



VARIABLES QUE INTERVIENEN EN LA CAPTURA DE AUDIO CON DISPOSITIVOS MÓVILES

Variables involved in capturing audio with a mobile

Juan Sebastián Murillo Sarmiento

*Corporación Ciencias Pedagógicas Latinoamericana y del Caribe
Colombia*

juan.murillo@conservatoriodeltolima.edu.co

https://orcid.org/0000-0002-9333-9219

Jorge Enrique Núñez Díaz

*Conservatorio del Tolima
Colombia*

jorge.enrique@conservatoriodeltolima.edu.co

https://orcid.org/0000-0002-6491-8545

Jorge Andrés Sánchez

*Conservatorio del Tolima
Colombia*

jorge.sanchez@conservatoriodeltolima.edu.co

https://orcid.org/0000-0003-0153-1455

Recibido: 16 de julio 2024

Aceptado: 28 de noviembre 2024

DOI <https://doi.org/10.48204/j.centros.v14n1.a6598>

Resumen

En la búsqueda de un protocolo eficiente para que músicos capturen con la máxima calidad posible el sonido de sus instrumentos mediante dispositivos móviles, se identificaron variables críticas. La disertación teórica y la colaboración con coautores revelaron que la distancia entre el músico y el micrófono del dispositivo, así como el momento del día de la captura, son las variables más influyentes. La distancia afecta la intensidad del sonido capturado, mientras que el momento del día contribuye a variaciones en las frecuencias, especialmente en el ruido de fondo.





Se analizó el espectro de frecuencias de un violín ejecutando una escala cromática de SOL, grabada en formato ".WAV" con la herramienta "PARROT" en dispositivos móviles. El análisis se realizó a dos distancias y en dos horarios diferentes para determinar las condiciones ideales de captura. Los resultados indican que la captura de mayor calidad ocurre a las 2 a.m., a una distancia de 50 cm entre el músico y el micrófono del dispositivo. Se recomienda enviar el archivo sin comprimir a través de Google Drive al encargado de postproducción para evitar pérdida de calidad. Este hallazgo simplifica la postproducción musical, proponiendo un protocolo aplicable a músicos en general.

Palabras clave: Espectro de frecuencias; octavas; decibeles; escala cromática.

Abstract

In the quest for an effective protocol enabling musicians to capture the highest audio quality from their instruments using mobile devices, critical variables were identified. Theoretical deliberation and collaboration with co-authors revealed that the distance between the musician and the device's microphone, along with the time of day during capture, are the most influential variables. Distance impacts the captured sound's intensity, while the time of day contributes to variations in frequencies, especially in background noise. The frequency spectrum analysis focused on a violin performing a chromatic scale of G, recorded in ".WAV" format using the "PARROT" tool on mobile devices. Analysis was conducted at two distances and two different times to pinpoint optimal capture conditions. Results indicate that the highest-quality audio capture occurs at 2 a.m., with a distance of 50 cm between the musician and the device's microphone. Recommending uncompressed file transfer via Google Drive to the post-production team prevents quality loss. This discovery streamlines musical post-production, proposing a protocol applicable to musicians at large.





Keyword. Frequency spectrum; octaves; decibels; chromatic scale.

Introducción

Debido a la pandemia del COVID-19 en 2020, los músicos del Conservatorio del Tolima tuvieron que realizar grabaciones desde casa utilizando recursos mínimos como teléfonos celulares y conexión wifi. Esta situación impulsó la necesidad de diseñar un protocolo para capturar audio de alta calidad, esencial para las grabaciones caseras enviadas a través de Google Drive para la postproducción orquestal.

Entre las obras grabadas se encuentran 'El Bunde Tolimense' (Murillo y Sánchez, 2020a) y 'La Carcoma' (Murillo y Sánchez, 2020b), disponibles en YouTube – Videos del Conservatorio del Tolima.

La falta de un protocolo unificado para la captura de audio desde casa ha generado dificultades en la postproducción, debido a variaciones en la distancia, horarios y dispositivos utilizados por los músicos. Esto complica la sincronización de archivos en términos de extensión, tipo, tempo y espectro de frecuencias, aumentando significativamente las horas de trabajo en la mezcla final.

Este estudio es crucial para comprender las variables clave en la captura de audio con dispositivos móviles en contextos domésticos, aplicable no solo a músicos del Conservatorio del Tolima, sino también a otros artistas contemporáneos como youtubers, tiktokers e instagramers, cuya influencia se ha ampliado con la pandemia del COVID-19.





Según García, A. (2024), las aplicaciones móviles para la captura de audio han ganado popularidad globalmente, especialmente durante la pandemia del COVID-19 (SARS-CoV-2), cuando los artistas necesitaron grabar música desde casa de manera eficiente para la postproducción.

Eargle, J. M. (2004), indica que la presencia de ruidos de fondo en las grabaciones requiere el filtrado de frecuencias no deseadas, mientras que la elección del software y el método de envío del archivo pueden impactar la calidad del audio debido a la compresión.

Se realizó una investigación teórica detallada para diseñar un protocolo que mejore los procesos de captura y reduzca los tiempos de edición. Se consultaron recursos audiovisuales específicos como Astegiano, N. (2018) y Jiménez, R. (2018), que ofrecen técnicas prácticas para mejorar la calidad de las grabaciones y fundamentan el protocolo propuesto.

Aunque parecen similares y están relacionados físicamente, la intensidad del sonido se refiere a la energía que transporta una onda sonora a una distancia específica de su origen, medida en Watt por metro cuadrado (W/m^2).

En contraste, el nivel de intensidad se asocia con la sonoridad percibida del sonido, medida en decibeles (dB). Esta medida es logarítmica respecto a la intensidad del sonido, lo que significa que el "volumen" percibido no varía linealmente con la distancia a la fuente (Serway, 1992).





Ecuación 1. Intensidad

$$I = \frac{P}{4\pi r^2}$$

En donde:

I: Intensidad del sonido en (W/m²)

P: Potencia de la fuente de sonido en (W)

r: Distancia a la cual se encuentra el dispositivo de captura de audio desde la fuente en (m)

En cuanto la ecuación 2, puede escribirse que la suma de dos niveles de intensidad diferentes NO puede realizarse de forma algebraica, por la misma razón que estos valores no tienen un comportamiento lineal sino logarítmico. Por otro lado, estas ecuaciones predicen que el nivel de intensidad de un sonido es dependiente de la distancia de la fuente y que, si se desea capturar ese sonido, la distancia a la cual esté ubicado el instrumento va a afectar la sonoridad de la captura.

Ecuación 2. Presión sonora

$$dB = 10 \text{Log} \frac{I}{10^{-12}}$$

Donde:

dB: Nivel de intensidad (sonoridad) del sonido en decibeles (dB)

Según Fletcher (1933), la percepción humana del sonido no solo depende de su intensidad física, sino también de factores psicoacústicos que afectan nuestra





percepción del "volumen" del sonido. Esto subraya la importancia de considerar aspectos psicoacústicos en la experiencia auditiva.

Además, la sonoridad, medida en decibeles, se presenta de manera logarítmica en relación con la intensidad del sonido. Es un atributo subjetivo influenciado por la edad, la familiaridad y la frecuencia del sonido, lo que puede resultar en percepciones diferentes entre personas expuestas al mismo sonido a la misma distancia (Hewitt, 2007).

Por otro lado, el fenómeno del enmascaramiento del sonido desempeña un papel fundamental en la percepción auditiva. Aunque hemos comprendido hasta ahora que la sonoridad depende de manera logarítmica e inversa a la distancia entre el receptor y la fuente sonora, diversos estudios respaldan la idea de que la ley del inverso de la distancia también afecta a la amplitud de la onda y, en consecuencia, a su sonoridad (Uriz, A. J., et al., 2013; Avallone, 2002).

La ecuación 3, como se muestra, ilustra que la amplitud de una onda se atenúa a medida que la distancia 'r' desde la fuente aumenta. Este fenómeno de atenuación se revela como un factor clave para comprender la compleja interacción entre la distancia y la percepción del sonido. Investigaciones adicionales, como las de Johnson (2012), respaldan esta perspectiva y aportan a nuestra comprensión sobre cómo las variaciones en la amplitud influyen directamente en la experiencia auditiva. Esta comprensión profunda del enmascaramiento del sonido es esencial para diseñar protocolos efectivos en la captura y reproducción de audio, especialmente en entornos donde la distancia entre la fuente y el receptor puede variar significativamente.





Ecuación 3. Amplitud

$$A = \frac{A_0}{r} \sin(Kr - \omega t)$$

Donde:

A: Amplitud de la onda en (m)

A₀: Amplitud inicial de la onda en (m)

r: Distancia a la cual se encuentra el dispositivo de captura de audio desde la fuente en (m)

K: Es una constante que depende de la longitud de onda en (rad/m)

ω: Frecuencia angular de la onda en (Rad/s)

t: Tiempo de oscilación de la onda en el espacio desde un t de referencia en (seg)

Berg, y Stork (2005) señalan que debido a que la amplitud de la onda es inversa a la distancia, es de esperarse que la sonoridad de un sonido cercano al receptor sea más alta que la de un sonido lejano. Esto se debe a la distribución de la energía en un área mayor y a la absorción de energía por cuerpos con masa que se antepone a los frentes de onda, por difracción. De esta manera, el receptor (micrófono) puede dar prioridad a la captura de sonidos cercanos, lo que resalta la importancia de obtener capturas de audio cercanas al receptor de sonido.

Con respecto a las frecuencias emitidas por los instrumentos musicales: Tono y Timbre de un instrumento, que es único y está determinado por su configuración geométrica y material, se distingue por las ondas parciales que se superponen a la onda principal de frecuencia. Esto significa que, aunque dos instrumentos puedan emitir la misma nota, cada uno tiene un timbre característico debido a estas ondas parciales (Hewitt, 2007; Marchiano y Martínez, 2017).





La limitación del rango de frecuencias emitidas por cada instrumento enfatiza su singularidad y contribuye a la diversidad sonora en una orquesta. Sin embargo, la presencia de ondas no armónicas, o ruido, puede afectar negativamente la calidad del sonido emitido. Estas ondas pueden interferir con la onda armónica principal, complicando su eliminación durante la postproducción (Hewitt, 2007).

Investigaciones recientes destacan cómo la configuración y el material de los instrumentos influyen en su capacidad para generar distintos tonos y timbres, subrayando la complejidad y la riqueza de la variedad sonora en la música (Scavone y Smith, 2021).

La eliminación del ruido en la captura de audio, mediante software de filtrado basado en análisis de Fourier, se vuelve más compleja a medida que aumenta la cantidad de fuentes de ruido y ondas parciales presentes (Avallone, 2002; Warren, 2013).

El violín, conocido por su versatilidad en el rango de frecuencias (200 Hz - 20 KHz) y su facilidad para abordar una escala cromática en clave de sol, es central en esta investigación (Serway, 1992).

Investigaciones destacadas en psicoacústica han profundizado en cómo el oído humano percibe las frecuencias musicales y la estructura tonal de los instrumentos. Por ejemplo, Wulf Díaz (2019) aporta una perspectiva cognitiva valiosa.

Además, expertos como Aguilar (2006) han explorado cómo los instrumentos moldean las frecuencias a lo largo del tiempo, enriqueciendo nuestra comprensión de las interacciones entre instrumentos y frecuencias.





Estas contribuciones fortalecen la base teórica de nuestra investigación en acústica y psicoacústica, ampliando nuestro entendimiento de las relaciones entre instrumentos musicales y las frecuencias que emiten.

Materiales y Métodos

Para comparar la captura de audio con y sin control de ruido de fondo, se realizaron las siguientes actividades:

- **Diagnóstico de Equipos:** Se encuestó a los músicos de la orquesta sinfónica del Conservatorio del Tolima para identificar el dispositivo móvil y el instrumento utilizado.
- **Análisis de Coincidencias:** Se verificó la ficha técnica de los dispositivos móviles informados para seleccionar músicos con micrófonos similares y que tocaran el violín.
- **Selección de Escala Cromática:** Se eligió una escala cromática para capturar todas las frecuencias posibles del violín, evaluadas posteriormente con un software de espectro de frecuencias.
- **Herramienta de Captura:** Se utilizó la aplicación gratuita PARROT para capturar el audio en formato .WAV.
- **Envío de Archivos:** Los archivos de audio se enviaron a través de Google Drive para evitar la compresión y pérdida de calidad.
- **Ampliación de Muestra:** Se convocó a violinistas de otras instituciones con dispositivos móviles similares.





- Protocolo de Captura: Se diseñó y capacitó a los músicos en un protocolo específico para capturar audio con condiciones controladas de ruido de fondo en un contexto doméstico.

Se diagnosticaron los dispositivos móviles de los músicos del Conservatorio del Tolima y otras universidades mediante una encuesta que evaluó la capacidad de almacenamiento y la calidad del micrófono.

De los encuestados, se seleccionaron 10 violinistas con dispositivos Samsung Galaxy A 2017. Estos músicos fueron capacitados en el protocolo para la captura de muestras de audio.

El análisis de los datos se realizó con Blue Cat's FREQUANALYST versión 2.31, seleccionada por su disponibilidad. Se hicieron 30 capturas de audio, 10 por músico, con las siguientes variables:

- Hora 2 p.m. distancia 50 cm.
- Hora 2 a.m. distancia 50 cm.
- Hora 2 a.m. distancia 250 cm.

El objetivo era comparar las capturas a las 2 PM y 2 AM manteniendo la distancia constante (50 cm) para analizar las frecuencias generadas por los ruidos de fondo. Además, se compararon capturas a las 2 AM a diferentes distancias (50 cm y 250 cm) para evaluar el efecto de la distancia entre el instrumento y el micrófono.





El análisis de las 30 capturas con FREQANALYST generó las figuras 4 a 7. Los resultados mostraron que las formas de onda capturadas a las diferentes horas y distancias coincidieron en su morfología para cada caso.

Resultados

La Figura 1 muestra que FREQANALYST puede analizar frecuencias de 10 Hz a 20 KHz. Sin embargo, el violín abarca frecuencias desde 47 Hz hasta 22 KHz, con las graves entre 47 Hz y 470 Hz.

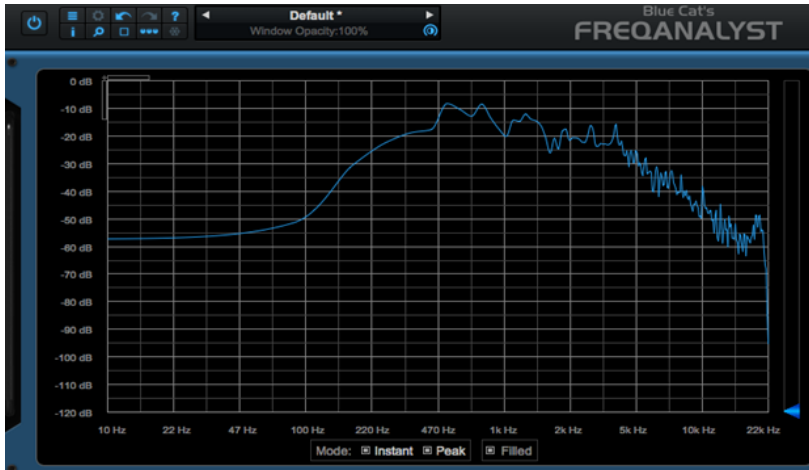
Después de 470 Hz, se observan picos armónicos adyacentes hasta 1000 Hz, alcanzando el valor máximo en decibeles permitido por el software. También se nota un extremo interesante cerca de los 22 KHz. La gráfica muestra una pendiente positiva entre 47 Hz y 470 Hz, y una pendiente negativa entre 470 Hz y 22 KHz, indicando que los armónicos pierden energía hacia el final del espectro.

Basado en la teoría, las frecuencias del violín deberían comenzar en 200 Hz. Las frecuencias capturadas entre 10 Hz y 196 Hz se consideran sonidos residuales, probablemente debido a la reverberación de la habitación y el ruido eléctrico del preamplificador o la cápsula del micrófono del celular. La Figura 1 destaca un pico máximo entre 470 Hz y 1000 Hz, indicando que la proximidad del violín a 50 cm no genera saturación en el micrófono, ya que no se presenta una lectura de 0 dB.



Figura 1.

Nivel de intensidad de un sonido en escala cromática de SOL generada por un violín capturado a 50 cm de la fuente a las 2 p.m.



En la Figura 2 se observa que el comportamiento de la respuesta armónica de un sonido capturado a las 2 a.m. con la misma escala cromática es similar al caso de la Figura 1, sin embargo, la diferencia entre las crestas y valles es menor comparada con el sonido capturado a las 2 p.m., en especial, puede notarse que la tiene un rango (eje vertical) desde -95 dB (ver la frecuencia de 22 KHz de la onda a la derecha de la gráfica) hasta un poco mayor a -10 dB

Figura 2.

Nivel de intensidad de un sonido en escala cromática de SOL generada por un violín capturado a 50 cm de la fuente a las 2 a.m.



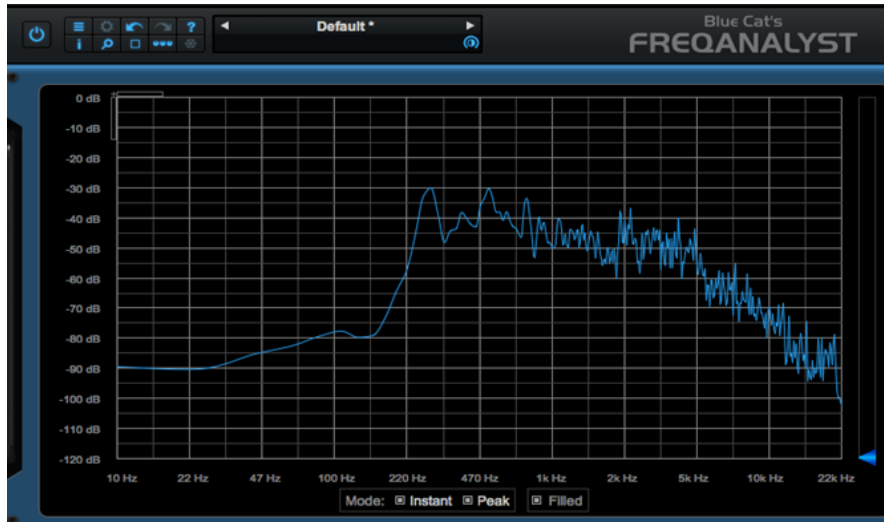
En el análisis del audio capturado a 250 cm del micrófono, no se observa claramente la nota más grave del violín (aproximadamente 196 Hz). La primera cresta de onda aparece entre 220 y 470 Hz, con una disminución en la amplitud de los picos, que en capturas a 50 cm se acercaban a -10 dB (Figuras 1 y 2).

La pendiente entre 470 Hz y 22 KHz es más pronunciada en altas frecuencias, a partir de 5 KHz, indicando una rápida pérdida de energía en este rango. Además, los niveles de intensidad son menores, con un intervalo de -100 dB a -30 dB, comparado con los casos anteriores.

La escala logarítmica de decibeles muestra que el menor rango en las capturas a 250 cm respecto a las de 50 cm evidencia grandes pérdidas de energía en la onda a mayores distancias entre el instrumento y el micrófono. (Figura 3).

Figura 3.

Nivel de intensidad de un sonido en escala cromática de SOL generada por un violín capturado a 250 cm de la fuente a las 2 a.m.



La Figura 4 muestra una gráfica comparativa del espectro de frecuencias capturado por el micrófono del dispositivo móvil al ejecutar la escala cromática de SOL con un violín interpretado por un músico con experiencia, en su orden de arriba hacia abajo así: 50 cm 2 pm, 2 am y 250 cm 2:00 am.

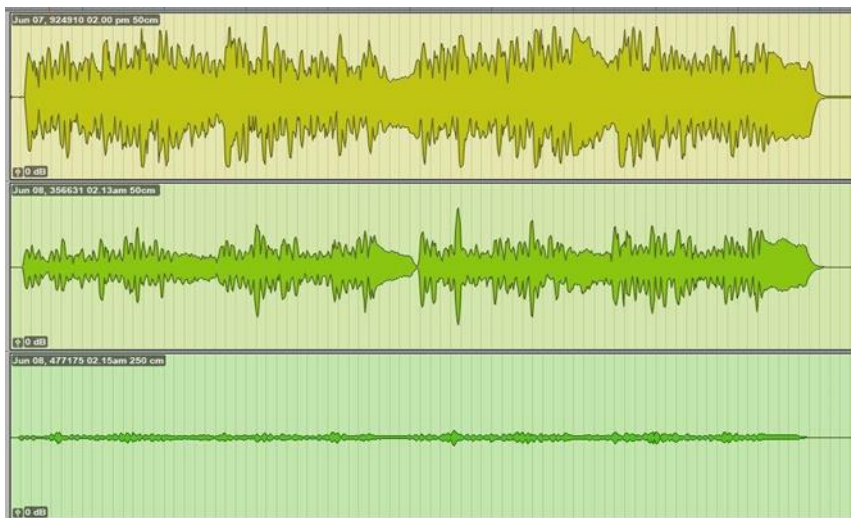
Se observa claramente que la cantidad de crestas que se presentan al capturar el audio a las 2 pm es mayor que la cantidad de crestas que posee un sonido de la misma naturaleza capturado a las 2 am. Esto pone de manifiesto que existe una gran diferencia entre los ruidos de fondo capturados durante las horas del día comparados con los sonidos capturados en las horas de la madrugada a una distancia definida, lo cual, dificulta, en el primer caso, el proceso de filtración de ruidos de fondo en el proceso de post producción.

Para el caso del sonido capturado a 250 cm a las 2 am, que corresponde al tercero en orden de arriba hacia abajo, comparado con el nivel de sonoridad del

sonido capturado a 50 cm a la misma hora de la madrugada, puede verse que el nivel de intensidad es menor, disminuyendo la amplitud de las crestas, perdiendo energía durante el desplazamiento de la onda lo cual dificulta el proceso de post producción ya que puede requerir incrementos de ganancia para lograr niveles de intensidad similares a sonidos capturados a distancias menores. Cabe resaltar que las frecuencias específicas son más difíciles de detectar y tratar con el software musical para la onda capturada a 250 cm que en los otros casos de captura a 50 cm.

Figura 4.

Análisis del espectro de frecuencias del sonido capturado a 50 cm 2 pm, 2 am y 250 cm 2:00 am de arriba hacia abajo



Conclusiones

El audio capturado a las 2 am tiene menos frecuencias de ruido de fondo que a las 2 pm, lo que mejora su calidad. Esto reduce significativamente el tiempo necesario para eliminar ruidos no deseados en la post producción.



Una distancia mayor al micrófono (250 cm vs. 50 cm) reduce la intensidad del sonido y el espectro de frecuencias capturadas. Esto disminuye la sonoridad del instrumento y aumenta la presencia de frecuencias de fondo no deseadas.

Introducir un micrófono en un lugar como un ropero con objetos absorbentes (almohadas, cobijas, prendas de vestir) mejora la calidad del audio al reducir la reverberación.

Se recomienda capturar audio con una aplicación que preserve todas las frecuencias del espectro musical en formato ".WAV". Además, hay que evitar cambiar la extensión del archivo al enviarlo, ya que podría comprimir el audio y limitar el dominio de frecuencias durante la post producción.

Por otro lado, utilizar Google Drive para enviar archivos asegura la conservación de la calidad original del audio.

Agradecimientos

A la Universidad del Conservatorio del Tolima, por su compromiso e importancia dada a los procesos de investigación en el clúster de la música y el arte. A los músicos que realizaron las tomas de las muestras y a todo el equipo de trabajo del semillero de Investigación AVANTE del Conservatorio del Tolima (Estudiantes y profesores)

Referencias Bibliográficas

Aguilar, A. (2006). Reflexiones sobre el análisis de música electroacústica. In *ANPPOM--XVI Congresso da Associação Nacional de Pesquisa e Pós-graduação em Música.*





https://anppom.org.br/anais/anaiscongresso_anppom_2006/CDROM/COM/07_Com_TeoComp/sessao02/07COM_TeoComp_0201-187.pdf

Astegiano, N., (Dir.). (2018). Video: Cómo grabar una voz. <https://youtu.be/VEw3gNvPXdo>

Avallone, E., y Baumeister III, T., (Dir.). (2002). *Marks Manual del Ingeniero Mecánico*. MéxicoD.F, México: Novena Edición, Mc Graw Hill.

Berg, R. E., y Stork, D. G. (2005). *The physics of sound* (3rd ed.). Pearson/Prentice Hall. ISBN-13: 9780131457898

Eargle, J. M. (2004). *The Microphone Book* (2nd ed.). Focal Press. URL: https://soma.sbccc.edu/users/davega/FILMPRO_181_AUDIO_I/FILMPRO_181_04_Reference_Notes/FILMPRO_181_Microphones/The%20Microphone%20Handbook.pdf

Fletcher, H., y Munson, W. A. (1933). Loudness, its definition, measurement and calculation. *Bell System Technical Journal*, 12(4), 377-430. Doi: <https://doi.org/10.1002/j.1538-7305.1933.tb00403.x>

García, A. (2024). *Las aplicaciones móviles para la captura de audio han ganado popularidad globalmente, especialmente durante la pandemia del COVID-19 (SARS-CoV-2)*. <https://hdl.handle.net/11362/45938>

Hewitt, P., (Dir.). (2007). *Física conceptual*. México D.F, México: Décima Edición, PearsonAddison Wesley.

Marchiano, M., y Martínez, I. C. (2017). Dimensiones acústicas y percepción del timbre en el procesamiento de grabaciones del análisis musical. In *XIII Simposio Internacional de Cognição e Artes Musicais (XIII SIMCAM) (Paraná, Brasil, 2017)*. Url: <http://sedici.unlp.edu.ar/handle/10915/63871>

Murillo, J. S., Sánchez, J. A., (Dir.). (2020). El bunde Tolimense. Conservatorio del Tolima. <https://youtu.be/FsDgFj8Egfg>

Murillo, J. S., Sánchez, J. A., (Dir.). (2020). La Carcoma. Conservatorio del Tolima. <https://youtu.be/Aq9ksIsUGGI>





- Scavone, G., y Smith, J. O. (2021). A landmark article on nonlinear time-domain modeling in musical acoustics. *The Journal of the Acoustical Society of America*, 150(2), R3-R4. <http://hdl.handle.net/10550/21230>
- Serway, A. R., (Dir.). (1992). *Física*. México D.F, México: Tercera Edición, Mc Graw Hill. Verdier, J. P., (Dir.). (2016). Video: 10 Tips para registro de sonido. <https://youtu.be/aeSEmZ03u1Y>
- Uriz, A. J., Agüero, P. D., Tulli, J. C., Moreira, J. C., González, E. L., y Hidalgo, R. M. Dispositivo de Bajo Costo para el Tratamiento de Acúfenos. *LIBRO DE ACTAS*, 82. https://www.aitadis.org/wp/wp-content/uploads/2015/01/Libro-de-Actas-IBERDISCAP-2013_Parte1.pdf#page=82
- Warren, R. M. (2013). *Auditory perception: A new synthesis* (Vol. 109). Elsevier. <https://books.google.es/books?id=1zyOAgAAQBAJ&lpg=PP1&ots=Rle-ZEiiWg&dq=Warren%20waves%20sound%20&lr&hl=es&pg=PP1#v=onepage&q=Warren%20waves%20sound&f=false>
- Wulf Díaz, F. (2019). Descripción de la musicalidad humana desde un enfoque interdisciplinario y sus relaciones con el desarrollo lingüístico y comunicativo temprano. <https://repositorio.uchile.cl/handle/2250/171025>

