

Santiago de Chile, 22 de mayo de 2013
Ricardo Pujol, Ingeniero en Sonido DuocUC

Análisis de audio presentado por el diputado de la República Bolivariana de Venezuela Ismael García el día 20 de mayo de 2013

Me sirvo en la libertad de realizar un análisis del audio presentado por lapatilla.com en su Soundcloud el día 20 de mayo de 2013, en mi calidad de ingeniero en sonido para esclarecer la veracidad del mismo o desmentirlo. El audio a analizar es el correspondiente al siguiente link y debería ser idéntico al presentado por el diputado Ismael García:

<https://soundcloud.com/lapatillaaudio/audio-completo#play>

Espectrograma

Uno de los gráficos que brinda más información sobre un archivo de audio es el espectrograma, el cual muestra la aparición, desaparición y desarrollo de todas las frecuencias del espectro audible. Este gráfico muestra la amplitud de cada frecuencia en función del tiempo, es decir muestra todos los parámetros que posee una forma de onda (exceptuando la fase, la cual no tiene importancia en este caso por tratarse de un audio mono y su nula influencia para comprobar la veracidad del audio) por lo cual, de ser un archivo editado o generado, existirían anomalías visibles en el gráfico.

El gráfico muestra en el eje X el tiempo, en el eje Y la frecuencia y según la intensidad del color se representa la amplitud de cada una de las frecuencias, por lo cual cada artefacto sonoro será visible en este gráfico

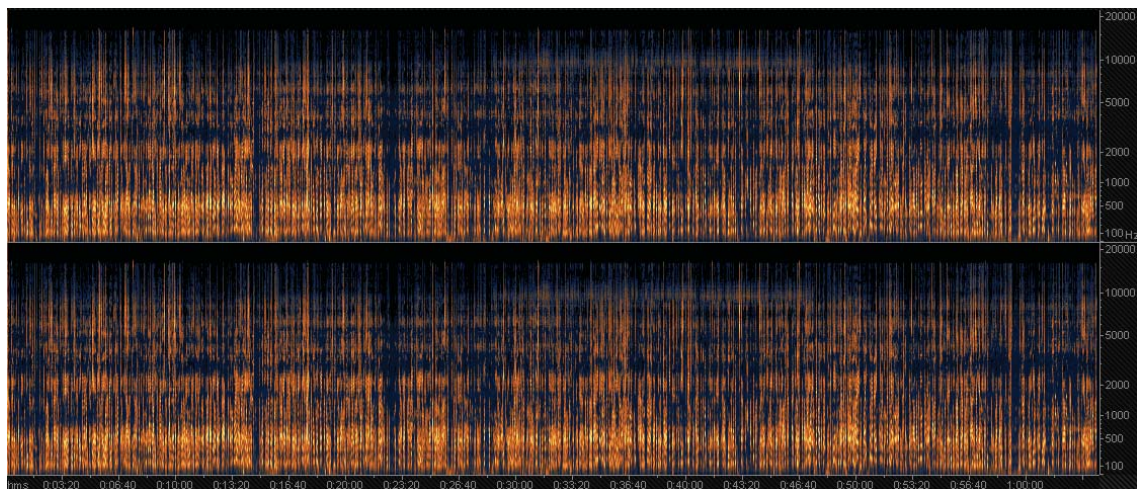


fig1: Espectrograma completo del archivo de audio.

Tratándose de un archivo de una hora de duración, totalmente coherente en su semántica lingüística y orden lógico, de ser editado existirían diferencias visibles en la composición armónica de la voz entre las partes que fueron cortadas.

El espectrograma se muestra muy uniforme, lo que sugiere que todo el audio fue grabado bajo las mismas condiciones acústicas y a la misma distancia del micrófono, cosa imposible de lograr dos veces.

Reducción de Ruido

El audio presentado tuvo que haber sido pasado por un proceso de reducción de ruido, es por esto que se escuchan algunos artefactos llamados “space monkeys” que hacen que suene poco natural. Me imagino que antes de presentar el audio quisieron realizar este proceso para aumentar la inteligibilidad de la palabra. La reducción de ruido es un proceso muy común, sobretodo cuando las condiciones acústicas no son óptimas y consiste en tomar un pedazo del audio de “silencio” (digo “silencio” porque en realidad no hay silencio sino únicamente ruido de fondo, sin palabra hablada ni sonidos como manipulación del micrófono o golpes a la mesa) y realizar un análisis espectral del mismo para luego restárselo a todo el archivo de audio, siendo la resultante únicamente la voz hablada, con ruido en menor medida pero con la aparición inevitable de los space monkeys. Mientras más reducción de ruido más presentes estarán estos artefactos.

Al escuchar (y observar) la uniformidad en el fondo de la grabación, certifico con toda seguridad y propiedad que no existen ediciones dentro del archivo de audio proviniendo en su totalidad de la misma persona, en el mismo lugar y durante un tiempo continuo.

Entonación

Una de las evidencias más notorias de un audio editado son los cambios en la entonación. La voz hablada es como una melodía (distintos acentos, distintas melodías) y al editar audio esta “melodía” se escucharía como cortada, o en casos de muy buena edición, al menos se vería en el espectrograma. En la figura 2 se ve el momento en que Mario habla de los rectores del CNE y del fraude

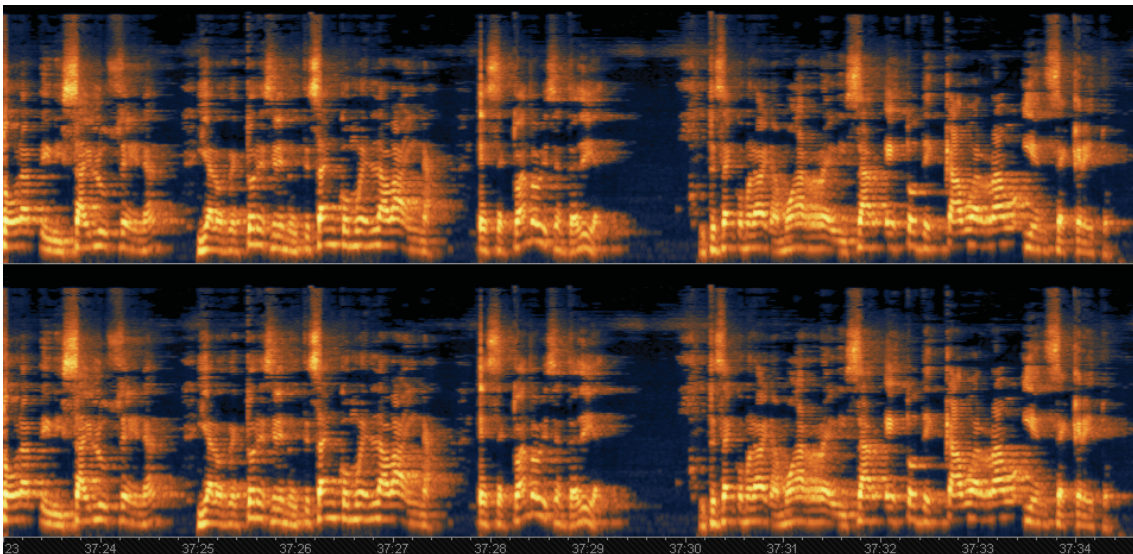


fig2: momento en que se habla de fraude

Al revisar meticulosamente se ve que la entonación es perfecta y que no hay elementos que denoten edición alguna. Incluso se ven artefactos que cruzan de forma horizontal audio lo cual denota continuidad. Lo mismo sucede con la hora entera de grabación

Reverberación

En la conversación grabada queda demostrado que Mario Silva y su interlocutor se encuentran en la misma habitación dado que la reverberación que los acompaña durante toda la grabación es auditiva y espectralmente igual (reverberación es el sonido resultante entre la interacción de la voz hablada y el recinto, común y erradamente llamado eco) Además se demuestra que la conversación no fue una intervención telefónica sino de un micrófono (posiblemente de celular) que se encontraba en la habitación. Esto queda demostrado al poseer frecuencias superiores a los 4Khz, ya que de ser una grabación telefónica se filtraría el sonido en esta frecuencia

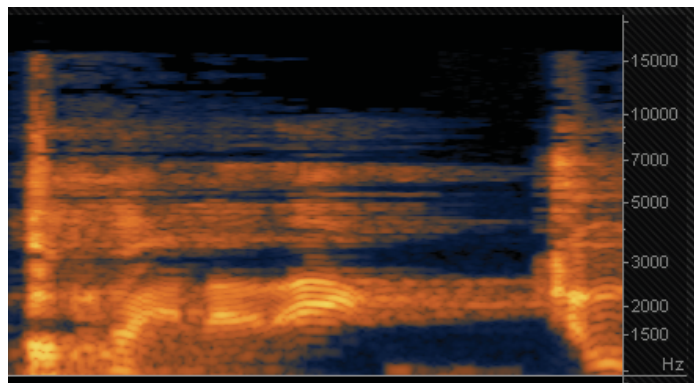


Fig3: el audio posee frecuencias hasta 15Khz

Queda demostrado también que Aramis Palacios se encuentra más distante del micrófono, no solo porque su voz se escucha a menor nivel, sino porque sus componentes armónicos comienzan a filtrarse a frecuencias inferiores a las de Mario. Voy a poner el ejemplo de que hablo por teléfono con alguien y luego hablo con esa misma persona pero a un metro de distancia del teléfono. El resultado será que mi voz tendrá menos agudos y más reverberación, cosa que sucede y queda demostrada en la figura 4.

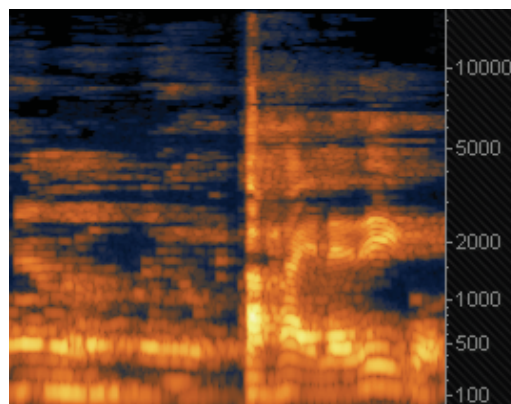


Fig4: comparación espectral entre interlocutores

Como se puede observar, el segmento de sonido a la izquierda corresponde a la voz de Aramís Palacios y el derecho a Mario Silva. La intensidad del sonido alrededor de los 500Hz es la misma para ambos, pero en frecuencias superiores Mario tiene mayor presencia.

Escuché la hora entera de audio viendo el espectrograma y no encontré ni un solo artefacto extraño que sugiriera edición. Además encontré que la entonación es perfecta en todo momento, cosa que sería imposible lograr sin lograr notorias distorsiones si fuera un collage de sonido. Además pude corroborar que el sonido del celular que suena en varias ocasiones proviene de la misma situación. Incluso pude comprobar que existe un remanente de la voz de la persona que habla con Aramís por teléfono (fig5).

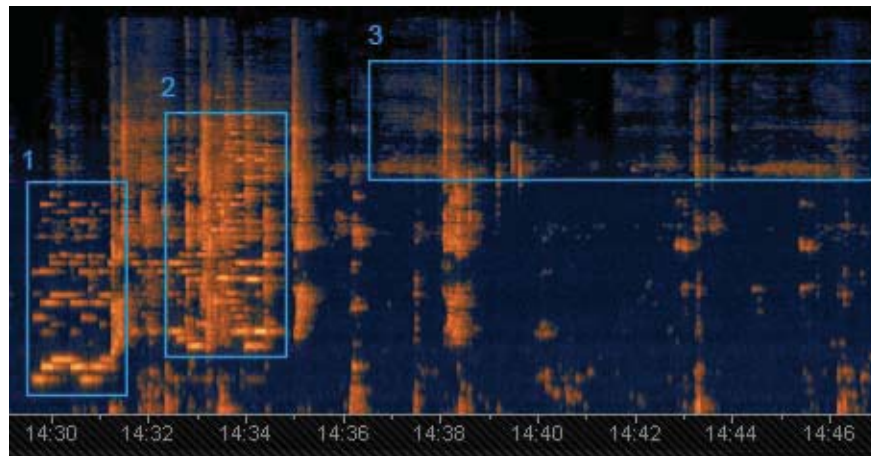


Fig5: llamada de celular minuto 14:30

En la zona 1 se ve el ringtone. En la zona 2 se ve el ringtone después de que mueve el celular (probablemente lo saca del bolsillo) por lo cual contiene armónicos en frecuencias superiores. En la zona 3 se ve la voz de la persona que habla con Aramís.

La Hojilla

Existen varias razones que me hacen tener la certeza de que este audio no proviene de las grabaciones de la hojilla, tal como explicó Mario Silva. En Televisión se utilizan los llamados micrófonos lavalier que se abrochan a la corbata o a la solapa de la camisa. El resultado de esto es que el sonido de la voz sea claro y de nivel uniforme. Si escuchamos el audio presentado por el diputado podemos escuchar claramente que el micrófono no estaba ubicado en el cuerpo de Mario Silva si no en una mesa o en otro lugar, ya que además de que no se escucha en ningún momento el roce del micrófono con la ropa (cosa que si se escucha en el programa), la voz se escucha con pequeñas variaciones de nivel y frecuencia, producto de que Mario giraba la cabeza, subía la voz o la bajaba. En la figura 6 se ve un segmento al azar del espectro sonoro de la voz de Mario Silva en su programa la Hojilla y en la figura 7 un segmento al azar de la grabación en cuestión para realizar una comparación

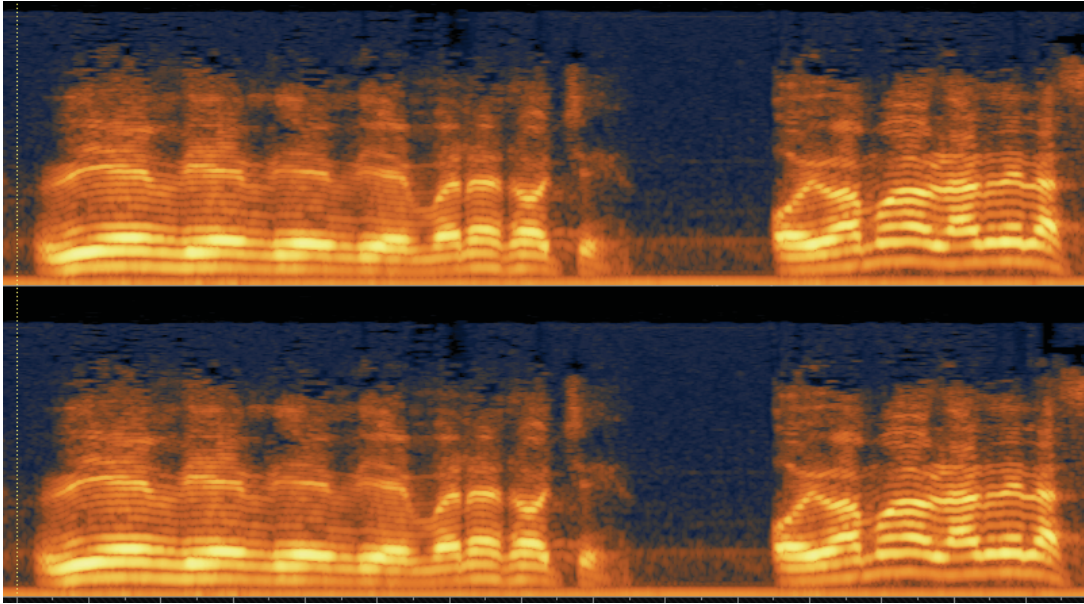


fig6: Espectrograma de la voz de Mario Silva en La Hojilla

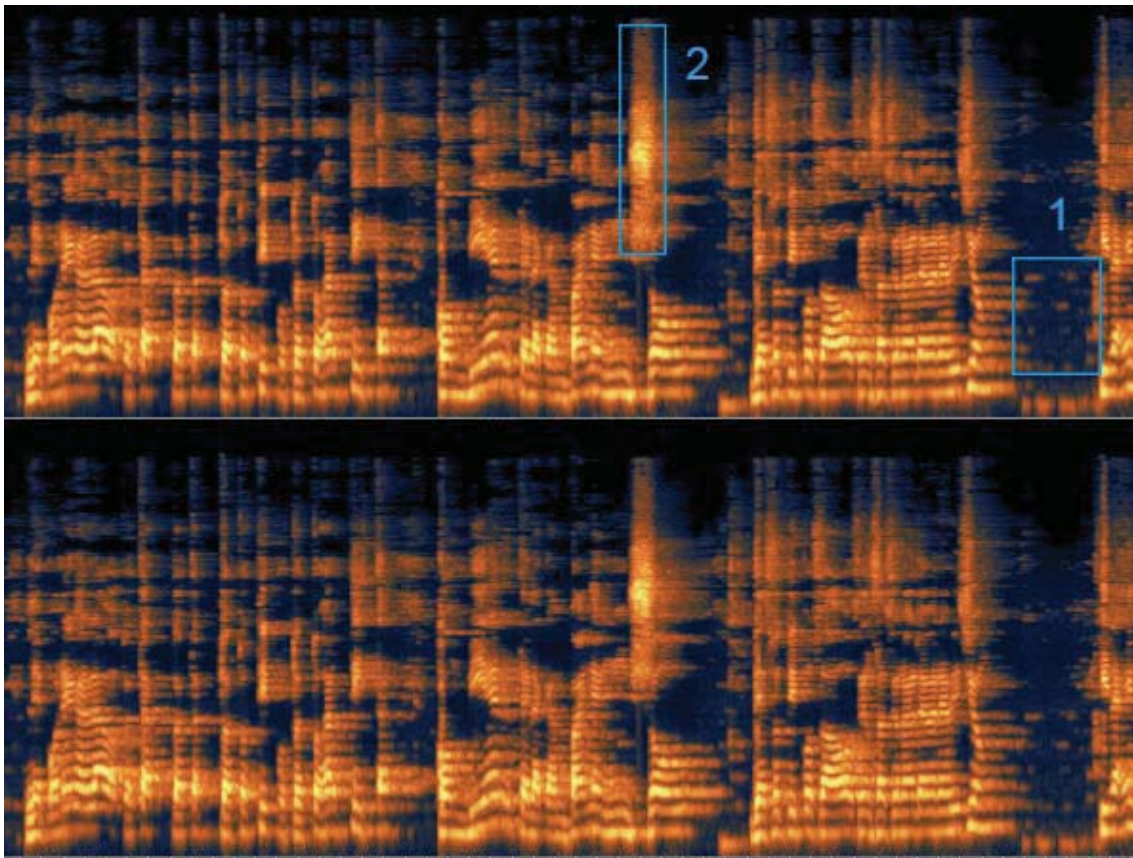


fig7: Espectrograma de la voz de Mario Silva en la grabación en cuestión

Las diferencias entre la fig6 y la fig7 son notorias. Si bien una de las diferencias más grandes es la de que el fondo en la figura 7 es más oscuro que en la fig6, esto es por la reducción de ruido, por lo cual pasa a ser una diferencia sin ningún significado para el análisis.

-Claramente las áreas coloreadas en la fig6 son más macizas que en la fig7, esto delata que la calidad del micrófono de la fig6 es mucho superior a las de la fig7.

-Otra gran diferencia, son las transientes (o el ataque) de la voz de Mario. En la fig6 son mucho más suaves y en la fig7 más agresivas. Esto se nota en las múltiples líneas verticales que se encuentran en la fig7, cualidad propia de micrófono de baja calidad y con un sistema que comprime la voz (un celular).

-Los espacios de silencio en la fig6 son uniformes, lo que demuestra que se grabó en un estudio de televisión donde se minimizan los ruidos por razones obvias. En cambio en la fig7 en los espacios de silencio (sector 1) se observan artefactos que son producto de las malas condiciones acústicas donde se grabó y de la baja calidad del micrófono.

-En el segmento 2 de la fig7 se observa una "S", al ser un micrófono de baja calidad, su respuesta de frecuencia hace que las letras S se disparen y originen ese tipo de figuras. En la figura 6 se mantiene la uniformidad en esta área del espectro, manteniéndose las S dentro del rango dinámico que sólo un micrófono profesional permite.

Conclusión

Después de analizar el audio, compararlo con el de la Hojilla y escuchar todo el archivo viendo el espectrograma, corroboro de forma profesional y con toda seguridad que no existe edición alguna y que es imposible generar un archivo de una hora a base de ediciones de audios de la Hojilla, que presente tal uniformidad y coherencia semántica.

Ricardo Pujol
Ingeniero en Sonido DuocUC, Santiago, CL.
ricardopujolh@gmail.com