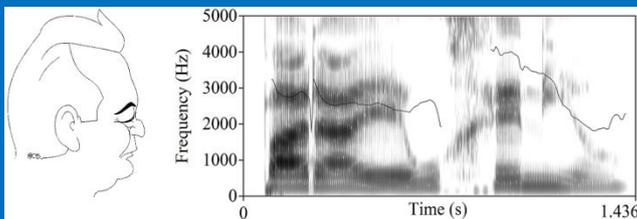


El espectrograma reasignado: creación, uso y utilidad en fonética

Wendy Elvira-García



Fernández Planas, A. Ma. (ed.) (2016): *53 reflexiones sobre aspectos de la fonética y otros temas de lingüística*, Barcelona, págs. 353-361.

ISBN: 978-84-608-9830-6.

El espectrograma reasignado: creación, uso y utilidad en fonética

Wendy Elvira-García
Universitat de Barcelona
wendyelvira@ub.edu

*Para Eugenio,
mi primer profesor de fonética.*

1. INTRODUCCIÓN

Cualquier persona que haya conocido al Dr. Martínez Celdrán, haya sido alumno o le haya escuchado hablar en los congresos, sabrá de su pasión por los espectrogramas de calidad, limpios y sin ruido de fondo. Este artículo es un homenaje a Eugenio y a su pasión por los espectrogramas, esa herramienta usada por todos los fonetistas pero al mismo tiempo a la que tan poca atención se presta.

Durante su vida académica, Eugenio Martínez Celdrán ha sido testigo de cómo el análisis espectrográfico pasaba de realizarse con el Sona-Graph 6061B de Kay, un gran equipo analógico que no permitía analizar más de 2,4 segundos y que dibujaba el espectrograma quemando el papel (figura 1), a los Digital Sona-Graph 7800 y DSP Sona-Graph 5500 grandes equipos digitales que ocupaban una mesa completa, hasta los laboratorios portátiles Kay CSL 4300 y 4500 y, en última instancia, el uso de software en su propio ordenador personal.

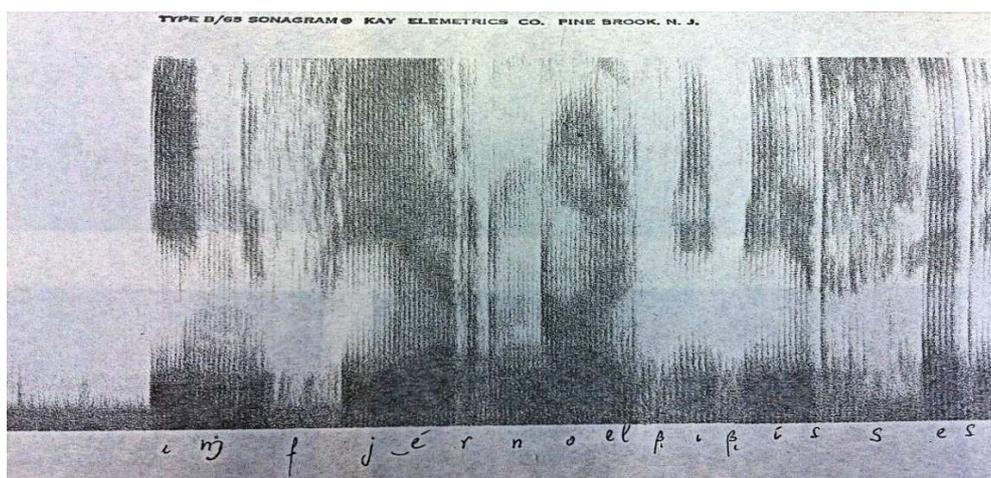
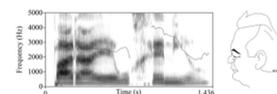


Figura 1. Espectrograma analógico realizado con Sona-Graph 6061B de Kay del Laboratori de Fonètica de la Universitat de Barcelona y anotado a mano por el Dr. Martínez Celdrán.



Sin embargo, y pese a las evidentes mejoras en la visualización de los datos, la filosofía detrás de los espectrogramas sigue siendo la misma que en la segunda guerra mundial: representar las frecuencias (Koenig, Dunn, y Lacy, 1946). A esto se refieren (Fulop y Fitz, 2006) cuando presentan su artículo llamado *A spectrogram for the twenty-first century* en que presentan el espectrograma reasignado.

Este trabajo es una presentación mediante un breve estado de la cuestión de esta técnica de espectrografía. El trabajo está dividido como sigue. En la siguiente sección, 2, se presenta la historia y se describe la técnica del espectrograma reasignado. La sección 3 se dedica a las posibilidades para el fonetista actual de realizar un espectrograma reasignado a partir de sus datos. La sección 4 trata las posibilidades de aplicación de la técnica al habla. Por último, en 5, se presentan las conclusiones.

2. EL ESPECTROGRAMA REASIGNADO

Como se ha adelantado, en el artículo de Fulop y Fitz (2006) se presenta un nuevo espectrograma que ellos definen en corto como el espectrograma del siglo XXI, es el espectrograma reasignado. En ese artículo se defiende que el espectrograma reasignado es una herramienta útil capaz de reflejar la información de los espectrogramas de manera más precisa. Sin embargo, la técnica que se usa en el artículo no era nueva, había sido propuesta 30 años antes por Kodera, Villedary y Gendrin (1976). Desde entonces, de vez en cuando algún investigador reclama que se le preste la debida atención. Algunas décadas más tarde de su invención se introdujeron cambios especialmente pensados para el análisis del habla (Auger y Flandrin, 1994; Plante, Meyer y Ainsworth, 1998).

El espectrograma reasignado consiste en realizar un espectrograma en que las frecuencias visibles en el gráfico son solo las más intensas, es decir, un espectrograma en que los picos espectrales (formantes o armónicos) se ven más oscuros y el resto de frecuencias pasan a prácticamente desaparecer. Esto se consigue, como se puede deducir de su nombre, mediante la reasignación de frecuencias. La reasignación de frecuencias consiste en mover las frecuencias que han recibido menos intensidad hasta los lugares donde se ha recibido más intensidad, por lo tanto, las frecuencias más intensas se refuerzan mediante las frecuencias cercanas menos intensas. La figura 2 muestra de manera esquemática cómo funciona el proceso.

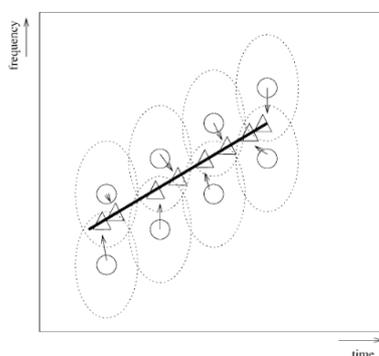


Figura 2. *Esquema del funcionamiento de la reasignación de frecuencias extraído de Plante et al. (1998).*



En un espectrograma clásico, al visualizar un sonido, el pico espectral, que sería el objeto de interés en la mayoría de casos, se ve difuminado. En la figura 3 se puede observar el espectrograma de un sonido de 440Hz (el La natural) creado de manera sintética donde resulta muy complicado determinar su frecuencia debido a que las frecuencias adyacentes al pico aparecen difuminadas de manera muy suave. Sin embargo, en el reasignado (en el centro de la figura 3) el mismo sonido se representa con una línea horizontal estrecha.

En la práctica eso hace que la determinación de, por ejemplo, la altura formántica de una vocal no se realice directamente a partir del espectrograma, sino a partir del espectro y usando, además, un método de predicción lineal (LPC) sobre este.

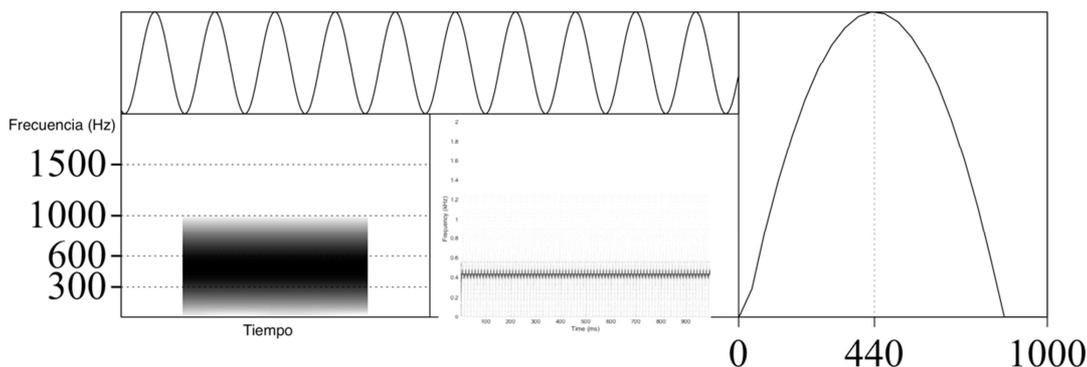


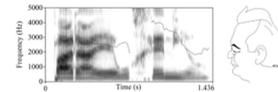
Figura 3. Oscilograma arriba y abajo espectrograma y espectrograma reasignado y a la derecha espectro de un sonido sintético equivalente al La natural. En el espectro se puede comprobar que su pico de frecuencia se encuentra a 440Hz, sin embargo el espectrograma muestra frecuencias de los 0 a los 1000Hz y una banda negra de los 300 a los 600 donde no se puede determinar el centro.

Fulop y Fitz (2006) publicaron una actualización del método especialmente pensada para el habla humana, esta es la que se sigue usando hasta la fecha. A partir de esta, Fulop (2010) muestra que la predicción de los formantes a partir de valores de vocales sintéticas se acerca más con esta técnica que con el uso de los espectros LPC, que se usan normalmente en estudios fonéticos. Específicamente, lo compara con los resultados obtenidos con Praat (Boersma y Weenink, 2015) a partir de análisis formántico LPC mediante el método de Burg.

Por lo tanto el espectrograma reasignado parece, en principio, una herramienta eficaz para paliar algunos de los problemas de análisis que presentan los espectrogramas clásicos.

3. REALIZACIÓN

La mayoría de fonetistas de hoy usan el software Praat para visualizar el espectrograma de una señal de manera rápida y fiable. Sin embargo, la técnica del espectrograma reasignado no está disponible en este momento en este *software*. Para extraer este tipo de



gráficos hay que recurrir a otro tipo de software también muy usado por fonetistas, si bien con menos asiduidad, como R o Matlab.

3.1. Espectrogramas reasignados en Matlab

Para realizar espectrogramas reasignados en Matlab se puede usar el código original desarrollado por Fitz y Fulop (2005, 2006, 2007). El paquete consta de todas las funciones necesarias para el análisis y la realización de gráficos en forma de scripts y está recogido en la siguiente página web: <http://www.cerlsoundgroup.org/Kelly/timefrequency.html>

Pero, además, Matlab incluye (a partir de la versión 2015b) el parámetro «reassigned» en su función «spectrogram». De esta manera, se pueden crear espectrogramas reasignados a partir de la función integrada en el programa sin necesidad de acudir a otros scripts.

En la figura 4 se incluye un ejemplo de código que usa esta función. El código es fácilmente adaptable a otros datos. Para ello es necesario cambiar en la segunda línea de código la ruta y el nombre de archivo del ejemplo por la ruta y el nombre de archivo que se quiere analizar, el archivo resultante con la imagen del espectrograma se guardará en la ruta indicada.

```
% Escribe en la línea siguiente tu ruta + nombre de archivo
file = '/user/Desktop/mi_wav.wav'

[wav_original, f_muestreo] = audioread(file); % Lee el audio
base = file.name(1:length(file.name)-4); % Crea una variable sin
la extensión

wav=resample(wav_original,10000, f_muestreo); % Frecuencia de
muestreo a 10000 para ver hasta 5000Hz
fs =10000;
spectrogram(wav, gausswin(0.005* f_muestreo),[], 1024,
f_muestreo, ...
'MinThreshold',-100,'yaxis'); % Crea y dibuja el espectrograma

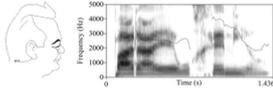
colormap(flipud(gray)); % Cambia la gama de color a escala de
grises
colorbar off; % Quita la barra de color

saveas(figSpectrogram,base,'png'); % Guarda la figura en PNG en
la misma carpeta y con el mismo nombre que el audio
```

Figura 4. Código para crear un espectrograma reasignado en Matlab2015.

3.2. Espectrogramas reasignados en R

En el programa R se puede obtener el espectrograma reasignado de archivo binario. Para ello es necesario usar (descargar e instalar) el paquete {GENEAread} (Zhou, 2014) y usar



la función `stft`. En el cuadro de la figura 5 se incluye un ejemplo de la sintaxis de la función.

```
stft (datos, start=0, end=1, length=NULL, time.format = c("auto"),
     type = c("mv", "svm", "sum"), mv.indices, date.col,
     reassign = TRUE, plot.it = FALSE,...)
```

Figura 5. Código para crear un espectrograma reasignado en R.

3. APLICACIONES

Como se ha comentado, el espectrograma reasignado tiene una ventaja clara contra el espectrograma clásico: la precisión (véase la figura 6). Es más preciso tanto en el eje frecuencial, como en el del tiempo, pero, además, permite ser más preciso al investigador, porque representa los formantes como una sola línea y no como una franja. Eso significa que es mucho más fácil dar el valor exacto de frecuencia en que se encuentra un formante.

En español su uso sería especialmente útil en los casos en que las vocales posteriores especialmente [u] muestran el primer y segundo formantes tan próximos que son difícilmente separables en el espectrograma.

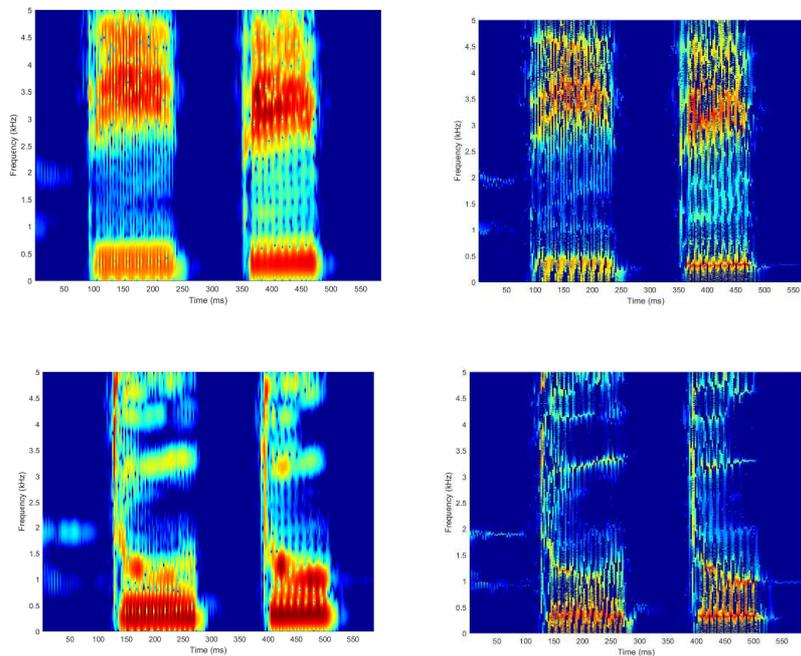
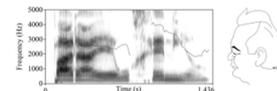


Figura 6. Ejemplos de espectrogramas clásicos (a la izquierda) y su espectrograma reasignado correspondiente (a la derecha). Arriba, la secuencia [ˈtiti]; abajo, [ˈtuɰu] pronunciadas por una hablante femenina de español septentrional en frase portadora.



Además, el espectrograma reasignado ha mostrado ser útil también en el ámbito forense, específicamente en la identificación de locutor (Fulop y Disner, 2007, 2009). Se ha afirmado que al «podar» (*prune* en el original), es decir, filtrar el espectrograma reasignado, se podría obtener una «huella vocal» más exacta que las que se obtienen con el espectrograma clásico partiendo de los pulsos glóticos. Sin embargo, un estudio más reciente (Enzinger, 2013) contrastó los valores obtenidos a partir del espectrograma reasignado con el estado del arte en fonética forense (valores obtenidos a partir de los coeficientes cepstrales mel de frecuencia –MFCC–) usando datos de triptongos del chino mandarín y comprobó que los resultados obtenidos por los dos métodos no diferían significativamente. Aunque los dos métodos no daban siempre los mismos resultados, el uso del espectrograma reasignado no hacía la herramienta más eficaz.

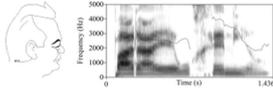
5. CONCLUSIÓN

Pese a los continuos avances tecnológicos y la revolución que ha supuesto para el mundo de la fonética contar con software de análisis acústico en los ordenadores personales, la principal herramienta del fonetista, el espectrograma, lleva inventada y sin cambios desde los años 40.

Ante este panorama el espectrograma reasignado ofrece la posibilidad de dar una mirada diferente a los datos. Sin embargo, aunque el espectrograma reasignado es una herramienta útil de visualización, cuando se trata de obtener a partir de él datos cuantitativos no supone una mejora respecto de los métodos actuales.

6. REFERENCIAS BIBLIOGRÁFICAS

- AUGER, F. y P. FLANDRIN (1994): «The why and how of time-frequency reassignment», en IEEE (ed.): *Time-Frequency and Time-Scale Analysis. Proceedings of the IEEE-SP International Symposium*, pp. 197–200.
- BOERSMA, P. y D. WEENINK (2015): *Praat: doing phonetics by computer*. <http://www.fon.hum.uva.nl/praat/> [15-12-2015]
- ENZINGER, E. (2013): «Testing the validity and reliability of forensic voice comparison based on reassigned time-frequency representations of Chinese /iau/», en T. Kalker y J. Huang (eds.): *Proceedings on the IEEE International Workshop on Information Forensics and Security (WIFS)*, pp. 13–18.
- FITZ, K. y S. FULOP (2005, 2006, 2007): *Matlab functions implementing time-frequency reassigned spectral analysis*. [Scripts de Matlab]
- FULOP, S. (2010): «Accuracy of formant measurement for synthesized vowels using the reassigned spectrogram and comparison with linear prediction», *The Journal of the Acoustical Society of America*, 127(4), pp. 2114–2117.
- FULOP, S. y K. FITZ (2006): «Algorithms for computing the time-corrected instantaneous frequency (reassigned) spectrogram, with applications», *Journal of the Acoustical Society of America*, 119, 360 – 371.
- FULOP, S. y S. DISNER (2007): «The reassigned spectrogram as a tool for voice identification», en J. Trouvain y W. Barry (eds.): *Proceedings of ICPhS XVI*, Saarbrücken, Alemania, pp. 1853–1856.



- FULOP, S. y S. DISNER (2009): «Advanced time-frequency displays applied to forensic speaker identification», *The Journal of the Acoustical Society of America*, 125, pp. 2530-2530.
- FULOP, S. y K. FITZ (2006): «A spectrogram for the twenty-first century», *Acoustics Today*, 2(3), pp. 26-33.
- KODERA, K.; C. De VILLEDARY y R. GENDRIN (1976): «A new method for the numerical analysis of non-stationary signals», *Physics of the Earth and Planetary Interiors*, 12(2), pp. 142–150.
- KOENIG, W.; H. K. DUNN y L. Y. LACY (1946): «The sound spectrograph», *Journal of the Acoustical Society of America*, 18(1), pp. 19–49.
- PLANTE, F.; G. MEYER y W. A. AINSWORTH (1998): «Improvement of speech spectrogram accuracy by the method of reassignment», *IEEE Transactions on Speech and Audio Processing*, 6(3), 282–287.
- <http://doi.org/10.1109/89.668821> [15-12-2015]
- ZHOU, F. (2014): *GENEAread* v.1.1.1. [Paquete de R]