



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

FIA2008-A047

## Un análisis alternativo de la proyección de la voz cantada

Cecilia Grisel Arone<sup>(a)</sup>,  
María Victoria Bongiovanni<sup>(b)</sup>,  
Alejandro de la Torre<sup>(b)</sup>,  
Gustavo Sánchez Sarmiento<sup>(b,c)</sup>.

(a) Universidad Tecnológica Nacional, Facultad Regional Haedo. París 532, Haedo, Argentina.

(b) Universidad Austral, Facultad de Ingeniería. Av. Juan de Garay 135, Buenos Aires, Argentina.

(c) Universidad de Buenos Aires, Facultad de Ingeniería. Av. Paseo Colón 850, Buenos Aires, Argentina.

### Abstract

A peak of energy about the frequency of 3 kHz, due to the cluster of superior formants, is observed in the male lyrical trained voices. This peak is called as "singer formant". In the present paper, authors propose a new method for finding the way in which the sound energy is distributed as a function of frequency. The energy of superior formants (EFS) is calculated as the ratio between the area below the curve of power vs. frequency, from 2.5 to 3.3 kHz and the same area from 0 to 4 kHz. Calculations are performed for several samples. The EFS coefficient has been measured for the vocals 'a' and 'i', being the last one systematically greater than the first one. EFS for trained singers results, in general, greater than for untrained ones. The biggest concentration of energy for nasal resonator was obtained. No differences between high and abdominal-diaphragmatic-rib breathing were found. To sing standing up not always allows major projection of the sound with respect to singing seated. Finally, significant differences between the mezzo-forte and forte were not observed, neither between the laid down position with respect to other positions.

### Resumen

En las voces masculinas, con entrenamiento en la técnica lírica, se observa un pico de energía alrededor de la frecuencia 3 kHz llamado formante del cantante, debido al *cluster* de los formantes superiores. El presente trabajo propone un nuevo método para hallar la forma en que se distribuye la energía sonora en función de la frecuencia. Se calcula la energía de los formantes superiores (EFS) a partir del cociente entre las áreas subtendidas por la curva de potencia en función de la frecuencia, de 2.5 a 3.3 kHz y de 0 a 4 kHz, para cada una de las muestras. El EFS se midió para las vocales "a" e "i", resultando para esta última sistemáticamente mayor respecto de la primera. Se obtuvo un valor de EFS para cantantes entrenados mayor que para no entrenados. Se obtuvo mayor concentración de energía para los casos de resonador nasal. No se detectó diferencias entre respiración alta y costo-diafragmática-abdominal. Se observó que cantar parado no siempre permite mayor proyección del sonido con respecto a cantar sentado. Finalmente, no se observaron diferencias significativas entre las sonoridades *mezzo-forte* y *forte*, ni entre la posición acostada con respecto a las demás posiciones.

## 1 Introducción

Los cantantes con entrenamiento en la técnica clásica deben proyectar su voz de tal manera que sobrepase el sonido de una orquesta sin realizar esfuerzo vocal. Dicha proyección de la voz se logra con el *apoggio*<sup>1</sup> diafragmático-abdominal. Acústicamente esta calidad sonora se debe al denominado formante del cantante, el cual es un pico prominente en el espectro cercano a los 3 kHz, que aparece en las voces de bajos, barítonos, tenores y contraltos, entrenados en la formación lírica.

El formante del cantante es un fenómeno resonatorio producido por el *cluster* de los formantes superiores (F3, F4 y F5) [1]. El rango de frecuencias del mismo es de 2,5 a 3,3 kHz [2] [3]. En este agrupamiento, F4 corresponde a una resonancia cercana asociada con el tubo laríngeo. La situación acústica que explica el formante del cantante, es que el tubo laríngeo sirve como un resonador cuasi-autónomo, con una frecuencia de resonancia en la vecindad de los 3 kHz y que no está influenciado por el resto del tracto vocal [4]. Esas frecuencias son cercanas a la resonancia del conducto auditivo y corresponden a la zona de máxima sensibilidad del oído.

El *LTAS* (Long - Term Average Spectrum) es aplicado para representar las características tímbricas y vocales. Es una medida en decibeles de un promedio en el tiempo de la potencia de la señal acústica para cada frecuencia. Cuando el cantante canta con una orquesta, el *LTAS* muestra un pico cercano a los 500 Hz y cae a razón de 9 dB/octava. Un pico claramente marcado en la región de frecuencias del formante del cantante puede ser seguramente atribuido a la voz cantada [2].

En las curvas de *LTAS* se puede encontrar diferencias entre voces masculinas y femeninas, cantantes y hablantes, voz solista y coral, cantantes de ópera y de otros estilos, basados en los cambios de energía necesarios para esas diferentes tomas vocales.

Existen diversos parámetros para medir cuantitativamente la calidad de la voz cantada basado en el análisis espectral. Para evaluar un *LTAS* se puede reducir la información que contiene a un número significativo, computando la razón a bajas y altas bandas de frecuencia.

Omori y otros [5] tomaron dos picos de armónicos entre los 0-2 kHz (llamados P1 y P2) y otros dos entre 2-4 kHz (P3 y P4). Estos picos corresponden al primero, segundo, tercero y cuarto formantes. El *SPR* (Singing Power Ratio), calculado en decibeles, es la diferencia de potencia de P1 y P4. Se demostró que los cantantes entrenados tuvieron un mayor *SPR* que los cantantes no entrenados y los no cantantes [5].

Otro parámetro a tener en cuenta es el *ER* (Energy Ratio), es el cálculo de la razón  $P_{hi}/P_{lo}$ . Donde  $P_{lo}$  es la potencia en la banda de frecuencias de 0-2 kHz y  $P_{hi}$  es la potencia en la banda de frecuencias de 2-4 kHz. Cuando los cantantes emplean una gran proyección hay un incremento significativo en la potencia acústica casi enteramente en la banda de altas frecuencias [6].

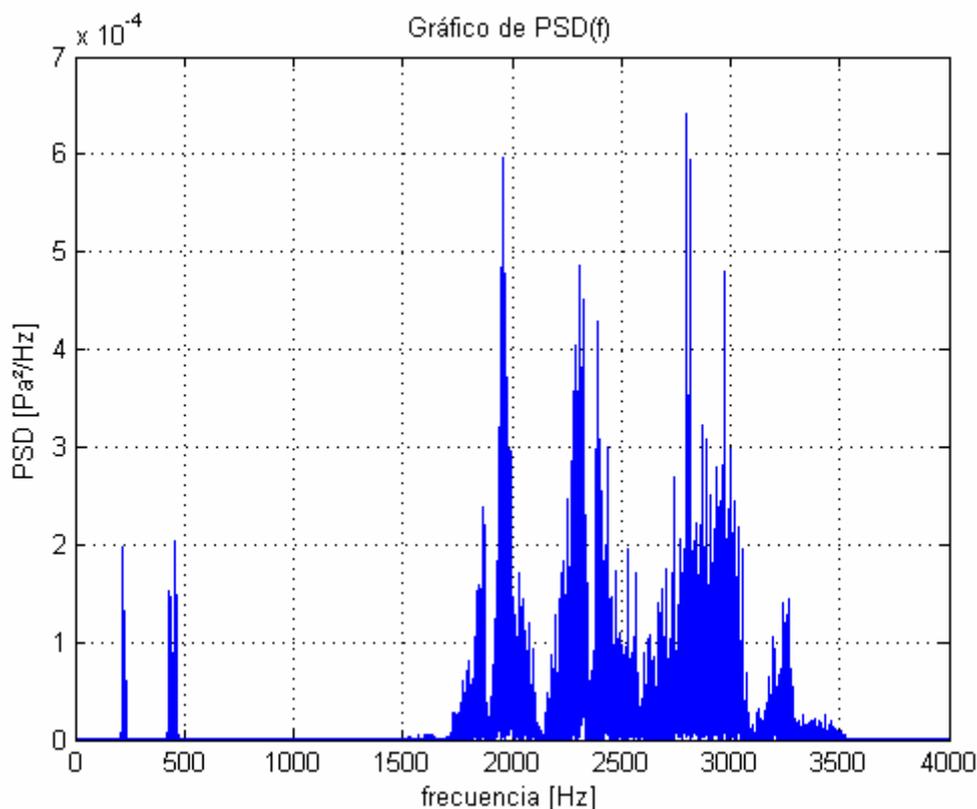
Kenny y Mitchell [7] realizaron un estudio sobre la relación se *SPR* y *ER* y la percepción de un grupo de expertos que juzgaron la técnica y belleza del canto respecto de la calidad vocal. Ellas argumentaron que la distribución de energía sobre los 2 kHz puede indicar una buena calidad de canto clásico y de ópera pero no define la calidad vocal en su totalidad.

El presente trabajo propone un nuevo método para hallar la forma en que se distribuye la energía sonora en función de la frecuencia. Se calcula la energía de los formantes

---

<sup>1</sup> *Apoggio*: Cantar sustentando el aire por los músculos abdominales sin permitir el abandono de los mismos durante la emisión.

superiores (*EFS*) [8] a partir del cociente entre las áreas subtendidas por la curva de densidad de potencia espectral (*PSD*) en función de la frecuencia, de 2,5 a 3,3 kHz y de 0 a 4 kHz, para cada una de las muestras, en cantantes masculinos. Se puede observar en la figura 1 un gráfico de *PSD* en función de la frecuencia. El *EFS* para la figura 1 es 0.4396.



**Figura 1.** Gráfico de *PSD* [Pa<sup>2</sup>/Hz] en función de la frecuencia para un cantante lírico profesional.

## 2 Metodología

### 2.1 Diseño del experimento

El diseño del experimento se puede dividir en tres partes: según el tipo de respiración, el tipo de posición y según los resonadores.

Cada una de estas partes consistió en emitir la vocal “a” e “i” en sonoridad *mezzo-forte* y *forte* (es decir, aproximadamente al 50 y 75 % del volumen sonoro que el cantante posee respectivamente). Estas vocales se emitieron a una determinada frecuencia fundamental que varía según la clasificación vocal. Los tenores cantaron un Re4 (294 Hz), los barítonos y bajos un La3 (220 Hz).

Los parámetros que se variaron para cada tipo de respiración son: alta y costo-diafragmática-abdominal. Para el tipo de posición: parada, sentada y acostada. Para el tipo de resonador: en máscara y nasal.

### 2.2 Instrumentos para la grabación

Los sonidos fueron grabados digitalmente con el programa “PRAAT doing phonetics by computer”, versión 4.6.06. Con una frecuencia de muestreo de 44100 Hz, en un canal mono a

16 bit. El micrófono que se utilizó es LEEA LE89A/B, tipo dinámico unidireccional cardioide, un rango de frecuencia de 70 Hz - 15 kHz, una sensibilidad de -78.5 dB y una impedancia de 500 Ohm.

Las muestras fueron tomadas mediante una PC con placa de sonido.

La sala en la que se hizo la grabación fue la misma para todos los cantantes.

## 2.3 Sujetos

Participaron tres tenores, dos barítonos y dos bajos. El rango de edades es de 22-52 años. El período de entrenamiento de los cantantes líricos profesionales varía entre 10 a 30 años y de los no profesionales de 5. Entre los participantes también se encontraba un estudiante de canto popular y coral de 3 años de estudio, sin técnica de canto lírico (cantante 3) y otro que no poseía ningún conocimiento de técnica (cantante 6). Al momento de la grabación, todos los cantantes gozaban de buena salud vocal. Los mismos cantaron dentro de su tesitura, en una fonación normal, con un color claro de voz para todas las muestras.

## 2.4 Análisis de datos

Una vez que se tomaron las muestras se editaron a través del programa Sound Forge versión 8.0b. La edición de cada sonido se realizó tomando la parte más estable del sonido cantado. Una vez realizado ello, se normalizó por el método RMS para poder comparar los espectros sonoros de todos los archivos.

El  $PSD_{dB}(f)$  fue obtenido para cada uno de los archivos con el programa PRAAT [9].

A partir de estos datos, se obtuvo el valor de  $PSD$  en  $Pa^2/Hz$  y se lo graficó en función de la frecuencia. A fin de facilitar el análisis comparativo de las muestras, se definió un coeficiente que posibilita cuantificar la energía obtenida en cada muestra en la zona del formante del cantante. Al mismo se lo llamó  $EFS$ , para diferenciarlo de los  $SPR$  y  $ER$  de la bibliografía. El mismo consiste en la razón de la energía bajo la gráfica de  $PSD$  vs frecuencia en el rango de 2,5 a 3,3 kHz sobre el rango total considerado, de 0 a 4 kHz. La elección del primer rango fue de acuerdo a donde se encuentra el formante del cantante [3] y del rango total de energía que se emite fue de acuerdo a Omori [5]. De esta forma se espera encontrar un mayor valor del coeficiente en los casos en que el formante del cantante esté más desarrollado.

## 3 Resultados

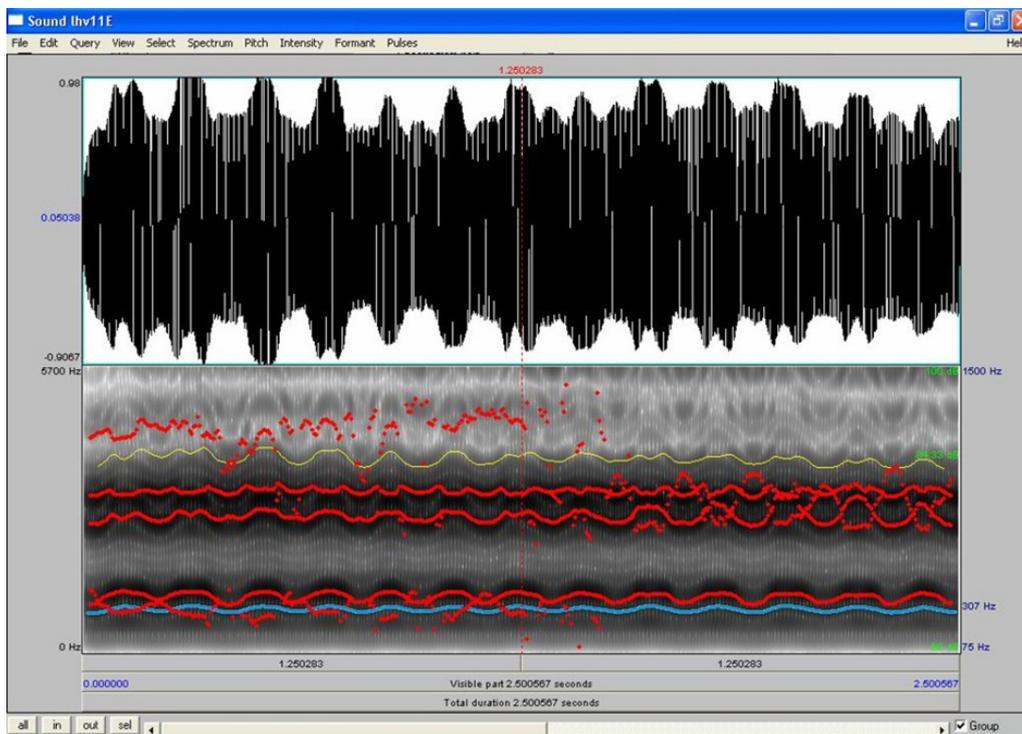
### 3.1 Análisis de los espectrogramas

En las figuras 2 (a) y (b) se observan los espectrogramas correspondientes a los cantantes 2 y 6, respectivamente. El cantante 2 es un cantante entrenado no profesional con una muy buena técnica lírica. El cantante 6 es un cantante *amateur* sin estudio de técnica. En la figura 2 (a) se puede percibir claramente las oscilaciones correspondientes al vibrato<sup>2</sup>. Además se observan zonas concentradas de energía alrededor de los 1000 y 3000 Hz, cuyos valores se ubican en el rango del primer formante y formante del cantante, respectivamente. Se puede observar el *cluster* de los formantes superiores en las líneas rojas. En cambio, en la figura 2 (b), hay zonas de menor intensidad pues las tonalidades de grises no son muy oscuras y se encuentran distribuidas más uniformemente dentro de todo el espectro. Los formantes 3 y

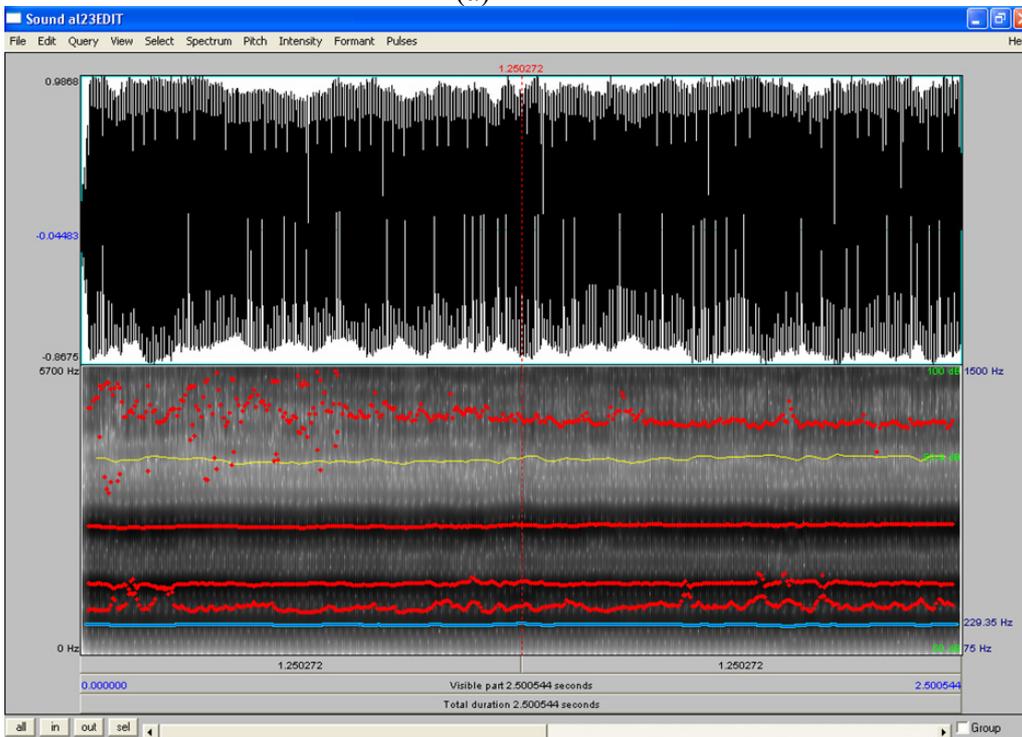
---

<sup>2</sup> Vibrato: Ondulación del sonido producida por una vibración ligera del tono.

4 se encuentran muy espaciados entre sí y no se observa un quinto formante. La separación de éstos está indicando que esta voz no posee formante del cantante. Tampoco se observa vibrato.



(a)



(b)

**Figura 2.** Espectrograma para la vocal “a” en sonoridad fuerte, posición sentado, respiración completa, correspondiente al (a) cantante 2 (b) cantante 6.

### 3.2 Análisis del EFS (Energía de los formantes superiores)

#### 3.2.1 Todos los parámetros con la vocal “a” e “i”

Se realizó el análisis de los datos según los siguientes parámetros:

En el eje horizontal:

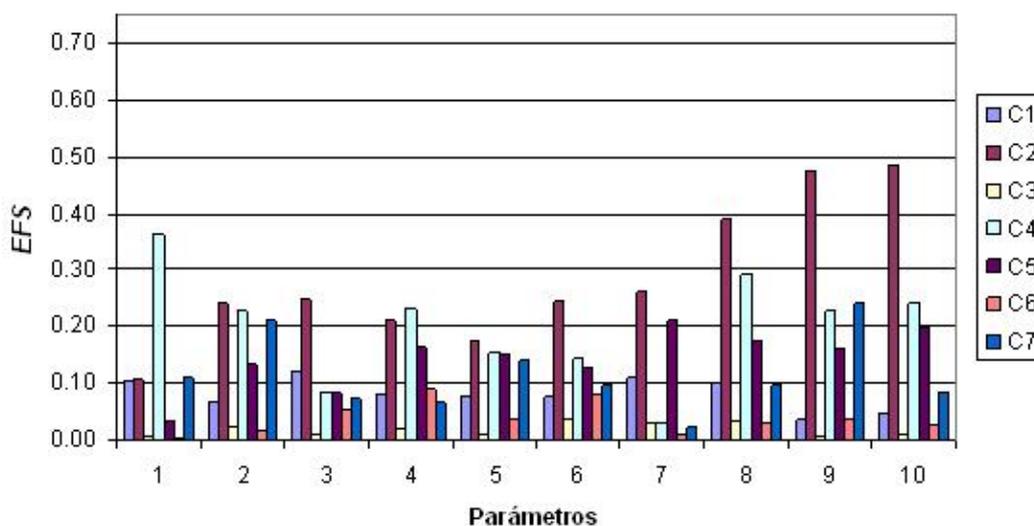
1. Respiración alta, parado, resonador en máscara, *mf*
2. Respiración alta, parado, resonador en máscara, *f*
3. Respiración costo-diafragmática-abdominal, parado, resonador en máscara, *mf*
4. Respiración costo-diafragmática-abdominal, parado, resonador en máscara, *f*
5. Respiración costo-diafragmática-abdominal, acostado, resonador en máscara, *mf*
6. Respiración costo-diafragmática-abdominal, acostado, resonador en máscara, *f*
7. Respiración costo-diafragmática-abdominal, sentado, resonador en máscara, *mf*
8. Respiración costo-diafragmática-abdominal, sentado, resonador en máscara, *f*
9. Respiración costo-diafragmática-abdominal, parado, resonador nasal, *mf*
10. Respiración costo-diafragmática-abdominal, parado, resonador nasal, *f*

Cantantes:

- C1. Tenor 1
- C2. Tenor 2
- C3. Tenor 3 no profesional
- C4. Barítono 1
- C5. Barítono 2
- C6. Bajo 1 no profesional
- C7. Bajo 2

En las figuras 3 y 4 se observa que la vocal “i” es más energética que la “a”. Este resultado es coherente con el que encuentra Fernández de la Gándara [10], en el que indica que cerca de la zona de los 3 kHz se concentra mayor energía en la vocal “i” que en la “a”.

**EFS para la vocal “a”**



**Figura 3.** Todas las muestras con la vocal “a”.

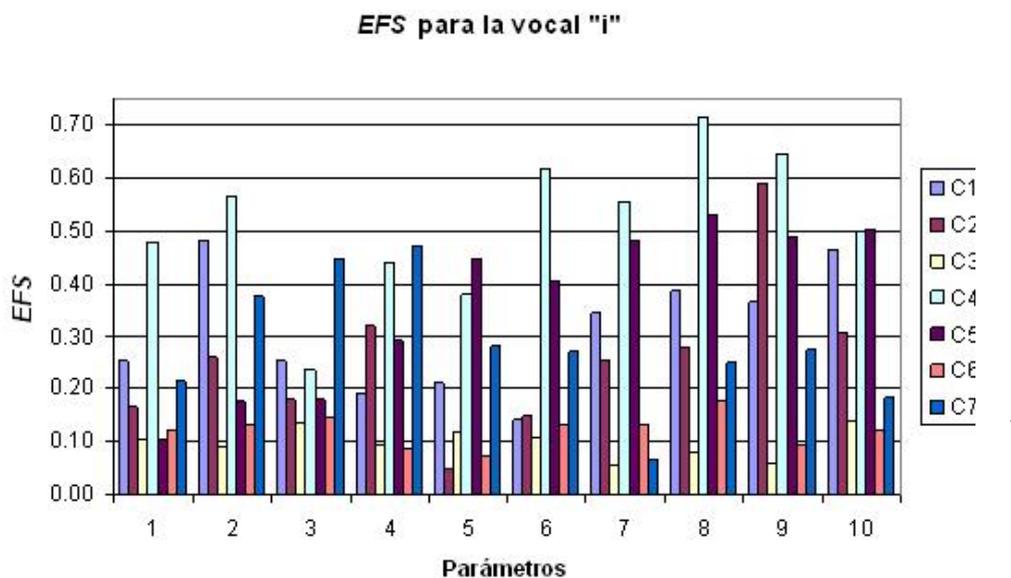


Figura 4. Todas las muestras con la vocal "i".

### 3.2.2 Tipos de respiración

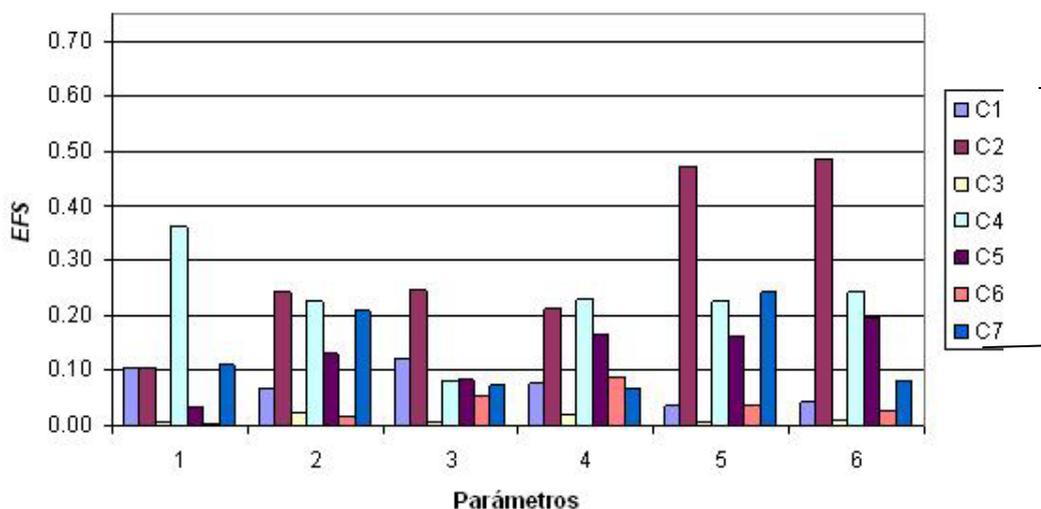
Los parámetros a tener en cuenta del eje horizontal son:

1. Respiración alta, parado, resonador en máscara, *mf*
2. Respiración alta, parado, resonador en máscara, *f*
3. Respiración costo-diafragmática-abdominal, parado, resonador en máscara, *mf*
4. Respiración costo-diafragmática-abdominal, parado, resonador en máscara, *f*
5. Respiración costo-diafragmática-abdominal, parado, resonador nasal, *mf*
6. Respiración costo-diafragmática-abdominal, parado, resonador nasal, *f*

Se puede ver en la figura 5, que los dos hombres que no realizaron estudio operístico exhiben una energía mucho más baja que los que sí tienen. Según está demostrado, los cantantes no clásicos no utilizan el formante del cantante y por lo tanto no poseen una mayor cantidad de energía en ese rango [3] [7].

La emisión nasal presenta más energía que la ubicada en la máscara (ambas con respiración costo-diafragmática-abdominal). Al ser nasal se baja el velo del paladar. Esto disminuye las cavidades de resonancia porque el aire no pasa a los senos superiores. El mismo flujo de aire debe pasar por un área más chica lo cual aumenta la velocidad de éste. Por eso se logra la mayor concentración de energía.

De este gráfico no se observaba diferencia entre respiración alta y costo-diafragmática-abdominal. Pero sí se observaba diferencia en calidad sonora en beneficio de la última.

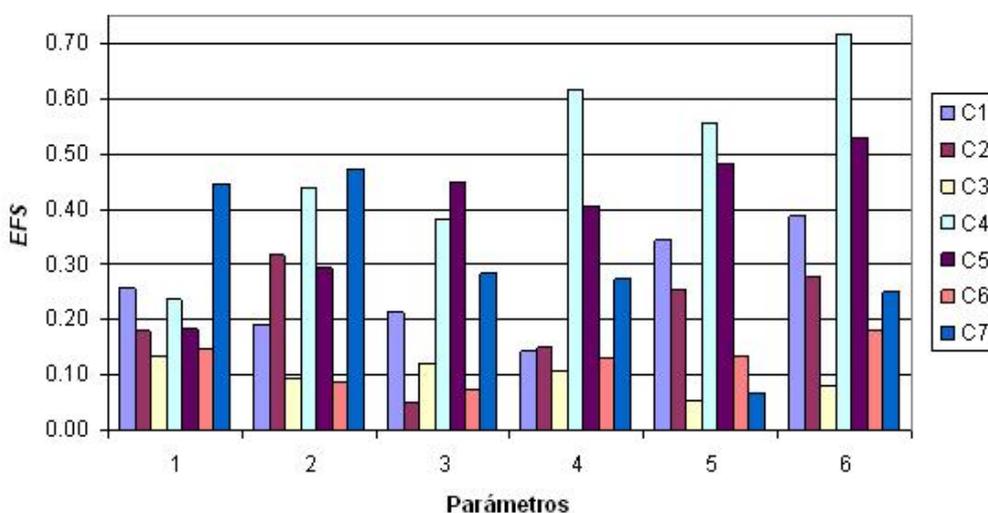
**EFS para la vocal "a" según el tipo de respiración y resonador**

**Figura 5.** Espectrograma para la vocal "a" para una respiración alta, costo-diafragmática-abdominal y con resonador nasal, las tres muestras se realizan de pie.

**3.2.3 Tipos de posición**

Los parámetros en el eje horizontal son:

1. Respiración costo-diafragmática-abdominal, parado, resonador en máscara, *mf*
2. Respiración costo-diafragmática-abdominal, parado, resonador en máscara, *f*
3. Respiración costo-diafragmática-abdominal, acostado, resonador en máscara, *mf*
4. Respiración costo-diafragmática-abdominal, acostado, resonador en máscara, *f*
5. Respiración costo-diafragmática-abdominal, sentado, resonador en máscara, *mf*
6. Respiración costo-diafragmática-abdominal, sentado, resonador en máscara, *f*

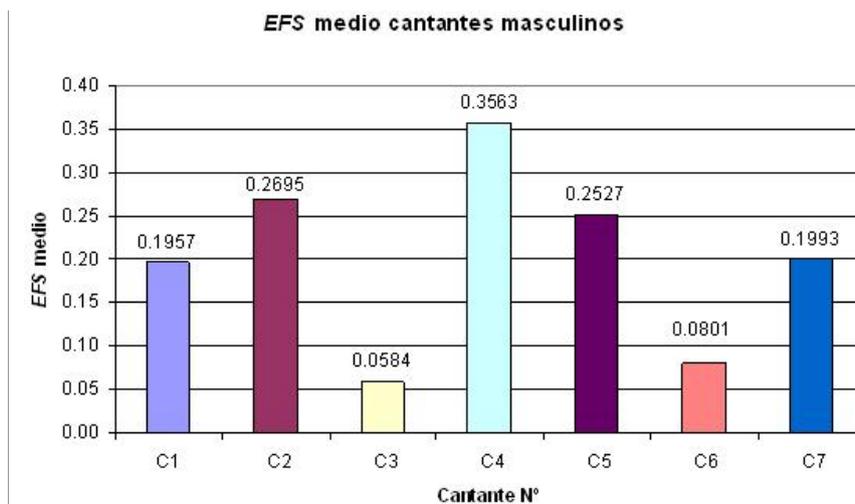
**EFS para la vocal "i" según el tipo de posición**

**Figura 6.** Espectrograma para la vocal "i" para distintas posiciones con respiración completa y de pie.

Para la figura 6 se puede indicar que en general, tanto en los *mf* como en los *f* hay más energía en la posición sentada que en las demás posiciones. Al estar sentado, el cuerpo tiene una mayor base en donde se puede sustentar mejor la columna de aire, mejorando la emisión del sonido y por lo tanto la proyección de la voz.

### 3.2.4 EFS medio para todos los cantantes masculinos

Se puede determinar un promedio de todos los coeficientes obtenidos para cada uno de los cantantes masculinos que se aprecia en la figura 7.



**Figura 7.** Promedio de todos los parámetros del *EFS* para cada uno de los cantantes masculinos.

En la misma se observa un mayor *EFS* en los cantantes entrenados, tal como se esperaba. Esto aporta evidencia a favor de la efectividad del método elegido. Dentro de los cantantes profesionales, se destaca el rendimiento del cantante 4, coincidentemente con un criterio estético al momento de tomar las muestras.

## 4 Conclusiones

Se definió un coeficiente llamado *EFS* a fin de cuantificar el grado de proyección de la voz cantada lírica. El *EFS* para cantantes entrenados dio mayor que para no entrenados. Se observó mayor energía para la vocal *i* que para la vocal *a*. Se obtuvo mayor concentración de energía para el caso de la respiración nasal. No hay diferencias entre respiración alta y costo-diafragmática-abdominal. Cantar parado no siempre permite mayor proyección del sonido con respecto a cantar sentado. En la posición acostada no se observaron diferencias significativas con respecto a las demás posiciones. Se observaron diferencias en el *EFS* entre la sonoridad *mezzo-forte* y *forte*.

La definición del coeficiente *EFS* resultó adecuada a fin de cuantificar las mediciones tomadas. Los resultados obtenidos se correspondieron con la percepción subjetiva al momento de tomar las muestras.

**Referencias**

- [1] Sundberg, Johan (1974). "Articulatory interpretation of the 'singing formant'". *The journal of the Acoustical Society of America* 55: 838-844.
- [2] Sundberg, Johan (2001). "Level and center frequency of the singer's formant". *Journal of Voice*, Vol. 15 N° 2, 176-186.
- [3] Cleveland, Thomas F.; Sundberg, Johan y Stone R. E. (Ed) (2001). "Long-Term-Average Spectrum Characteristics of Country Singers During Speaking and Singing". *Journal of Voice*, Vol. 15, N° 1, 54-60.
- [4] Sundberg, Johan (2003). "Research on the singing voice in retrospect". *TMH-QPSR*, 45(1), 011-022.
- [5] Omori, Koichi; Kacher, Ashutosh; Carroll, Linda M.; Riley, William D; Blaugrund, Stanley M. (1996). "Singing Power Ratio: Quantitative Evaluation of Singing Voice Quality". *Journal of Voice*, Vol. 10 N° 3, 228-235.
- [6] Thorpe, C. William; Cala, Stephen J.; Chapman, Janice; Davis, Pamela J. (2001). "Patterns of Breath Support in Projection of the Singing Voice". *Journal of Voice* Vol 15 N° 1, 86-104.
- [7] Kenny, Dianna T.; Mitchell, Helen F. (2005). "Acoustic and Perceptual Appraisal of Vocal Gestures in the Female Classical Voice". *Journal of Voice*. Article in press.
- [8] Arone, Cecilia G. (2007). "La voz humana cantada: Características y diferencias según los emisores. Un análisis acústico". Tesis de licenciatura. UTN Regional Haedo, Argentina.
- [9] PRAAT doing phonetics by computer", versión 4.6.06, por Paul Boersma y David Weenink, [www.praat.org](http://www.praat.org)
- [10] Fernández de la Gándara, G.; Lorente, M (1998). "Acústica Musical" Colec. Música Hispana Textos. Manuales. Public. del Instituto Complutense de Ciencias Musicales. Edic. del ICCMU. Segunda edición. 1ª parte. Madrid, España.