



VI Congreso Iberoamericano de Acústica - FIA 2008  
Buenos Aires, 5, 6 y 7 de noviembre de 2008

FIA2008-A226

## **Estimador de distorsión no lineal usando características psicoacústicas aplicado a una situación de altavoces**

Víctor Manuel Espinoza Catalán<sup>(a)</sup>,

(a) Laboratorio Sala Anecoica, Universidad Tecnológica de Chile. Brown Norte 290, Santiago, Chile. E-mail: victor.espinoza11@docentes.inacap.cl

Centro Tecnológico, Facultad de Artes, Universidad de Chile, Compañía 1264, Santiago, Chile. E-mail: vespinozac@gmail.com

### **Abstract**

One way to improve the quality of sound reproduction through loudspeakers focuses on developing techniques for reducing the nonlinear distortion. This type of phenomenon it is showed in a subjective test by trained auditors and distortion estimators (THD, IMD, Ruzz&Buzz and others). The estimators measures the energy produced by the distortion almost without unweighted the psychoacoustic features of human hearing. At present work shows some of these features and then uses them with the classical estimators concept. For this achievement, it is employed a model of loudness with artificial neural networks and a model of simultaneous masking using Matlab language. Finally, it will show the measurements made in anechoic chamber using different loudspeakers and then compare them with classical estimators.

### **Resumen**

Una de las formas para mejorar la calidad de reproducción sonora mediante altavoces se centra en el desarrollo de técnicas para la disminución de la distorsión no lineal. Este tipo de fenómeno se evidencia en una evaluación subjetiva por parte de auditores entrenados y por estimadores de distorsión (THD, IMD, Rub&Buzz, etc.). Los estimadores clásicos consideran la medición de energía producida por la distorsión casi sin ponderar las características psicoacústicas del oído. El presente trabajo consiste en incluir algunas características psicoacústicas y combinarlas con el concepto asociado a los estimadores clásicos. Para esto, se empleará un modelo de sonoridad implementado con redes neuronales artificiales y un modelo de enmascaramiento simultaneo en frecuencia programado en Matlab. Se presentan mediciones realizadas en cámara anecoica a distintos altavoces y se comparan con los estimadores clásicos.

## 1 Introducción

En la actualidad, para determinar el grado de distorsión de un dispositivo electroacústico, se emplean índices porcentuales de la cantidad de energía producida por frecuencias armónicas o no armónicas. Conocidos son el “Total Harmonic Distortion” %THD y el “Inter Modulation Distortion” %IMD. Estos estimadores suelen tener valores bajos en dispositivos electrónicos, pero altos en sistemas de altavoces. Esto se debe a la naturaleza no lineal de la transducción electromecanoacústica; sobretodo en altavoces electrodinámicos. Klippel (2006), describe con claridad la dependencia de los parámetros Thiele-Small con la excursión y temperatura alcanzada por el altavoz en los parámetros  $Bl(x)$ ,  $Le(x)$ ,  $Kms(x)$ . De igual forma, la distorsión no lineal puede evidenciar defectos en la construcción del altavoz. Ejemplo de esto es: una bobina ubicada fuera del punto de equilibrio de las líneas de fuerza y sistema móvil (spider, diafragma, suspensión) no alineado al chasis. Los diseñadores de sistemas de altavoces consideran el grado de distorsión de estos, pero en ocasiones valores altos de distorsión son evaluados con “mejor” sonido, que otros con bajos valores. Otros estimadores más recientes, como %Rub&Buzz (Temme 2000) ponderan en mayor grado las frecuencias en zonas medias altas, pero sin ser muy específicos en otras zonas de frecuencia. El presente artículo propone un estimador que considere características psicoacústicas del oído para la medida de distorsión “audible” de un dispositivo.

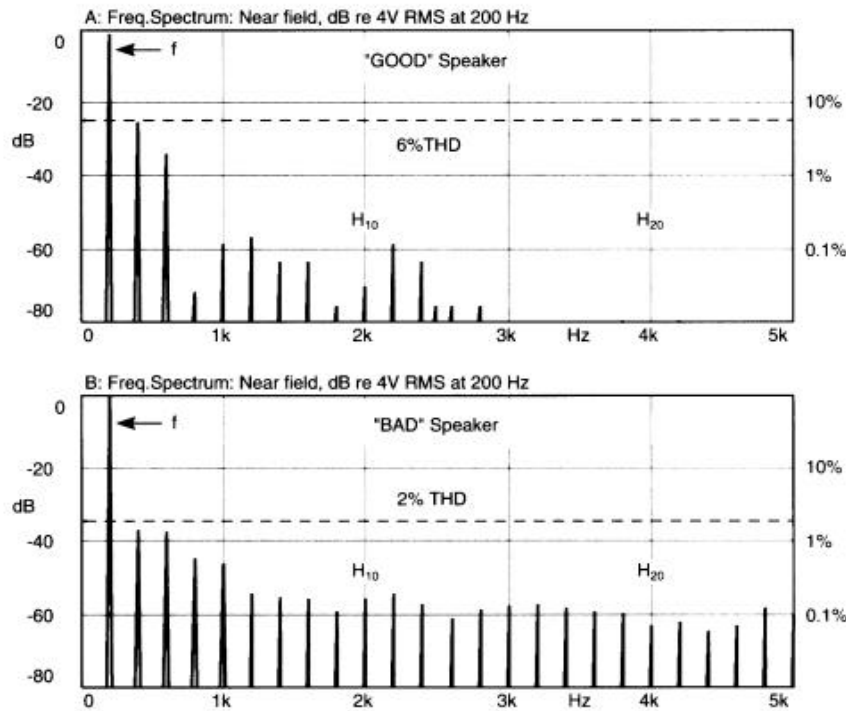
## 2 Antecedentes Teóricos

### 2.1 Estimadores de distorsión no lineal

En general, los estimadores de distorsión no lineal consideran la relación de energía de entrada versus salida de un dispositivo bajo test. Entre los estimadores más utilizados se encuentran %THD, %IMD, %Rub&Buzz, entre otros. Para un dispositivo absolutamente lineal estos valores son cero (0%), es decir, sin contribución de frecuencias armónicas o moduladas. Consideremos las ecuaciones (1) y (2) que calculan dos de estos estimadores: %THD y %Rub&Buzz. El primero considera toda la energía de los armónicos producidos en el rango audible. El segundo solamente pondera armónicos superiores (entre el N° 15 y 20), que tienen mayor posibilidad de ser escuchados, o sea, considera características psicoacústicas del oído. Comúnmente, un valor alto de THD significa mayor distorsión frente a otro de bajo valor. Debido a que este estimador no considera características auditivas del oído, falla en la evaluación subjetiva por parte de un auditor, es decir, un altavoz con THD mayor que otro es considerado como de menor “distorsión subjetiva”. Temme (2000) muestra esta diferencia con un ejemplo (ver Figura 1).

$$\%THD = 100 \times \sqrt{\frac{\sum_{i=1}^{n-1} H_{i+1}^2}{\sum_{i=1}^n H_i^2}} \quad (1)$$

$$\%Rub \& Buzz = 100 \times \sqrt{\frac{\sum_{i=15}^{20} H_i^2}{\sum_{i=1}^n H_i^2}} \quad (2)$$



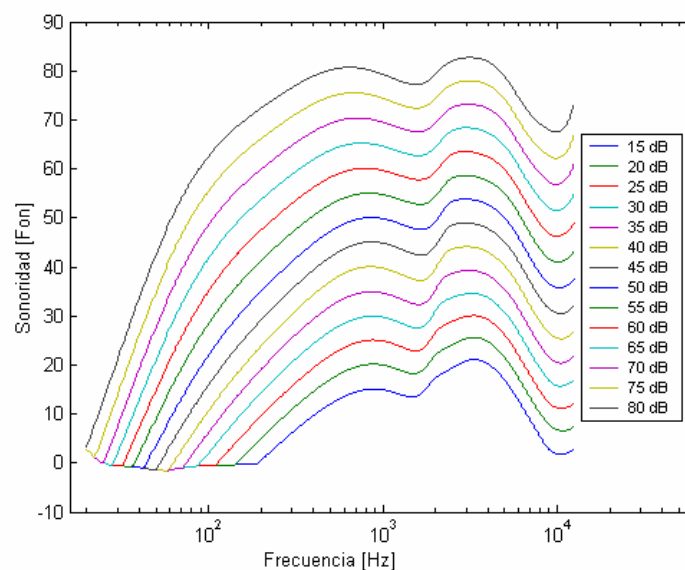
**Figura 1.** Ejemplo donde un altavoz con menor %THD, es juzgado como “Malo”. (Temme 2000)

Los valores de %Rub&Buzz de la figura 1 son: 0.02%, para altavoz “Bueno”, y 0.3% para altavoz “Malo”. En este caso el %Rub&Buzz de cada medición pretende guardar directa relación con la calidad subjetiva. Una hipótesis de lo que ocurre, es que los armónicos cercanos a la frecuencia de prueba tienen poca incidencia en la valoración subjetiva.

## 2.2 Modelo de Sonoridad

A partir de un estudio realizado por Suzuki y Takeshima (2004), fueron calculadas nuevamente las curvas isofónicas. Este estudio fue la base para de la normativa de referencia ISO 226:2003. En tal investigación, los autores revisan estudios anteriores y determinan nuevos valores para las curvas. Errores de hasta 14 dB NPS son detectados, los cuales ocurren principalmente en baja frecuencia. Usando la característica de aproximador universal de las Redes Neuronales Artificiales, Espinoza, Venegas y Floody (2006) determinan los niveles de sonoridad en fones de cualquier frecuencia y amplitud en el rango de 0 a 80 fones, y 20 a 12500 Hz, mediante la función NeuLoud(frecuencia,dB NPS)=Fones. Esto permitirá ponderar

la amplitud de cada frecuencia, y determinar la percepción aproximada por parte del oído. Un ejemplo de las curvas de ponderación aplicadas se presenta en la figura 2.



**Figura 2.** Curvas de igual contorno de NPS versus Frecuencia. (Espinoza, Venegas y Floody, 2006).

### 2.3 Modelo de Enmascaramiento Simultáneo

En ciertas situaciones el oído es incapaz de detectar señales audibles, que sí podría detectar en otras. Esto sucede cuando algunos sonidos ocultan o enmascaran a otros. Se habla entonces de “enmascaramiento” del sonido, que es la incapacidad del oído de detectar un sonido frente a otro. Están bien documentados, mediciones y modelos matemáticos que describen este comportamiento bajo ciertas condiciones. Una detallada explicación de estos y otros fenómenos psicoacústicos, puede encontrarse en Zwicker y Fastl (2007). A continuación se revisará los datos y modelos utilizados en este artículo.

La figura 3, muestra las curvas de umbral de enmascaramiento para tres frecuencias. A estas curvas se les denominan bandas críticas. Se define tono enmascarador (línea azul y nivel mayor), al tono que genera una determinada curva de enmascaramiento, bajo la cual el oído no percibe al tono enmascarado (línea roja y nivel menor). Notar que el tono de 250 Hz enmascara al de 500 Hz, es decir, este último no se escucharía frente al de 250 Hz. No sucede lo mismo si se considera al tono de 1000 Hz. Este no enmascara al de 500 Hz.

Mediante mediciones se ha detectado que el comportamiento de este fenómeno es variable con la frecuencia y el nivel, como la muestra la figura 4. Los anchos de banda crítica varían y las pendientes sobre y bajo el tono enmascarador son no simétricas.

Otra característica psicoacústica a utilizar para el modelo de enmascaramiento es la escala de sensación relativa a la altura (pitch-related scale). En la figura 5 se aprecia la membrana basilar del oído con diferentes medidas. Una de estas es el Bark, unidad con la cual comúnmente se mide la sensación relativa de altura. Cabe observar que esta medida es proporcional respecto a la posición en la membrana basilar; no así la frecuencia. Un modelo matemático de la relación entre Bark y frecuencia es presentado en la ecuación (3).

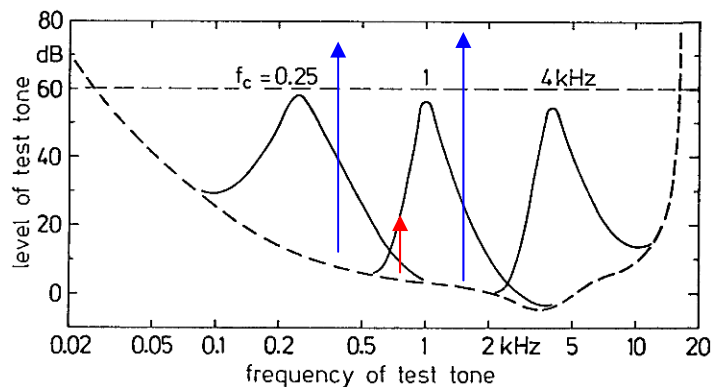


Figura 3. Ejemplo de enmascaramiento simultáneo (Zwicker y Fastl, 2007).

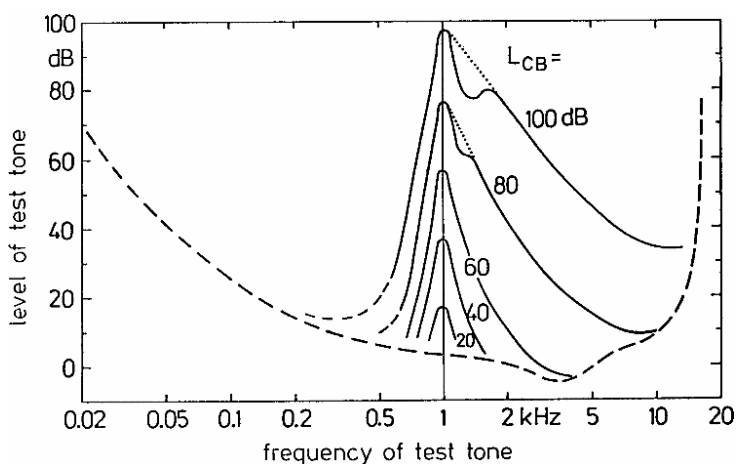


Figura 4. Variación de la banda crítica centrada en 1 kHz a distintos niveles.

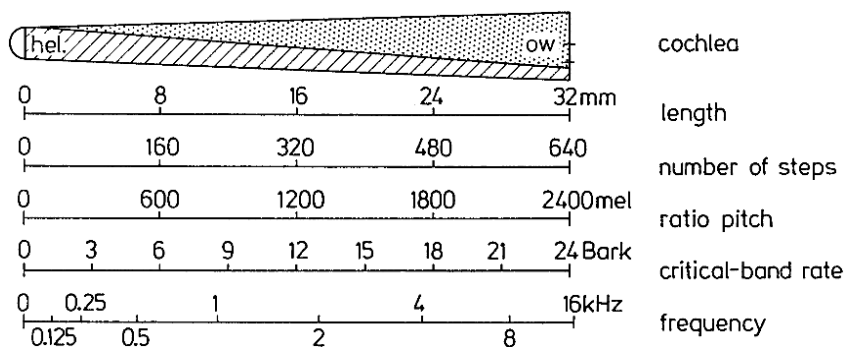


Figura 5. Membrana basilar con sus medidas más utilizadas.

$$z = 13 \times \text{atan}(0,76 \times f) + 3,5 \times \text{atan}(f / 7,5)^2 \tag{3}$$

Donde  $f$ : frecuencia en kHz.

Si se convierte el eje de frecuencias en unidades Bark, y se grafican varias bandas críticas (como se muestra en la figura 6) se observa que las pendientes de la banda a un mismo nivel (en el ejemplo 60 dB), son similares, y pueden ser aproximadas a un modelo graficado en la figura 7,

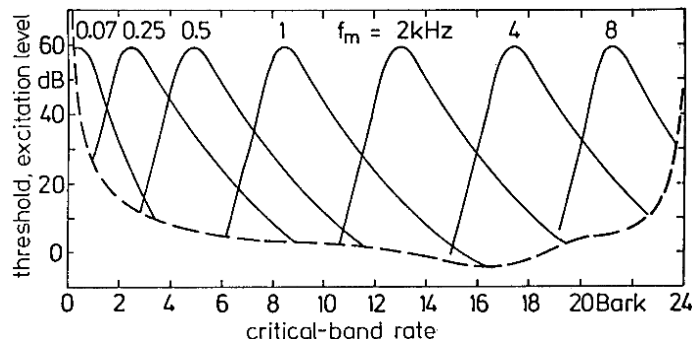


Figura 6. Curvas de enmascaramiento para varios Bark. (Zwicker y Fastl, 2007)

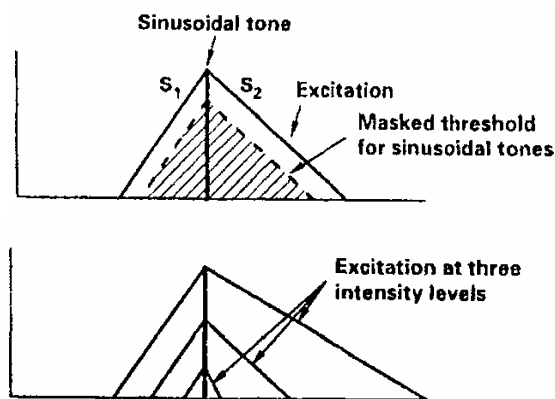


Figura 7. Modelo de enmascaramiento. (Kahrs y Brandenburg, 2002).

Terhardt (1979), citado en Kahrs y Brandenburg (2002), muestra un modelo matemático para bandas críticas, como el mostrado en la ecuación (4)

$$\begin{aligned}
 S_1 &= 31 \text{ dB/Bark} \quad ; \quad f_M > f_T \\
 S_2 &= (22 + \min(230/f, 10) - 0,2 \times L) \text{ dB/Bark} \quad ; \quad f_M \leq f_T
 \end{aligned}
 \tag{4}$$

donde,

$f$  : frecuencia en Hz

$f_M$  : frecuencia tono enmascarador en Hz.

$f_T$  : frecuencia en banda crítica en Hz.

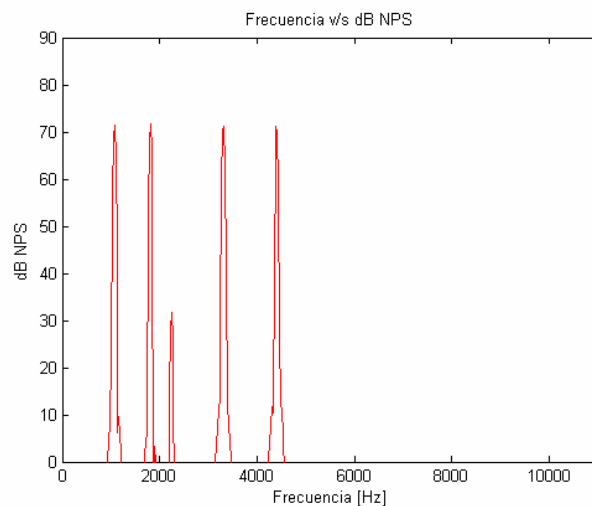
$L$  : nivel de tono enmascarador en dB NPS

La ecuación anterior considera la variación en la pendiente superior (S2) con el nivel y frecuencia, tal como se evidencia en la figura 8 inferior. Mientras la pendiente inferior S1 permanece constante.

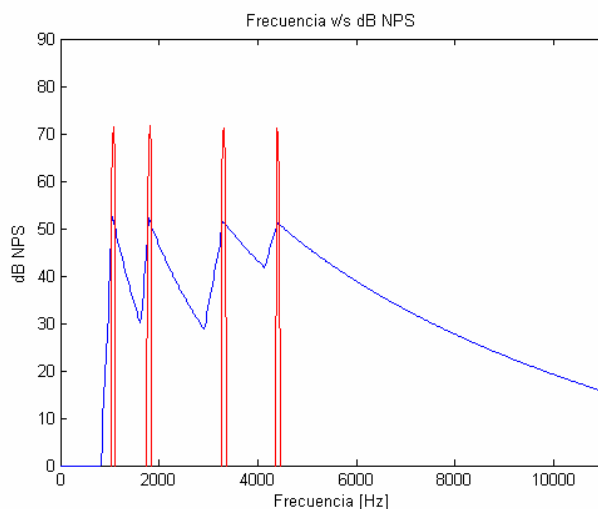
### 3 Implementación

Ambos modelos, de sonoridad y enmascaramiento son implementados como funciones en Matlab. Para la determinación del espectro de frecuencias se utilizó la transformada rápida de Fourier (FFT), con 4096 puntos, 22050 Hz de frecuencia de muestreo y ventana Blackmann.

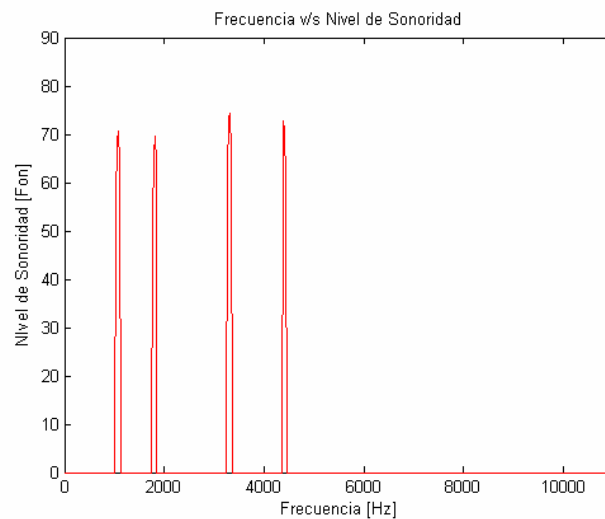
Debido a que la distorsión genera frecuencias armónicas al tono de prueba, se necesita construir un umbral de enmascaramiento global de amplio espectro. Para esto se superponen las bandas críticas sobre cada tono y se mantienen los valores máximos de éstas. No se considera el fenómeno de “adición” de enmascaramiento simultáneo en las zonas de cruce. Para determinar las pendientes S1 y S2 se transformó la escala de frecuencias en Hz a Bark, usando la ecuación (3). Luego se determinan las pendientes de la banda crítica asociada y se transforma de Bark a Hz, según ecuación (4). Esto mostrará que las pendientes tendrán una leve curva debido a esta transformación. Las figuras 8 y 9 muestran un ejemplo de tonos y sus bandas críticas. Notar que uno de los tonos desaparece debido a enmascaramiento, y que la diferencia tono enmascarador a banda crítica es cercana a 20 dB (Zwicker y Fastl, 2007).



**Figura 8.** Señal de prueba con varios tonos.



**Figura 9.** Señal de prueba y su curva de enmascaramiento. Notar que 2000 Hz desaparece.



**Figura 10.** Señal de prueba ponderada en fon.

Determinado el umbral global, y la eliminación de información psicoacústica irrelevante según enmascaramiento simultáneo, se muestra en la figura 10 la ponderación de cada NPS a su respectivo valor de nivel de sonoridad en fones, usando el modelo de sonoridad en base a Redes Neuronales Artificiales de Espinoza, Venegas y Floody, (2006). Finalmente, se calcula el nuevo estimador usando el mismo enfoque utilizado en la ecuación (1). El índice propuesto se llamará:  $\%THD_{mask+phon}$ , y se calcula usando la ecuación (5).

$$\%THD_{mask+phon} = 100 \times \sqrt{\frac{\sum_{i=2}^n h_i^2}{h_1^2}} \quad (5)$$

Donde,  $h_i^2 = 10^{\frac{phon_i}{10}}$  y  $phon_i$ , corresponden a las cantidades del  $i$ -ésimo armónico, respectivamente.

#### 4 Ejemplos

Para ejemplificar el uso del nuevo estimador, se realizaron mediciones en la sala anecoica de la Universidad Tecnológica de Chile INACAP, Sede Pérez Rosales. Keele (1973) y Struck y Temme (1992), describen procedimientos útiles para la medición en ambientes no anecoicos.

Se evaluaron 2 altavoces, cuyos datos aparecen en Tabla 1.

**Tabla 1.** Datos Altavoces.

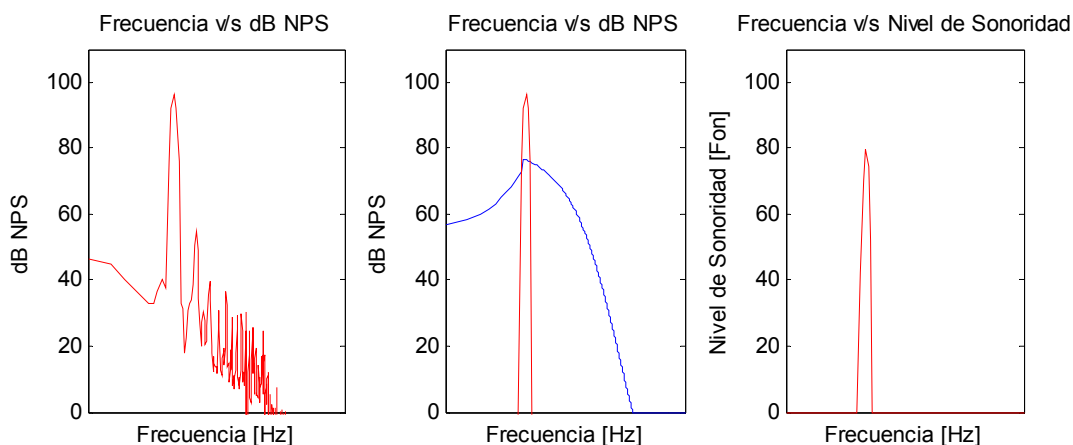
| Altavoz   | NPS | frecuencia Hz | %THD | %THDmask+phon |
|-----------|-----|---------------|------|---------------|
| Altavoz 1 | 96  | 63            | 1.6  | 0             |
| Altavoz 2 | 87  | 120           | 1.9  | 0.81          |



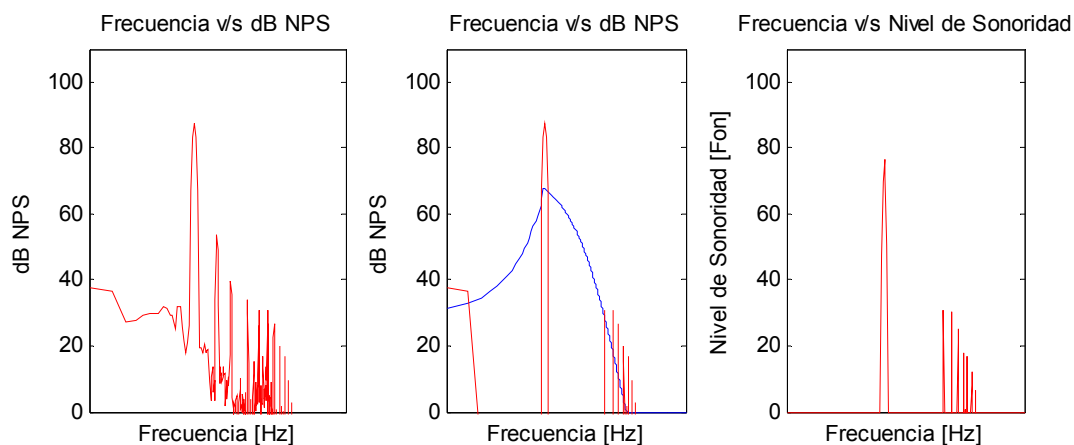
El altavoz 1 es un dispositivo bien construido, mientras el altavoz 2 tiene su bobina, spider y suspensión, desalineada del centro del motor magnético.

Los valores del estimador  $\%THD_{mask+phon}$ , evidencian que la distorsión es más audible en el altavoz 2.

A continuación se muestran las gráficas de ambos altavoces. Para claridad, se usa frecuencia en escala logarítmica.



**Figura 11.** Espectro de frecuencias Altavoz 1



**Figura 12.** Espectro de frecuencias Altavoz 1, y su curva de enmascaramiento.

## 5 Conclusiones

Se ha presentado un nuevo estimador de distorsión utilizando características psicoacústicas del oído.

La implementación se realizó utilizando un modelo de sonoridad y enmascaramiento simultáneo; características muy conocidas y que ponderan la audibilidad de los sonidos tonales estacionarios.

Es necesario establecer, con mayor cantidad de mediciones, cuales son los valores umbrales de  $\%THD_{mask+phon}$ , que determinan si el altavoz es juzgado subjetivamente como “Bueno” o Malo”.

El ejemplo presentado muestra que a  $\%THD$  similares, el nuevo estimador puede distinguir cual altavoz podría tener distorsión audible.

### Referencias

- Temme, Steve (2000); “Are You Shipping Defective Loudspeakers to Your Customers? ”, Presented at the ALMA 2000 Winter Symposium, Las Vegas, Nevada, USA.
- Zwicker, Eberhard; Fastl, Hugo (2007). “Psychoacoustics-Facts and Models”, Springer-Verlag, Berlin-Heidelberg, Germany.
- Kahrs, Mark; Brandenburg, Karlheinz (Editors) (2002). “Applications of Digital Signal Processing to Audio and Acoustics”, Kluwer Academic Publishers, New York, USA.
- Suzuki, Y. and Takeshima, H. (2004). “Equal-loudness-level contours for pure tones”, Journal of Acoustics Society of America, 116, 918–933.
- Espinoza, Victor; Venegas, Rodolfo; Floody, Sergio (2006). “Modelo de Sonoridad Usando Redes Neuronales”(A027). V Congreso Iberoamericano de Acústica, Santiago, Chile.
- Klippel, Wolfgang (2006); “Tutorial: Loudspeaker Nonlinearities—Causes, Parameters, Symptoms”. Journal of the Audio Engineering Society (JAES), Volume 54, Número 10, 907-039.
- Temme, Steve “Brüel & Kjær Application Note: Audio Distortion Measurements”. This application note is based on a paper presented at the AES 11 th International Audio Test and Measurement Conference, Portland, Oregon, U.S.A., May 31, 1992.
- Foley, Dan; Celmer, Robert; Sachwald, Benjamin; Anthony, James, Pagliaro, Tony; Thompson, Shane (2004), “Higher Order Harmonic Signature Analysis for Loudspeaker Defect Detection”, Presented at the 117th Convention of the Audio Engineering Society, 2004, October 28–3, 1 San Francisco, CA, USA
- Struck, Christopher J.; Temme, Steve (1992); “Simulated Free Field Measurements”, Presented at the 93rd Convention of the Audio Engineering Society, 1992, October 1-4, San Francisco, CA, USA.
- Keele, D.B. Jr. (1973); "Low-Frequency Loudspeaker Assessment by Near-Field Sound Pressure Measurement" (1973), Presented at the 45rd Convention of the Audio Engineering Society, 1973, May 15, Los Angeles, CA, USA.
- Beranek, Leo (1986); “Acoustics”; American Institute of Physics for the Acoustical Society of America, Cambridge, MA. USA.
- ISO 226:2003, Acoustics, “Normal equal-loudness-level contours”
- Metzler, Bob (2005) “Audio Measurement Handbook”, Audio Precision, Oregon, USA.
- Terhardt, E. (1979). “Calculating virtual pitch”. Hearing Research, 1:155–182.