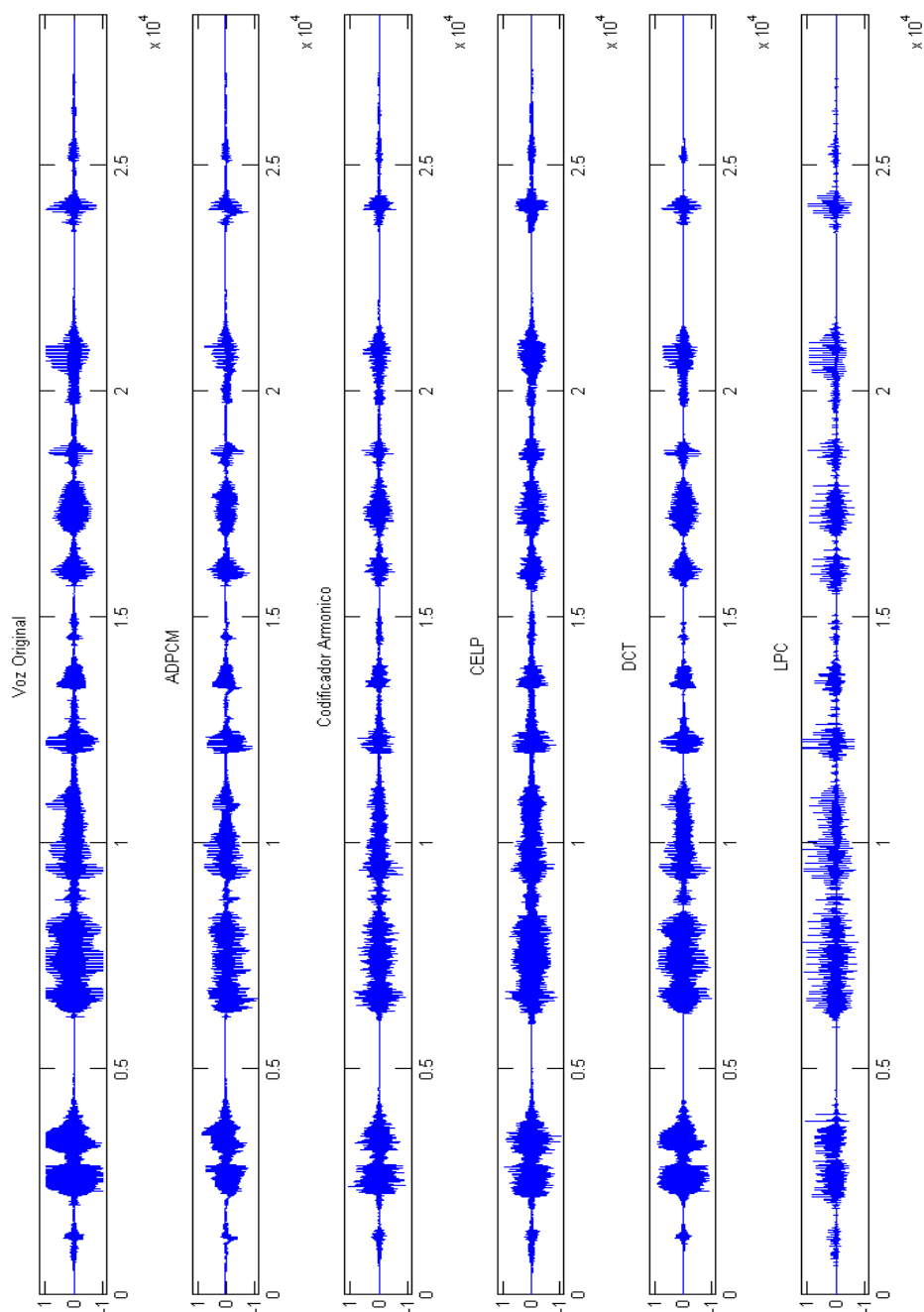


Figura 66: Experimento 10

4.11. EXPERIMENTO 11.

Datos de entrada.

Nombre del fichero	sonido13.wav
Tipo de voz	mujer
Transcripción	"A través del follaje perenne, que oír deja rumores extraños..."
Duración	3.87 segundos
Tasa de entrada	64000 bits/segundo

Datos de salida.

<i>Codificador\Datos</i>	Tasa de salida (bits/segundo)	Ratio de compresión (%)	SNR (dB)
ADPCM	56350	11.95	19.68
Armónico	32000	50.00	2.96
CELP	7500	88.28	-0.61
DCT	20000	68.70	29.44
LPC	4133	93.54	-6.98

<i>Codificador\Datos</i>	Tiempo de computación	Percepción subjetiva (MOS)
ADPCM	42	4
Armónico	1	3
CELP	19	2
DCT	1	4
LPC	2	2

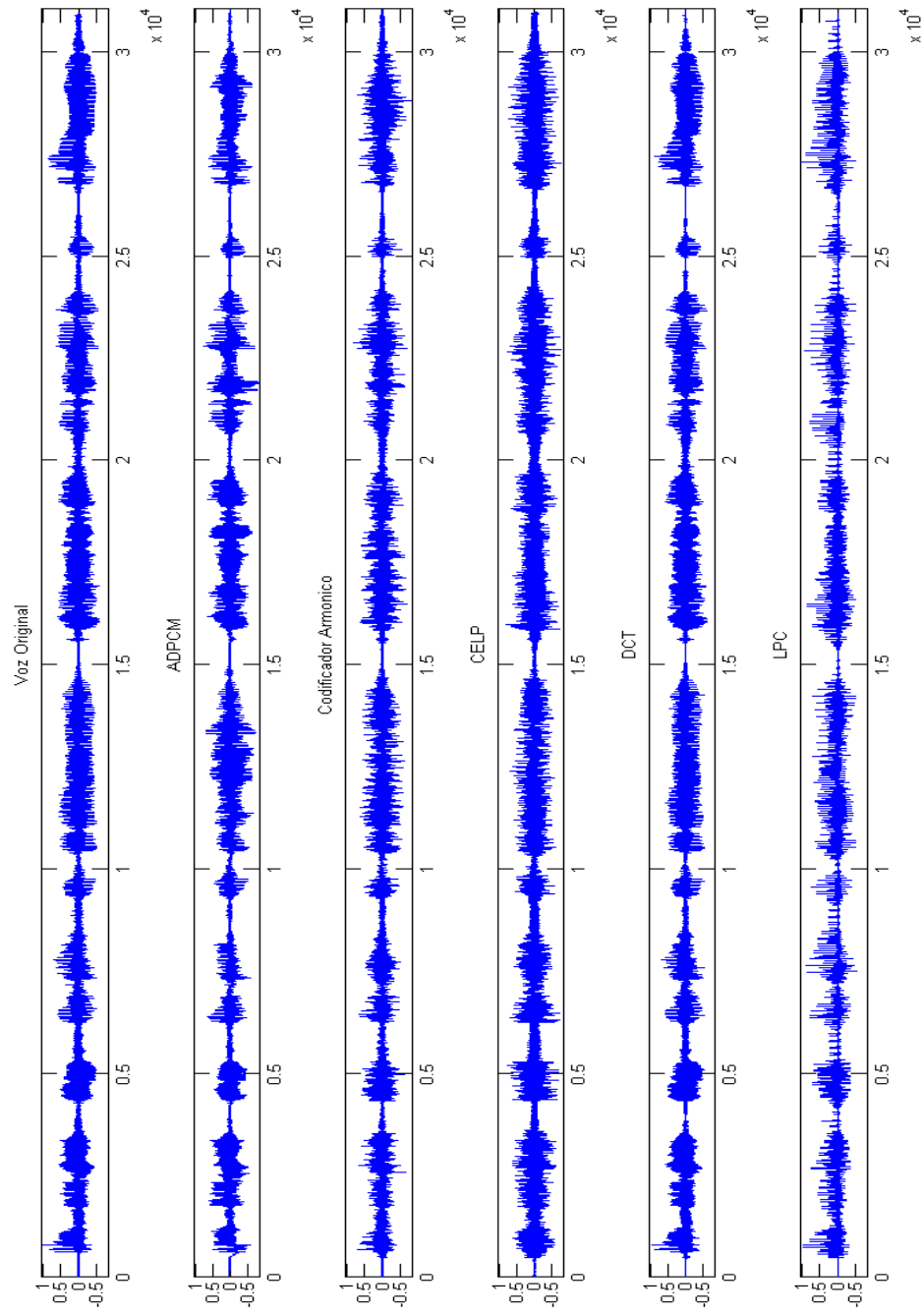


Figura 67: Experimento 11

4.12. EXPERIMENTO 12.

Datos de entrada.

Nombre del fichero	sonido14.wav
Tipo de voz	mujer
Transcripción	"¡Socorro!"
Duración	1.34 segundos
Tasa de entrada	64000 bits/segundo

Datos de salida.

<i>Codificador\Datos</i>	Tasa de salida (bits/segundo)	Ratio de compresión (%)	SNR (dB)
ADPCM	56350	11.95	27.40
Armónico	32000	50.00	-0.42
CELP	7500	88.28	-0.91
DCT	33200	48.07	32.47
LPC	4133	93.54	-5.97

<i>Codificador\Datos</i>	Tiempo de computación	Percepción subjetiva (MOS)
ADPCM	85	4
Armónico	2	3
CELP	38	1
DCT	1	4
LPC	3	1

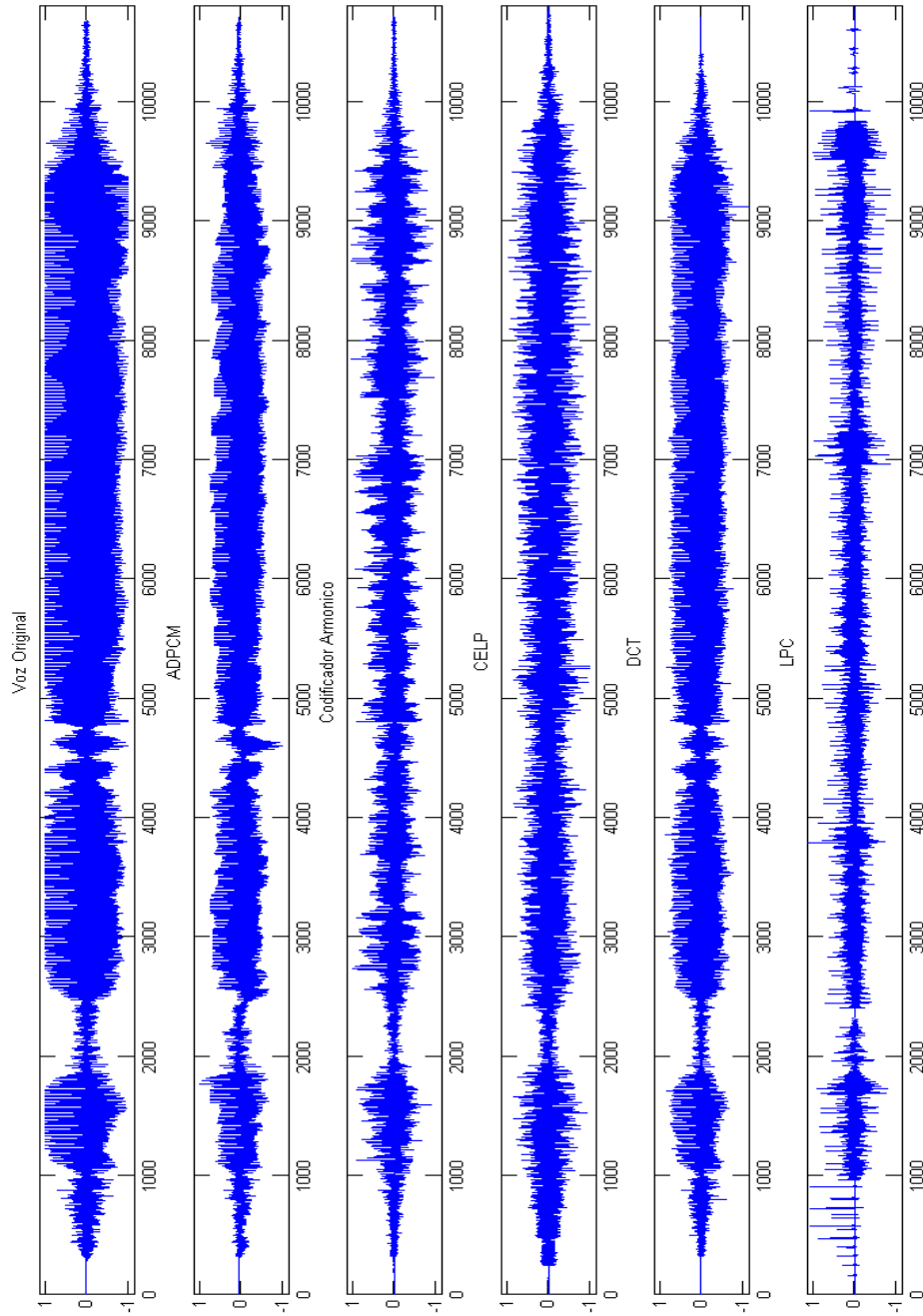


Figura 68: Experimento 12

4.13. EXPERIMENTO 13.

Datos de entrada.

Nombre del fichero	sonido15.wav
Tipo de voz	mujer
Transcripción	"¿Quién es el último?"
Duración	0.99 segundos
Tasa de entrada	64000 bits/segundo

Datos de salida.

Codificador\Datos	Tasa de salida (bits/segundo)	Ratio de compresión (%)	SNR (dB)
ADPCM	56350	11.95	11.27
Armónico	32000	50.00	3.29
CELP	7500	88.28	-0.94
DCT	17000	73.55	27.61
LPC	4133	93.54	-7.31

Codificador\Datos	Tiempo de computación	Percepción subjetiva (MOS)
ADPCM	83	5
Armónico	3	3
CELP	36	2
DCT	1	3
LPC	3	2

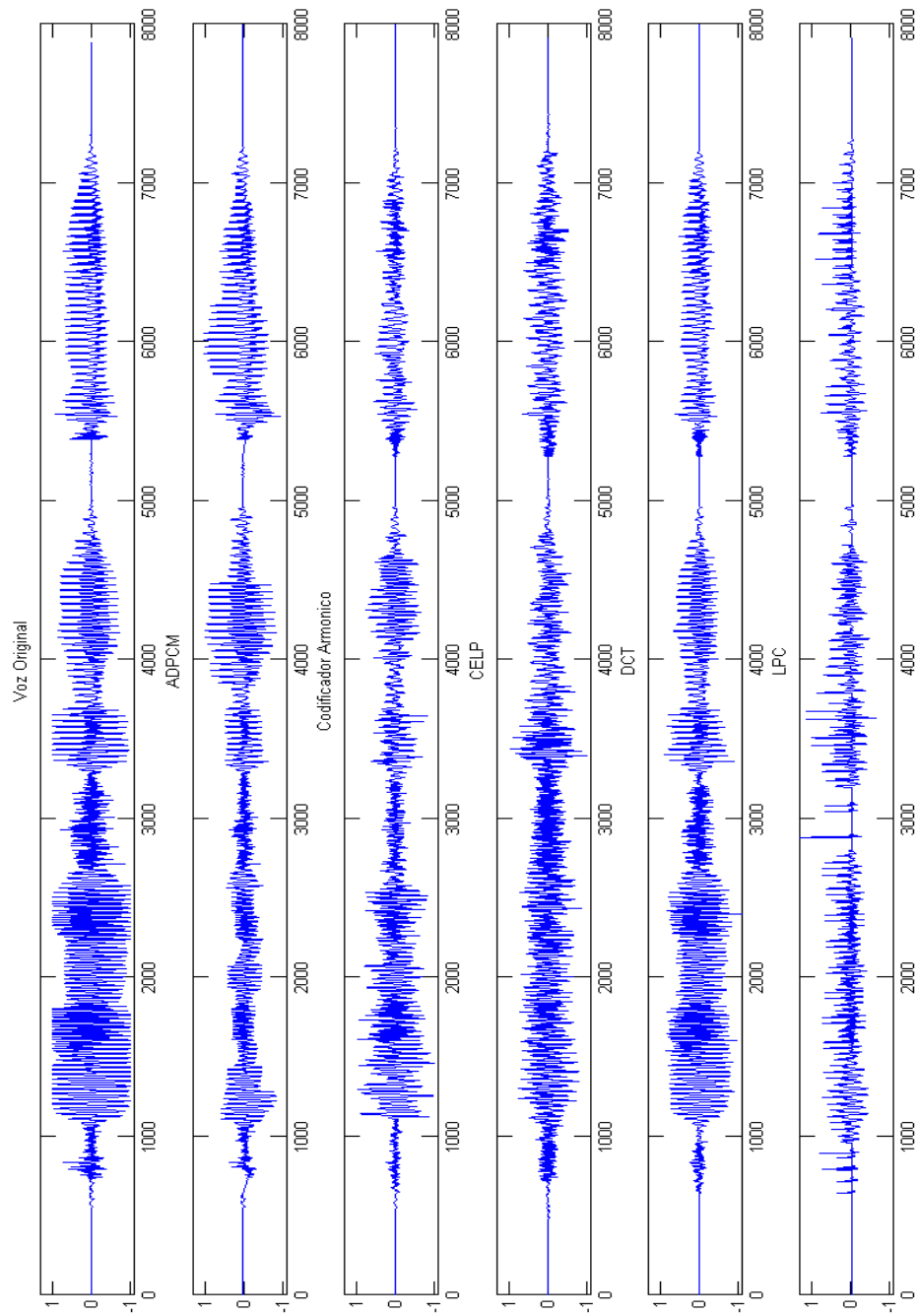


Figura 69: Experimento 13