

Inteligencia artificial y machine learning

Educación musical

Producción musical y creación artística

Acústica

cordia
ouos

artes sonoras
producción
musical

Revista Volumen 3 / N.º 9
MAYO 2024
ISSN 2953-657X

Revista Volumen 3 / N.º9

Mayo 2024

ISSN 2953-657X



AES LAC 2023

Conferencia Latinoamericana
de la Audio Engineering Society

Artes
EDICIONES



Rector: William Herrera

Vicerrector Académico: Bradley Hilgert

Vicerrectora de Posgrado e Investigación: Olga del Pilar López

Director Escuela de Artes Sonoras (EAS): Javier Andrés Ocampo Cardona

Director de Sonocordia: Luis Pérez-Valero

Editor invitado: Diego Benalcázar

D. R. © Universidad de las Artes



Este obra está bajo una licencia de Creative Commons
Reconocimiento-NoComercial-CompartirIgual 4.0 Internacional.



Dirección: José Miguel Cabrera Kozisek

Diseño y maquetación: José Ignacio Quintana

Corrección de textos: Silvia Daniela Zeballos Manosalvas

MZ14, Av. 9 de Octubre y Panamá

Guayaquil, Ecuador

editorial@uartes.edu.ec



Vol. 3 N.º 9 / Guayaquil
mayo 2024
ISSN 2953-657x

SONOCORDIA. Revista de artes sonoras y producción musical

Universidad de las Artes

Volumen 3, Número 9. Mayo 2024

ISSN 2953-657X

Director: Luis Pérez-Valero

Universidad de las Artes (Ecuador)

Coordinación editorial:

Omar Domínguez Castro y Pedro Segovia Universidad de las Artes (Ecuador)

Editora asociada: Bernarda Ubidia

Universidad de las Artes (Ecuador)

COMITÉ CIENTÍFICO EDITORIAL

Johanna Abril

Universidad de las Américas (Ecuador)

Jannet Alvarado

Universidad de Cuenca (Ecuador)

Claudia Fallalero

Gabinete de Patrimonio Musical Esteban Salas (Cuba)

Pablo Freiberg

Universidad de las Artes (Argentina)

Adina Izarra

Universidad de las Artes (Ecuador)

David de los Reyes

Universidad de las Artes (Ecuador)

Ketty Wong

Universidad de Kansas (Estados Unidos)

CONSEJO ASESOR

Miguel Álvarez-Fernández

Ars Sonora - Radio Nacional de España (España)

Andrey Astaiza

Universidad de las Artes (Ecuador)

Susan Campos-Fonseca

Universidad de Costa Rica (Costa Rica)

Miriam Escudero Suásteegui

Gabinete de Patrimonio Musical Esteban Salas (Cuba)

Ana María Ochoa

Universidad de Columbia (Estados Unidos)

Carmen Pardo

Universidad de Girona (España)

Rodrigo Sigal

Centro Mexicano para la Música y las Artes Sonoras (México)

Sonocordia. Revista de artes sonoras y producción musical

Universidad de las Artes

Malecón Simón Bolívar y Francisco Aguirre

Guayaquil, Ecuador. CP 090313

revista.sonocordia@uartes.edu.ec

www uartes.edu.ec/sitio/sonocordia

Índice

Introducción

Jorge Azama 7

Prólogo

AES LAC 2023

Diego Benalcázar 9

Artículos

Optimizando la mezcla: Answer Set Programming

en la producción musical inteligente

Flavio Everardo, Ileana Benítez, Yamil Burguete 11

Estudio comparativo de enfoques probabilísticos

para la generación de progresiones armónicas

Juan Núñez, Cristian Bajaña, Mateo Yepes, Daniel Gómez 27

Implementación de una aplicación software

para la emulación de una consola mezcladora de audio

como herramienta de aprendizaje significativo

Michael Acevedo, Josué Ordóñez, Juan Prieto,

Daniel Layton, Eric Rodríguez, Daniel Jara-Hurtado 47

Software adaptivo para el entrenamiento auditivo técnico

en las aulas de clases

Juan Corredor 59

La sonificación de datos como herramienta tecnológica

para la creación artística

Diego Molina, Isabella Álvarez, Lina González, Jamir Moreno 69

Pódcast inmersivo para la divulgación de historias de niños,

niñas y adolescentes víctimas del conflicto armado colombiano

María Castañeda, Julia Cubides, Laura Pineda 87

Estudio de viabilidad del residuo del cultivo de cáñamo

como material de acondicionamiento acústico

Timothy Carrera 97

Miscelánea (Fuera de tono)

Epílogo

Renato Zamora 113

Memorias AES LAC 2023

Juan Jiménez 115

Biografías de los autores 121

Guía rápida para los autores 125

Introducción

Jorge Azama
Vicepresidente regional para América Latina
Audio Engineering Society (AES)

Entre los días 18 y 21 de septiembre de 2023, se llevó a cabo en las bellas instalaciones de la sede Granados de la Universidad de las Américas (UDLA), en la ciudad de Quito, Ecuador, la Conferencia Latinoamericana de la Audio Engineering Society (AES LAC) 2023, organizada por el equipo de AES Ecuador, encabezado por mis queridos colegas y amigos Renato Zamora y Juan Jiménez, quienes fueron los *co-chairs* de la conferencia.

Este evento —el más importante del año en el calendario de actividades de AES en América Latina— fue una gran oportunidad para volver a reunir a profesionales, investigadores y estudiantes de todo el mundo, pero especialmente de nuestra región, en jornadas intensas pero muy enriquecedoras en las que se compartieron y difundieron muchos conocimientos en diferentes aspectos del vasto mundo del audio profesional.

Además, en esta oportunidad, la AES LAC se vio muy honrada de contar con la presencia de autoridades de la Audio Engineering Society, como Bruce Olson, presidente de la AES a nivel mundial, quien estuvo encargado de ofrecer el discurso de apertura, haciendo un repaso de los 75 años de historia de la institución, y Agnieszka Roginska, expresidenta de la AES y actual integrante del Comité de Coordinación de Eventos de AES (ECC, por sus siglas en inglés). Ellos pudieron comprobar *in situ* el gran trabajo de organización realizado por nuestros queridos hermanos de AES Ecuador, anfitriones del evento.

Durante los días que duró la LAC 2023, pudimos asistir a ponencias y exposiciones de renombrados profesionales latinoamericanos (de países como Colombia, Chile, Argentina, Perú, México, Bolivia, Brasil, Venezuela y Ecuador) y de otras latitudes del planeta (Estados Unidos, España, Canadá o Alemania). Entre las tantas y variadas ponencias podemos mencionar, por ejemplo, la del reconocido ingeniero estadounidense Bob McCarthy (Meyer Sound), quien nos habló sobre el «Diseño e implementación del sistema de sonido utilizado por la banda Metallica durante su Gira Mundial M72»; el ecuatoriano Joffre Amaguaña, quien cubrió el tema «Diseño sonoro para musicales y ópera»; el español Miguel Domínguez con su charla sobre «Implementación de sistemas de audio inmersivo»; la canadiense Hazel Burns con «Técnicas de grabación clásicas para estudio»; la colombiana Marcela Zorro que nos ilustró sobre «El inventario sonoro de los órganos tubulares en Colombia»; el brasileño Carlos Freitas con «¿Por qué es importante masterizar tu música en Dolby Atmos?»; el boliviano Marcelo Navía quien cubrió el tema «Emulación digital de procesos analógicos»; y un largo etcétera.

Otro aspecto para resaltar fue que, en esta conferencia, tuvimos la suerte de contar con una «sala inmersiva», especialmente acondicionada por nuestros queridos amigos de Genelec, importante empresa finlandesa que se ha convertido es una de las principales auspiciadoras de la LAC desde hace varios años. Allí tuvimos un sistema Dolby Atmos 7.1.4 con monitores activos de dicha compañía, calibrado por sus ingenieros, que permitió

que los asistentes pudieran conocer más y escuchar ejemplos en el lugar y en tiempo real gracias a las presentaciones de reconocidos profesionales con amplia experiencia de trabajo en este formato de audio inmersivo como Miguel Domínguez (España), Ivo Markovic (Argentina), Christopher Manhey (Chile), Francisco Bissone (Argentina), Daniel Pasquel (Ecuador), Eric Horstmann (Alemania), Carlos Freitas (Brasil), Ken Lewis (Estados Unidos) y Fernando Richard (Argentina).

Y, como ya es costumbre en todas las ediciones de la AES LAC, también pudimos presenciar las etapas finales de la Competencia de Grabación Estudiantil. En estas, estudiantes de distintas instituciones educativas de países diversos de Latinoamérica presentaron sus trabajos ante exigentes jurados conformados por profesionales presentes en el evento, quienes compartieron sus conocimientos, retroalimentación y críticas constructivas para que sigan mejorando su trabajo. Siempre es muy grato ver que chicas y chicos de distintos países, más allá de la competición misma, terminan conociendo nuevos amigos y ampliando su red de contactos a nivel internacional, un aspecto muy importante que la Audio Engineering Society también ayuda a promover entre sus miembros. En este punto, también aprovechamos para agradecer a todas la compañías y empresas que patrocinaron esta competencia; su gestión permitió que los estudiantes ganadores pudieran llevarse a casa premios en cada categoría.

Finalmente, cabe agradecer y reconocer de manera muy especial a aquellos investigadores latinoamericanos que acudieron al llamado para la presentación de *posters* y *papers*, trabajos que fueron expuestos y compartidos con la audiencia durante la conferencia. En esta publicación se podrán revisar aquellos trabajos gracias a la paciente y exigente labor del profesor Diego Benalcázar, quien fue el *chair* de esta con-

vocatoria y cuyo arduo trabajo de más de un año se ve plasmado en estas páginas. De esa forma, se produjo un documento que quedará a disposición de generaciones futuras como medio de consulta y aprendizaje.

Solo me queda felicitar nuevamente a todas y todos mis colegas de AES Ecuador por haber llevado a cabo esta impecable conferencia y sin cuyo gran y silencioso trabajo y esfuerzo, siempre tras bambalinas, no se hubiera logrado que la AES LAC 2023 fuera un rotundo éxito. Quisiera agradecer también a todas las empresas ecuatorianas e internacionales que colaboraron con la organización de la LAC y a todos los profesionales —que en muchos casos se movilizaron desde lugares muy distantes— que nos brindaron su tiempo y conocimientos para compartirlos con los asistentes a la LAC. Debo agradecer también al público asistente, profesionales, estudiantes y amantes del audio profesionales que colmaron las instalaciones de la UDLA durante estas jornadas.

Solo me queda invitarlos a las AES LAC 2024, que se llevará a cabo en el mes de septiembre de este año en la ciudad de La Paz, Bolivia. Para ese evento ya venimos trabajando muy duro con los colegas hermanos de AES Bolivia, quienes han asumido la enorme responsabilidad de ser los anfitriones de esta nueva edición, que estoy seguro de que será tan exitosa como las anteriores. Estén atentos a las redes sociales de AES Bolivia, porque en breve estaremos anunciando toda la información respecto a la LAC 2024, que, desde este año, además pasará de ser una conferencia a una convención: una de las tres convenciones internacionales en el calendario anual de la Audio Engineering Society junto con la de Europa (en junio en Madrid, España) y la de Estados Unidos (que se llevará a cabo en octubre en Nueva York).

Reciban todas y todos un afectuoso saludos, jnos vemos en La Paz!

Prólogo

AES LAC 2023

Diego Benalcázar Vega, Ph. D.

University of the Arts London

d.benalcazar@lcc.arts.ac.uk

En esta edición especial de *Sonocordia*, me complace presentar una compilación excepcional de investigaciones presentadas en la reciente conferencia latinoamericana de la Audio Engineering Society celebrada en Quito, Ecuador, en septiembre de 2023. Este encuentro, que congregó a mentes brillantes en el ámbito del sonido, la música y la producción musical, ha sido un espacio de divulgación e innovación, donde la tecnología se encuentra con el arte para explorar y expandir las fronteras del audio y el sonido.

Los trabajos seleccionados abarcan una diversidad temática que refleja no solo la profundidad y amplitud del campo musical y sonoro, sino también la particular resonancia que estas investigaciones tienen en el contexto latinoamericano. Desde nuevos métodos de producción musical inteligente hasta la exploración de la sonificación de datos como una nueva forma de expresión artística, cada ponencia ofrece una ventana única a las últimas tendencias y desarrollos en la producción musical, educación, acústica y sonido.

Dentro de la educación de audio, se destaca el desarrollo y la implementación de software para la emulación de consolas mezcladoras de audio y para el entrenamiento auditivo técnico. Estos son ejemplos claros de cómo la intersección entre tecnología y pedagogía puede fomentar un aprendizaje significativo en nuestras aulas.

De la misma manera, tenemos textos sobre la creación artística a través del audio, como la utilización del audio inmersivo en un pódcast para compartir las historias de jóvenes afectados por el conflicto armado en Colombia, lo que ejemplifica el poder del audio para construir empatía y comprensión.

Las investigaciones también se extienden al ámbito de la acústica y la ingeniería del sonido, con textos sobre el estudio de materiales sostenibles para la acústica, con un enfoque particular en el potencial del residuo del cultivo de cáñamo, que subraya la relevancia de la acústica y su aplicación práctica en diversas regiones de Latinoamérica.

Este conjunto de artículos no solo resalta la importancia de la investigación en audio, música y sonido en Latinoamérica, sino que también celebra nuestra capacidad de contribuir significativamente a la discusión global desde una perspectiva enriquecida por nuestra cultura y arte únicos. Es un testimonio del espíritu innovador y la creatividad sin límites que

caracterizan a nuestra comunidad científica y artística.

Esta compilación no solo sirve como un escaparate de la excelencia académica y creativa de nuestra región, sino también como un puente para el diálogo intercultural e interdisciplinario. Al leer estos trabajos, nos sumergimos en un viaje que trasciende lo meramente técnico y nos adentra en una reflexión sobre cómo la música y el sonido reflejan, influencian y transforman nuestras sociedades.

La diversidad de temas abordados en esta edición refleja la riqueza y complejidad de los trabajos investigativos dentro de Latinoamérica. La investigación en estos campos es fundamental tanto para el avance tecnológico y académico como para la preservación, interpretación y evolución de nuestro quehacer cultural y académico. Al realizar estos trabajos en nuestra región y explorar nuevas formas de expresión y comprensión, estas investigaciones contribuyen al diálogo global sobre la música y el sonido, destacando la singularidad y la universalidad de nuestras experiencias.

Además, la inclusión de temas como la sostenibilidad y el impacto social en la investigación musical y sonora subraya la responsabilidad que tenemos como comunidad académica y artística de abordar los desafíos contemporáneos. La música y el sonido, en

su capacidad de conectar a las personas y contar historias, poseen un poder único para generar conciencia, fomentar la empatía y promover el cambio social.

Al celebrar estas investigaciones, también reconocemos los retos y oportunidades que enfrenta nuestra región en el campo de la música y la producción musical. La tecnología, que ocupa un lugar central en muchos de estos trabajos, ofrece herramientas poderosas para superar barreras, democratizar el acceso a la educación musical y fomentar la innovación en la creación y producción artística. Sin embargo, también nos enfrentamos al desafío de asegurar que estos avances sean accesibles para todos y reflejen la diversidad de voces y experiencias dentro de nuestras comunidades.

Con esta edición especial, anhelamos inspirar a futuras generaciones de investigadores, ingenieros, músicos, productores y educadores a continuar explorando las infinitas posibilidades que el sonido y la música nos ofrecen. Esta compilación es la continuación de los procesos académicos de AES Latam y una inspiración para nuevas investigaciones, colaboraciones y exploraciones artísticas que continúen enriqueciendo el panorama musical y sonoro de Latinoamérica y del mundo.



Vol. 3 N.º 9 / Guayaquil
mayo 2024
ISSN 2953-657x

Optimizando la mezcla: Answer Set Programming en la producción musical

Flavio Everardo
Escuela de Humanidades y Educación,
Tecnológico de Monterrey Campus Puebla
Puebla, México
Instituto de Ciencias de la Computación,
Universidad de Potsdam
Potsdam, Alemania
flavio.everardo@tec.mx

Ileana Benítez
Escuela de Humanidades y Educación,
Tecnológico de Monterrey Campus Puebla
Puebla, México

Yamil Burguete
Escuela de Humanidades y Educación,
Tecnológico de Monterrey Campus Puebla
Puebla, México

Resumen

En la producción musical, el proceso de mezcla es una tarea compleja donde se deben tomar decisiones interconectadas, tanto técnicas como creativas, para lograr un resultado apropiado. La producción musical inteligente (*Intelligent Music Production*) es un campo emergente que combina métodos cognitivos con procesos de creación y producción musical, que van desde la composición hasta la masterización. Answer Set Programming (ASP) es un paradigma declarativo de programación del área de representación de conocimiento y razonamiento capaz de resolver problemas complejos de optimización, el cual ha demostrado prometedores avances en procesos de producción musical. Este artículo propone un sistema basado en ASP llamado *frmixerr* como un enfoque inteligente para optimizar el proceso de mezcla, con generación automática de mezclas considerando criterios de optimización derivados de un análisis estadístico y una extracción de características en la respuesta en frecuencia. Se realizó una comparación sobre la calidad subjetiva de mezclas generadas por profesionales y por *frmixerr*. No hay diferencias sustanciales en la evaluación subjetiva entre la mezcla generada por el algoritmo y las realizadas por profesionales con diferentes años de experiencia en mezcla.

Palabras clave: Answer Set Programming, *frmixerr*, inteligencia artificial, mezcla automática, producción musical inteligente,

Abstract

In music production, the mixing process is a very complex task where interconnected technical and creative decisions must be made to achieve an appropriate result. Intelligent Music Production is an emergent field that combines cognitive methods with music creation and production processes, ranging from composition to mastering. Answer Set Programming (ASP) is a declarative programming paradigm in the Knowledge Representation and Reasoning area capable of solving complex optimization problems, which has shown promising advances in music production processes. This article proposes an ASP-based system called *frmixerr* as an intelligent approach to optimize the mixing process, automatically generating mixes that consider optimization criteria derived from statistical analysis and frequency response feature extraction. Subjective quality was compared between mixes generated by professionals and *frmixerr*. There are no substantial differences in the subjective evaluation between the mix generated by the algorithm and the ones generated by professionals with different years of mixing experience.

Keywords: Answer Set Programming, *frmixerr*, artificial intelligence, automatic mixing, intelligent music production

1. Introducción

En el contexto de la producción musical, es bien sabido que la mezcla es una tarea muy desafiante y de optimización. Para que una mezcla suene bajo los estándares de la industria, debemos asumir que los pasos previos de producción, como la grabación y su edición correspondiente, se realizaron correctamente para que el ingeniero de mezcla pueda tomar decisiones que implican desafíos técnicos.¹ Algunos de estos retos involucran asegurar el balance entre las fuentes, su audibilidad mediante la minimización del enmascaramiento, la distribución de elementos en el campo panorámico, limitar el rango dinámico o incluso garantizar una amplitud estéreo deseada. Además, existen otros aspectos no técnicos a considerar, como decisiones creativas o artísticas, que derivan decisiones como la aplicación de efectos de audio como reverberación artificial o la automatización de la frecuencia de corte en un filtro.

Las decisiones de mezcla no suelen tomarse de forma aislada sin afectar a otras fuentes o comprometer el resultado. La interacción de estas decisiones de mezcla permite un constante equilibrio en el intento de lograr el mejor resultado posible.² Hoy en día, con la democratización de la tecnología, los productores tienen un mayor acceso a herramientas de producción musical. Ya sea analógico o digital, el encargado de la mezcla realiza los mismos procesos, como ecualización, compresión, distribución panorámica o ajuste del nivel general para optimizar la mezcla, por mencionar algunos. De manera análoga, con el avance de diferentes áreas de la inteligencia artificial (IA) —particularmente enfocándose en representación del conocimiento y razonamiento (*Knowledge Representation*

and Reasoning; KRR)³; el aprendizaje automático, también conocido como aprendizaje de máquina (*Machine Learning; ML*)⁴ o el procesamiento del lenguaje natural (*Natural Language Processing; NLP*)⁵— contamos con herramientas inteligentes que apoyan las actividades diarias humanas, como los automóviles de conducción o manejo autónomos, asistentes de voz inteligentes e incluso los sistemas de recomendación de música.⁶ En la música y en sus procesos internos, como la composición o la producción, nos hemos encontrado con inclusiones de la IA en todo el proceso de creación musical⁷ desde la composición hasta la masterización. En otras palabras, existen aplicaciones inteligentes en el mundo del audio y de la música debido al concepto de cognificación, el cual busca la aplicación de algoritmos de IA y que ha permitido el desarrollo de campos emergentes como la producción musical inteligente (*Intelligent Music Production; IMP*)⁸.

Para el alcance de este trabajo, presentamos *frmixerr*, un sistema con base en los avances y la injerencia que ha tenido el

3 M. Gelfond y V. Lifschitz, «The Stable Model Semantics for Logic Programming», *ICLP/SLP*, vol. 88 (agosto 1988): 1070-1080.

4 C. M. Bishop y N. M. Nasrabadi, *Pattern Recognition and Machine Learning* (Nueva York: Springer, 2006).

5 Manning, C. y H. Schütze, *Foundations of Statistical Natural Language Processing* (MIT Press, 1999).

6 M. Coeckelbergh, *AI Ethics* (MIT Press, 2020). Sameer Balaganur, «How Spotify's Algorithm Manages To Find Your Inner Groove», *Analytics India Mag*, 6 de enero de 2020, <https://analytic-sindiamag.com/how-spotifys-algorithm-manages-to-find-your-inner-groove/>

7 Carolina Sacristán-Ramírez, Flavio Everardo, Yamil Burguete y Bretch De Man, «AI in Music: Implications and Consequences of Technology Supporting Creativity», en *What AI Can Do*, edición de Manuel Cebral-Loureiro, Elvira G. Rincón-Flores y Gildardo Sánchez-Ante, 233-252 (Nueva York: Chapman and Hall/CRC).

8 Bretch De Man, Ryan Stables y Joshua D. Reiss, *Intelligent Music Production* (Nueva York: Routledge, 2019).

David Moffat y Mark B. Sandler, «Approaches in Intelligent Music Production», *Arts*, vol. 8, n.º 125 (2019).

A. Palladini, «Intelligent Audio Machines» (charla magistral en 4th Workshop on Intelligent Music Production WIMP-18, Huddersfield, Reino Unido, septiembre de 2018).

Eduardo Reck Miranda (Ed.), *Handbook of Artificial Intelligence for Music: Foundations, Advanced Approaches, and Developments for Creativity* (Springer Nature, 2021).

1 A. Case, *Mix smart: Professional Techniques for the Home Studio* (Taylor & Francis, 2012).

2 M. Senior, *Mixing Secrets for the Small Studio* (Routledge, 2018).

área de KRR en la producción musical inteligente, particularmente mediante uno de sus paradigmas llamado *Answer Set Programming* (ASP).

ASP⁹ es un formalismo basado en reglas para modelar y resolver problemas combinatorios (de optimización) intensos en conocimiento.¹⁰ Ha aumentado rápidamente el número de aplicaciones tanto en el ámbito académico como en la industria, como la planeación, planificación (*scheduling*), diseño de configuraciones, biología, logística¹¹ e incluso la música.¹² ASP y la música han demostrado aplicaciones principalmente en la composición, pero ha habido poco progreso en áreas de postproducción, particularmente en la mezcla.¹³

Retomando nuestra motivación, al ver la mezcla como un problema de optimización complejo, el cual embona perfectamente en ASP, este artículo describe las siguientes contribuciones:

- Presentar la metodología de *frmixerr*, un sistema inteligente, capaz de generar mezclas de forma automática. Este enfoque inteligente propone balances de niveles mediante la comparación de la magnitud general de frecuencias de la mezcla con una referencia. Para nuestros propósitos, usamos el ruido rosa como referencia.
- Demostrar la fiabilidad de nuestra propuesta inteligente mediante la imple-

9 Gelfond y Lifschitz, «The Stable Model...»; Torsten Schaub y Stefan Woltran, «Answer Set Programming Unleashed!», *KI-Künstliche Intelligenz*, vol. 32, n.º 2 (2018): 105–108.

10 Chitta Baral, *Knowledge Representation, Reasoning and Declarative Problem Solving* (Cambridge University Press, 2003).

11 Esra Erdem, Michael Gelfond y Nicola Leone, «Applications of Answer Set Programming», *AI Magazine*, vol. 37, n.º 3 (2016): 53–68.

Andreas Falkner *et al.*, «Industrial Applications of Answer Set Programming», *KI-Künstliche Intelligenz*, vol. 32, n.º 2 (2018): 165–176.

12 Sacristán-Ramírez *et al.*, «AI in Music...».

13 Flavio Everardo, «Over a Decade of Producing Music with Answer Set Programming. A Survey» (charla en Trends and Applications of Answer Set Programming TAASP Workshop, Klagenfurt, Austria, noviembre 2020).

mentación de una evaluación de escucha subjetiva comparando nuestra solución con mezclas hechas por personas con variedad en sus años de experiencia.

El resto del artículo se compone de la siguiente manera: la sección II presenta los antecedentes necesarios para el desarrollo de *frmixerr*, la que se detalla en la sección III. Las secciones IV y V explican el diseño e implementación de la prueba subjetiva, así como sus resultados, respectivamente. Finalmente, la sección VI cierra el texto discutiendo los resultados, así como futuras investigaciones.

2. Antecedentes

En esta sección, se describen los conceptos necesarios que sustentan nuestra propuesta. Principalmente, nuestra investigación se fundamenta en IMP y ASP.

2.1. Producción musical inteligente

Producción musical inteligente (*Intelligent Music Production*; IMP)¹⁴ es un área de investigación que incorpora IA en procesos de producción y postproducción como grabación, mezcla o *mastering*. En palabras de David Moffat y Mark B. Sandler¹⁵, la IMP tiene la premisa de cambiar fundamentalmente la manera en la que ingenieros y consumidores interactúan con la música, no solo permitiendo la colaboración entre humanos y sistemas inteligentes, sino también explorando y comprendiendo nuevas dimensiones e ideas dentro del espacio de la producción musical. Alineado con los conceptos de IA establecidos por Russell y

14 De Man, Stables y Reiss, *Intelligent Music Production...*

Moffat y Sandler, «Approaches in Intelligent...».

Palladini, «Intelligent Audio Machines».

Miranda, *Handbook of Artificial...*

15 Moffat y Sandler, «Approaches in Intelligent...».

Norvig¹⁶, estos sistemas inteligentes deben ser capaces de tomar decisiones automáticamente, pensar y razonar sobre esas decisiones, sacar conclusiones y aprender de la experiencia al acumular conocimiento.

Adicionalmente, un sistema general de IMP sigue una arquitectura como se describe en la figura 1. En principio, un sistema inteligente para la producción musical considera diferentes tipos de datos de entrada que se clasifican en materiales sonoros y no sonoros. Estos últimos corresponden a partituras o textos, mientras que los sonoros son archivos de audio. El sistema toma esta información en conjunto con la interacción que viene por parte del usuario para su respectivo análisis. Esto desencadena decisiones validables por el usuario y, en caso de aprobarlo, se realiza una acción específica antes de que el resultado sea devuelto al usuario.

es visto como una caja negra y no se requiere interacción humana. En otras palabras, el sistema toma el control total de los procesos internos y la toma de decisión.

2. **Independiente:** asistente en el que se delegan algunas tareas específicas para proporcionar un resultado cercano al producto terminado, con el usuario actuando como supervisor.
3. **Sugestivo:** la participación del sistema se limita al análisis y la propuesta de recomendaciones, mientras que el usuario tiene control completo sobre el resultado.
4. **Perceptivo:** el sistema proporciona información adicional, datos textuales, visualizaciones, para aconsejar al usuario sobre información o conocimiento sin realizar modificaciones a los datos de entrada.

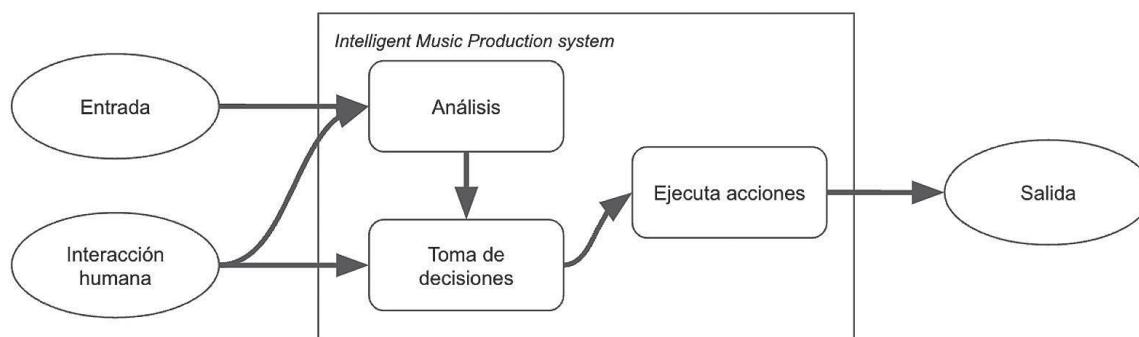


Figura 1: diagrama de un sistema de producción musical inteligente tomado de *AI in Music: Implications and Consequences of Technology Supporting Creativity*¹⁷.

La IA en un sistema de esta naturaleza no se aplica de la misma manera. Se requieren diferentes niveles predefinidos de automatización o grados de participación de la IA para cumplir los propósitos definidos. Un sistema de IMP se puede clasificar en una de estas cuatro áreas¹⁸:

1. **Automático:** un enfoque completamente automatizado donde el sistema

En otras palabras, la participación de la IA en un sistema de IMP varía desde ser un sistema de asesoramiento que utiliza información y conocimiento, resultando en decisiones más rápidas o mejores para que el usuario las tome, hasta controlar todo el proceso de los datos de entrada siguiendo un enfoque de caja negra. En el contexto de la mezcla, podríamos ejemplificar un sistema de IMP capaz de únicamente recomendar un procesamiento *mid-side* para corregir los problemas de fase en una señal estéreo, hasta el punto de poder realizar una mezcla en su totalidad aplicando balances, paneo, compresión y ecualización. En el medio, encontramos niveles que re-

16 Stuart Russell y Peter Norvig, *Artificial Intelligence: A Modern Approach* (Pearson, 2020).

17 Sacristán-Ramírez et al., «AI in Music...».

18 Sacristán-Ramírez et al., «AI in Music...».

lajan el grado de autonomía para resolver problemas específicos, como hacer exclusivamente el balance de las fuentes sonoras o gestionar el rango dinámico de un número de pistas.¹⁹

Diferentes áreas de la IA han sido probadas en diversas tareas dentro del proceso de creación o composición musical. A su vez, la IA ha mostrado avances en el campo de la composición e incluso para la producción musical.²⁰ En este artículo nos centramos en las cualidades que tiene *Answer Set Programming* para integrarse en un sistema de producción musical inteligente.

2.2. Answer Set Programming

ASP es un formalismo basado en reglas para modelar y resolver problemas combinatorios (de optimización) intensos en conocimiento²¹ con orígenes en el área de KRR. Sus raíces parten de la programación lógica, permitiendo que ASP tenga una sólida teoría, así como implementaciones de alto rendimiento. Es por eso que ha demostrado ser exitoso en el ámbito académico como en la industria, resolviendo problemas complejos como planeación, planificación (*scheduling*), diseño de configuraciones, biología, logística²² e incluso la música.²³

ASP tiene más de una década de historia en la música, particularmente en composición,²⁴ con reglas y restricciones específicas para un estilo musical o enfoques independientes del género con conocimiento especializado para cubrir reglas musicales generales. Sin embargo, en tér-

¹⁹ Bretch De Man, «Audio effects 2.0: Rethinking the music production workflow (ADC'17)», video en YouTube, 35:17, noviembre 20 de 2017, <https://www.youtube.com/watch?v=yIcqNR7by9o>.

Moffat y Sandler, «Approaches in Intelligent...»; Palladini, «Intelligent Audio Machines».

Sacristán-Ramírez et al., «AI in Music...».

²⁰ Sacristán-Ramírez et al., «AI in Music...».

²¹ Baral, *Knowledge Representation...*

²² Erdeim, Gelfond y Leone, «Applications of Answer Set Programming».

Falkner et al., «Industrial Applications...».

²³ Everardo, «Over a Decade...».

²⁴ Everardo, «Over a Decade...».

minos de manejo de señales de audio, se ha hecho poco progreso en la postproducción, específicamente en la mezcla. Everardo²⁵ realizó un primer esfuerzo de un sistema inteligente capaz de proponer el balance y paneo utilizando un conjunto de conocimientos sobre las mejores prácticas de mezcla de ingenieros provenientes de la literatura.

Uno de los principales beneficios de ASP como paradigma de lenguaje de programación, en contraste con la programación tradicional, es que nos permite especificar el problema por medio de reglas y restricciones en lugar de codificar la solución de un problema en cuestión. En la práctica, dejamos que un solucionador de ASP (*solver*) busque por medio de inferencias lógicas, aprendizaje y razonamiento, si el problema es satisfacible y si tiene una o varias soluciones. En este sentido, podemos enseñar al *solver* a entender las reglas y restricciones para realizar la mezcla en lugar de hacerla por nuestra cuenta.

En la sección siguiente, explicaremos la metodología de cómo nuestro sistema inteligente realiza una mezcla inicial por medio de ASP.

3. *frmixerr*

Esta sección presenta la metodología utilizada para generar mezclas por parte del sistema inteligente llamado *frmixerr*. El sistema, por ahora, se encuentra en la intersección de las categorías de un sistema automático y uno independiente. Se dice que es un sistema automático ya que el usuario no supervisa ni dirige ninguno de los procesos internos y la IA es responsable del análisis, toma de decisiones y la

²⁵ Flavio Everardo, «Towards an Automated Multi-track Mixing Tool Using Answer Set Programming» (ponencia en 14th Sound and Music Computing Conference, julio de 2017).

realización de las acciones pertinentes. Por otro lado, se puede decir que es un sistema independiente ya que se puede ver como un asistente al que se le ha delegado una actividad específica, por ejemplo, el balance de diferentes pistas de un *multitrack*. Con ello, no se busca que proporcione un resultado final, sino cercano a terminado, con el usuario actuando como supervisor. La finalidad de *frmixerr* a largo plazo es que sirva como un asistente para el proceso de mezcla propiciando un diálogo e intercambio de ideas junto al productor o ingeniero. Es evidente que hoy en día el proceso de la producción musical es hecho en su totalidad por personas en donde el *hardware* o *software* no proporcionan ningún tipo de comunicación o colaboración, así como ningún atributo que explique lo que la IA está realizando.²⁶ Las aplicaciones más apetecibles para los usuarios, en su interés por utilizar IA en la música, son aquellas que, en la práctica, consumen mucho tiempo, son tediosas y de naturaleza técnica: es incierto el efecto que tendrá en el resultado, pero ahorrará tiempo y energía para el músico o productor que quiere centrarse en desarrollos creativos. Estas son las áreas donde la mayoría de los especialistas creativos prefieren la implementación de soluciones de IA. Por otro lado, se oponen mucho a las propuestas no supervisadas de tareas creativas.²⁷ En otras palabras, la mayoría de los profesionales prefieren el uso

de la IA como asistente que mejora su flujo de trabajo para acelerar su proceso creativo.

Tradicionalmente, el ingeniero de mezcla realiza el balance de las diferentes pistas en la sesión ajustando cada uno de los *faders* o atenuadores en una superficie de control, consola de mezcla o dentro de una estación digital de trabajo (*Digital Audio Workstation* o DAW). Sea analógico o digital, la equidad de las fuentes sonoras respeta un balance en las frecuencias de la mezcla de manera similar al comportamiento del ruido rosa (*pink noise*).²⁸ El ruido rosa, también conocido como $1/f$, al momento de ser visto en un analizador con filtros de octava, se puede apreciar que todas las bandas de octava tienen el mismo nivel sonoro y cuyo nivel de presión sonora se caracteriza por una densidad espectral inversamente proporcional a la frecuencia (decreciendo 3 dB por octava) como se puede ver en su periodograma en la figura 2. El ruido rosa es utilizado con diversos fines como en mediciones acústicas, pero también se encuentra en el balance de frecuencias en la música. Voss y Clarke²⁹, demostraron que las fluctuaciones de tono e intensidad en el habla y

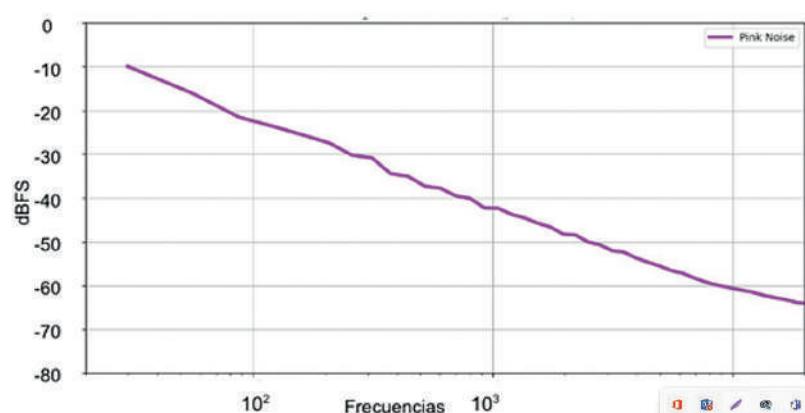


Figura 2: periodograma del ruido rosa. La música corresponde al ruido rosa.

26 David Gunning et al., «XAI—Explainable artificial intelligence», *Science Robotics*, vol. 4, n.º 37 (2019).

Explainable AI (XAI): conjunto de procesos y métodos que permiten a los usuarios humanos comprender y confiar en los resultados creados por IA.

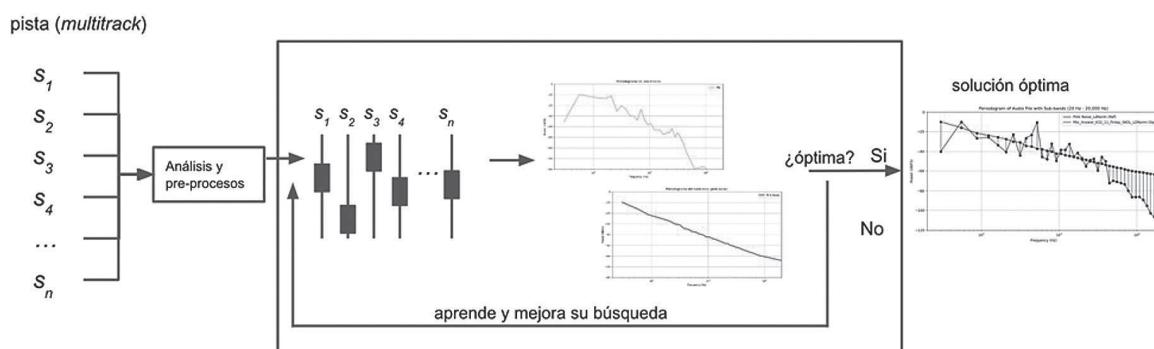
27 Andreas Pfeiffer, *Creativity and technology in the age of AI: Research Report (USA, Europe and Japan)* (Pfeiffer Report, 2018), <https://www.pfeifferreport.com/essays/creativity-and-technology-in-the-age-of-ai/>

28 Bobby Owsinski, *The Mixing Engineer's Handbook (5th ed.)* (Bobby Owsinski Media Group, 2022).

29 Richard F. Voss y John Clarke, «“1/f noise” in music and speech», *Nature*, 258 (1975): 317–318.

Richard F. Voss y John Clarke, «“1/f noise” in music: Music from 1/f noise», *The Journal of the Acoustical Society of America*, vol. 63, n.º 1 (1978): 258–263.

El sistema *frmixerr*, programado en *python* realiza las acciones desplegadas en la figura 3. Primero comienza analizando el *multitrack* dado por el usuario normalizando en pico (*peak normalization*) cada una de sus pistas o señales ($s_1 \dots s_n$) para tener la mayor amplitud posible en la mezcla. Después, entra al proceso iterativo en donde el *solver*³⁰ propone un balance en decibeles a escala completa (dBFS) y se realiza la atenuación de cada señal al mismo tiempo de que se genera la suma de señales (Anexo I). Una vez generada la mezcla, se calcula su periodograma y se compara contra el periodograma del ruido rosa. El sistema genera mezclas hasta encontrar la óptima. Se define una solución óptima a la configuración de *faders* cuyo periodograma de dicha mezcla es la más cercana al ruido rosa. Si durante la búsqueda de una mejor solución se encuentra con una diferencia de periodogramas mayor a la que tenga como óptima en



ese momento, se descarta esa solución y aprende de esa configuración no fructífera para mejorar su búsqueda. La figura 4 muestra un gráfico de cuatro periodogramas en donde se ilustra la constante mejora acercando la respuesta en frecuencia de la mezcla al ruido rosa.

La diferencia entre periodogramas se hace en 42 bandas siguiendo las me-

³⁰ Se usó *clingo* como *solver* del grupo de Potsdam Answer Set Solving Collection (Potassco). <https://potassco.org/clingo/>

didas de anchos de filtros utilizadas en el ancho de banda rectangular equivalente (*Equivalent Rectangular Bandwidth*; ERB³¹) el cual es una medida utilizada en el campo de la percepción auditiva para cuantificar la resolución de frecuencia del sistema auditivo humano.

El ERB representa el ancho de banda de un filtro que tiene una forma equivalente al filtro auditivo en una frecuencia central específica. La excitación total que cae dentro de cualquier filtro ERB dado se comprime de manera equivalente, esto resulta en una contribución uniforme al volumen total. Es por esto que ERB proporciona un mecanismo relacionado a la compresión, el enmascaramiento y la sonoridad. Finalmente, nuestro sistema proporciona la configuración en dBFS de cada uno de los tracks. La tabla 1 nos muestra los diferentes valores por pista usados en las mezclas mostradas en la figura 4.

Figura 3: arquitectura de *frmixerr*. El sistema genera mezclas buscando refinar en cada interacción la diferencia entre ambos periodogramas (el generado en la mezcla y el ruido rosa) hasta que la solución es óptima. De lo contrario, se descarta y se procede a buscar una mejor opción.

³¹ Brian Glasberg y Brian C. J. Moore, «A Model of Loudness Applicable to Time-Varying Sounds», *Journal of the Audio Engineering Society*, vol. 50, n.º 5 (2002): 331-342.

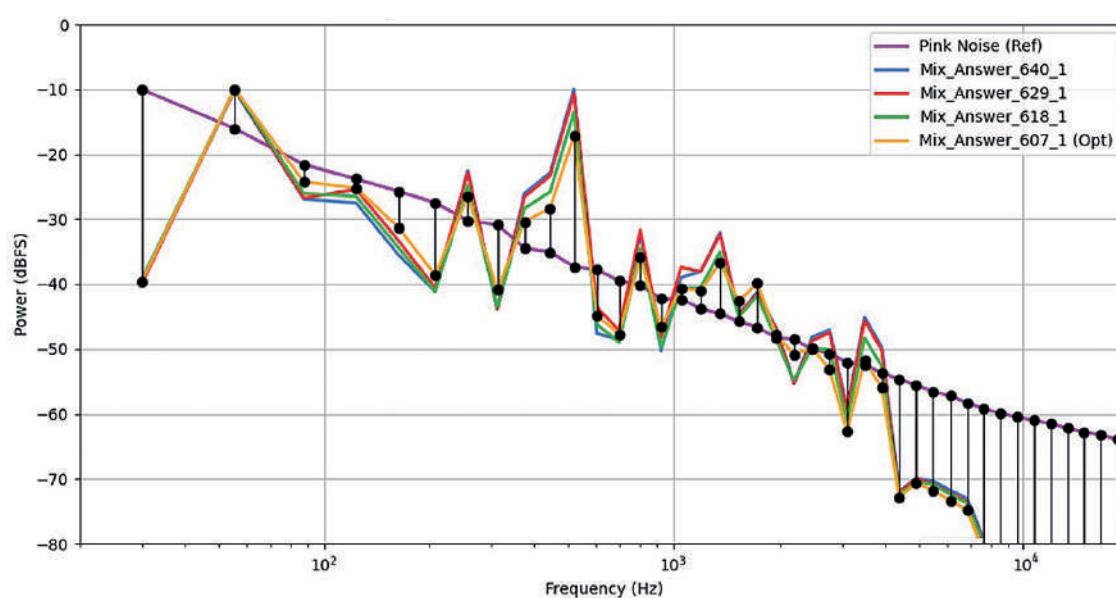
A. J. Simpson, M. J. Terrell y Joshua D. Reiss, «A Practical Step-By-Step Guide to the Time-Varying Loudness Model of Moore, Glasberg, and Baer (1997; 2002)», *Audio Engineering Society Convention 134* (2013).

4. Prueba de escucha-evaluación subjetiva

4.1. Estímulos

Como un primer acercamiento, para tener una referencia inicial del rendimiento de nuestro sistema, se comparó una de las mezclas generadas con mezclas realizadas por una suma simple de señales normalizadas y por cuatro ingenieros de mezcla humanos con diferentes niveles de desarrollo

que no se permitió utilizar ningún efecto o proceso de audio. Para hacerlo más homogéneo, las mezclas resultantes se ajustaron a -23 LUFS para obtener el mismo volumen (*loudness*) y para eliminar el sesgo hacia muestras más fuertes (o más suaves) durante la prueba de audición. Dentro de las múltiples soluciones generadas por *frmixer*, la mezcla a utilizar fue elegida por los autores del presente artículo, previo a tener las mezclas generadas por humanos.



profesional. A cada ingeniero de mezcla se les pidió que hicieran una mezcla de la versión editada de *Same kind of life* por *Finlay*.³² Dicha mezcla del género pop de 24 segundos de duración consiste en solo el balance en mono de las 16 pistas que contiene el *multitrack*. Esta sesión cuenta con elementos como voz, bajo, guitarras, teclados, bombo, caja y micrófonos de batería (*overheads*). De la misma manera, nuestra mezcla generada por el sistema y la suma de señales también fueron generadas en mono. La razón principal para realizar esta prueba en monoaural, es para evitar distracciones del panorama. De igual forma, cabe resaltar

Figura 4: proceso de optimización de la mezcla. Los valores de 640_1 hasta 607_1 indican la suma de todas las diferencias en dBFS de las 42 bandas. La mezcla Mix_Answer_640_1(azul) es la más alejada del ruido rosa, mientras las mezclas 629 (roja) y 618 (verde) son resultados más cercanos. La mezcla 4 Mix_Answer_607_1(amarillo) es la mezcla óptima.

4.2. Diseño y ejecución de la prueba

El método de prueba de escucha que utilizamos fue una prueba de estímulos múltiples con ancla oculta. En otras palabras, se trata de una prueba MUSHRA (estímulo múltiple con ancla y referencia oculta³³) excepto que en este caso no había referencia. El ancla

³² Multitracks disponibles bajo una licencia Creative Commons, <https://www.cambridge-mt.com/ms/mtk/>

³³ C. Völker et al., «Modifications of the Multi Stimulus Test with Hidden Reference and Anchor (MUSHRA) for Use in Audiology», *International Journal of Audiology*, vol. 57, n.º 3 (2018): S92-S104.

oculta es la suma monofónica del audio sin procesar, donde cada canal ha sido normalizado en pico. La intención de esta mezcla es que sea identificable como la más baja de todas. Sin embargo, es posible que incluso sujetos expertos en el área califiquen esta ancla por encima de la calificación menor o incluso aceptable dependiendo de la percepción de las otras mezclas.³⁴ Para esta prueba se utilizó una escala de nueve valores en total a través de cinco identificadores: mala, pobre, razonable, buena y excelente. Los otros cuatro valores son las ponderaciones intermedias entre cada identificador. La escala se define con valores desde 1 para una mezcla considerada como mala hasta 9 para una excelente.

El orden de las canciones, así como su identificador fueron aleatorios para evitar cualquier tipo de sesgo o diferencia de interpretación de los temas relacionada con el orden de reproducción. El estudio consistió de una sola sesión *offline* con una duración de entre 10 y 20 minutos. Es decir, los audios fueron puestos a disposición de los evaluadores para su escucha y juicio.

Un total de 18 profesionales de la industria (3 mujeres y 15 hombres) con diferentes perfiles, pero con experiencia en mezcla fueron invitados a participar en esta prueba. A cada participante se les pidió llenar un consentimiento informado dando por hecho la realización del ejercicio utilizando audífonos o monitores de estudio. La edad media de los participantes fue de 31.5 años y ninguno identificó sufrir algún padecimiento auditivo. Se pidió a los sujetos su calificación de las seis mezclas del mismo fragmento de canción de acuerdo con lo que para cada quien era sinónimo de calidad. Los sujetos no tenían ninguna información sobre el contenido del audio o el objetivo de la investigación antes de realizar la prueba.

³⁴ Bretch De Man y Joshua D. Reiss, «A Knowledge-Engineered Autonomous Mixing System», *Audio Engineering Society Convention*, 135 (2013).

	Mezclas generadas por frmixerr. (dBFS)			
Pistas del multitrack <i>Same kind of life</i> por Finlay	640_1	629_1	618_1	607_1
01_Kick.wav	-28	-28	-28	-28
02_KickSample.wav	-30	-30	-30	-30
03_SnareSample.wav	-27	-27	-27	-27
04_Tom1.wav	-28	-28	-28	-28
05_Tom2.wav	-28	-20	-21	-29
06_Tom3.wav	-30	-30	-30	-30
07_Cymbal.wav	-27	-27	-27	-27
08_Bass.wav	-27	-27	-27	-27
09_ElecGtr1.wav	-30	-30	-30	-30
010_ElecGtr2.wav	-27	-27	-27	-27
011_ElecGtr3.wav	-30	-30	-30	-30
012_Piano1.wav	-28	-28	-28	-28
013_Piano2.wav	-30	-21	-28	-28
014_Organ.wav	-30	-30	-23	-22
015_LeadVox.wav	-18	-18	-21	-22
016_BackingVox.wav	-27	-27	-27	-27

Tabla 1: ejemplo del proceso de optimización de los periodogramas mostrados en la figura 4. El sistema busca reducir la diferencia en dBFS entre su periodograma con respecto al del ruido rosa. Se puede notar entre los cambios más sobresalientes entre las mezclas 607 y 640, que 12 de las 16 pistas no ajustaron sus valores del *fader*. Sin embargo, es notorio el ajuste en las pistas Piano2, Organ y LeadVox.

5. Resultados

Se tomaron las calificaciones para cada una de las seis mezclas por parte de los 18 evaluadores. La mezcla 2 fue generada por nuestro sistema *frmixerr* y, como se mencionó en la sección anterior, usamos un ancla negativa identificada como la mezcla 5. En otras palabras, se puede interpretar como una manipulación experimental con expectativa de resultados bajos. Las mezclas 1, 3, 4 y 6 corresponden a las producciones realizadas por humanos especialistas en mezcla con diversa variedad en sus años de experiencia (desde seis meses hasta diez años). Existen trabajos³⁵ en los que hacen énfasis en que es posible que el ancla pueda ser evaluada al mismo nivel que los otros estímulos.

La escala de sus calificaciones se encuentra entre la calificación mala, con un

³⁵ De Man y Reiss, «A Knowledge-Engineered...».

valor de 1, y excelente, con un valor de 9. El análisis de datos se realizó de la siguiente manera: se calculó la estadística descriptiva pertinente (ver figura 5), posteriormente se realizó un análisis de varianza (ANOVA) de una vía para detectar las diferencias entre grupos. Todo el trabajo de análisis estadístico se realizó con el software Jamovi³⁶, que es un programa basado en el lenguaje R³⁷.

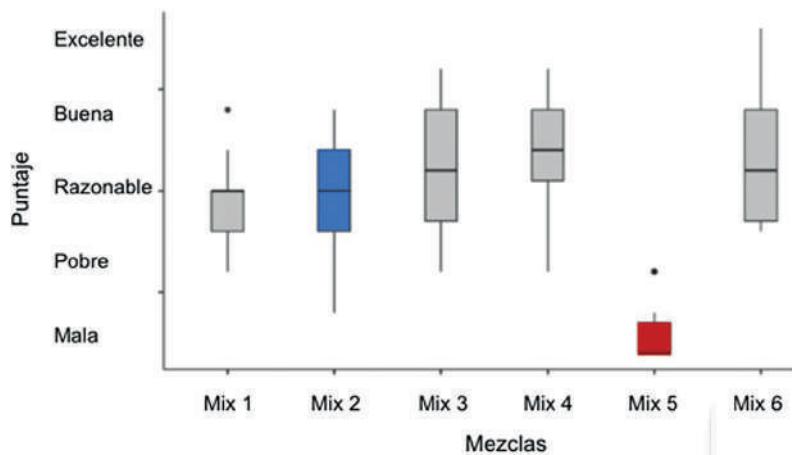


Figura 5: se presentan los diagramas de caja de la evaluación para cada una de las mezclas. La mezcla 5 (rojo) es una manipulación experimental con un resultado bajo esperado (ancla). La mezcla 2 (azul) fue producida por nuestro sistema *fmixerr*.

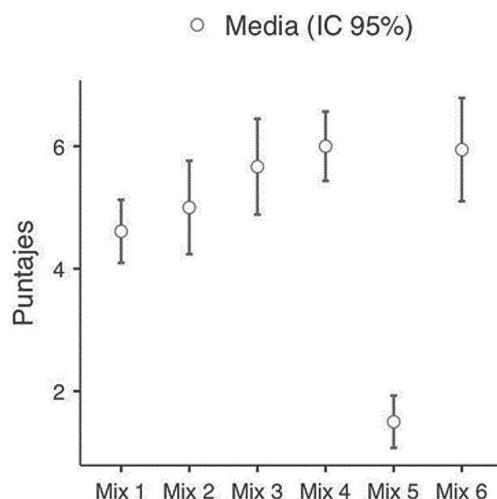


Figura 6: se observan los puntajes medios por mezcla con los intervalos de confianza a 95 %. Se puede observar que la mezcla 5 se encuentra distante del resto de las medias.

Para el análisis descriptivo, se calculó el promedio y desviación estándar de la evaluación de cada mezcla (ver tabla 2). En general se observa que las medias se mantienen entre los valores de 4.6 hasta 6.0 salvo por la mezcla ancla. Este control experimental sirve para ver que todos los evaluadores tienen la capacidad de identificación de las mezclas además de un control de atención. Por su parte, la mezcla de nuestro sistema (mezcla 2), se encuentra aproximadamente al nivel de las otras mezclas producidas por especialistas humanos. De igual

manera, las desviaciones estándares se encuentran en el rango entre 0.85 y 1.69. Con los resultados anteriores se procedió al análisis inferencial para ver si las diferencias detectadas en la variación de las medias son significativas. Por las condiciones de la muestra (18 observaciones por cada mezcla) se decidió correr algunas pruebas para verificar los supuestos de la prueba estadística. Por los valores obtenidos en la prueba de Shapiro-Wilk (W = 0.981, p-valor = 0.135), se puede asumir que la muestra presenta un comportamiento normal. En el caso de la prueba de Levene (Fox y Weisberg, 2020) de igualdad de varianzas ($F = 3.87$, $gl_1=5$, $gl_2=102$, p-valor = 0.003) se identificó que las varianzas son desiguales, por lo tanto, el ANOVA de un factor utilizó la corrección de Welch para desigualdad de varianzas.

³⁶ The Jamovi Project, «jamovi (Versión 2.3) [Computer Software]», Jamovi, 2022, <https://www.jamovi.org>.

³⁷ R Core Team, «R: A Language and environment for statistical computing. (Version 4.1) [Computer software]», Cran.r Project, 2021, <https://cran.r-project.org> («R packages» extraídos de MRAN snapshot 2022-01-01).

Mezcla	Mezcla	Mezcla
Mix 1	4.6	1.0369
Mix 2	5.0	1.5339
Mix 3	5.7	1.5718
Mix 4	6.0	1.1376
Mix 5	1.5	0.8575
Mix 6	5.9	1.6968

Tabla 2. Se observan los valores de las medias y desviaciones estándares para cada una de las mezclas. No hubo datos faltantes, todos los evaluadores calificaron todas las mezclas. La mezcla 2 es la generada por el algoritmo y la mezcla 5 es la mezcla ancla con calificación baja esperada.

Al realizar el ANOVA de un factor, se encontró que existe al menos una diferencia entre las mezclas ($F = 50.2$, $gl_1 = 5$, $gl_2 = 47.1$, p -valor < 0.001). Por lo anterior, se decidió graficar los valores de las medias con sus intervalos de confianza al 95 % (ver figura 6). Como se puede observar, la mezcla 5 tuvo un comportamiento significativamente diferente al resto de los grupos, todas las comparaciones por pares tuvieron un p -valor < 0.001. La única diferencia significativa entre el resto de las mezclas fue entre la mezcla 1 y la 4 (diferencia = -1.39, p -valor = 0.007), las cuales correspondieron a las preferencias de los evaluadores a dos mezclas realizadas por ingenieros de mezcla.

Se puede concluir que no existió una diferencia en la evaluación subjetiva entre las mezclas realizadas por profesionales y la mezcla realizada con el algoritmo. Los resultados también muestran que los evaluadores identificaron y calificaron adecuadamente la mezcla 5, que es una manipulación con baja respuesta esperada. Por lo tanto, la mezcla generada por *frmixerr* tiene una aceptación similar a la de profesionales con diferentes años de experiencia en mezcla.

6. Discusión

ASP ha demostrado su viabilidad en la resolución de un problema de optimización de alta complejidad como la mezcla, posicionándose como una alternativa dentro del área de producción musical inteligente. Para el alcance de este trabajo, se propuso un sistema llamado *frmixerr* para generar soluciones balanceadas siguiendo como criterio de optimización una respuesta en frecuencia. Para este caso, el ruido rosa. El sistema es capaz de entregar una configuración inicial en un formato legible para humanos para que cualquier ingeniero de

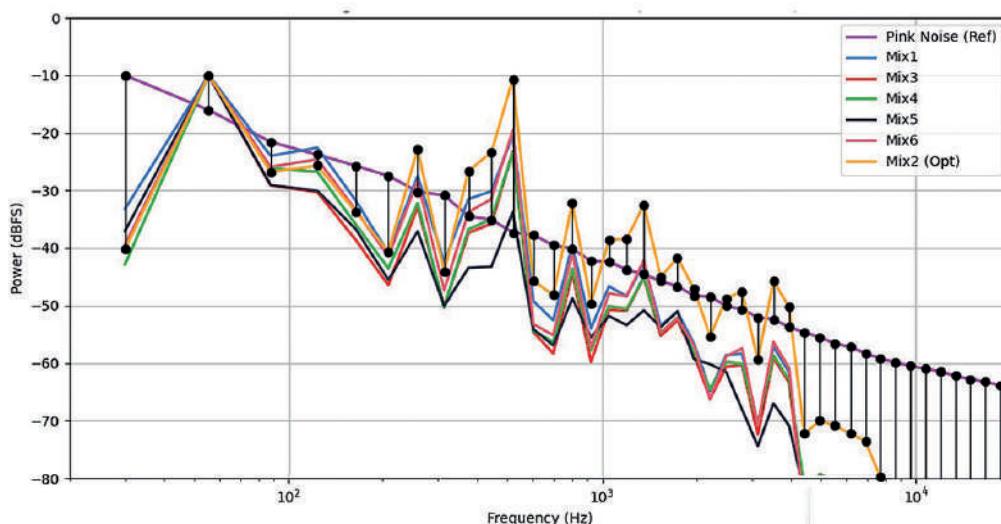


Figura 7: comparación de periodogramas de las seis mezclas utilizadas en la prueba. Cabe mencionar que la mezcla 2 es la más cercana al ruido rosa. La mezcla ancla (5) es la más alejada a la curva de la respuesta objetivo.

audio especializado comience a abordar la mezcla. Los resultados de la evaluación subjetiva demostraron criterios de elección variados, pero coincidiendo en cuatro de seis

mezclas estando la mezcla generada por *frmixerr* dentro de las elegibles y haciendo énfasis en el gusto del evaluador sin ningún otro criterio adicional. En otras palabras, los resultados avalan que nuestra mezcla es igual de probable de poder ser elegida como cualquiera generada por humanos con años de experiencia. No existe una diferencia significativa con respecto a las generadas por humanos.

Es importante destacar que este primer experimento, no da una resolución definitiva, sino que motiva a extender la prueba para conocer el potencial de una herramienta inteligente en la mezcla. Para ello, se listan algunas de las consideraciones inmediatas: ampliar el número de *multitracks*; utilizar un ambiente controlado de escucha; extender *frmixerr* con la posibilidad de generar mezclas que tomen en cuenta otros procesos de mezcla, como paneo, amplitud estéreo, minimización de enmascaramiento, rango dinámico, entre otros. Para ello, es inminente expandir el conjunto de módulos de medición, para permitir una mejor escucha y procesamiento, así como aplicar la correlación cruzada para medir cancelaciones de fase y amplitud estéreo y continuar con el modelo de bandas ERB para medir y minimizar enmascaramiento, por mencionar algunos.

Para este trabajo, se utilizó el ruido rosa como la curva de frecuencia estándar. Sin embargo, vale la pena analizar diferentes respuestas en frecuencia por género. La mezcla generada por *frmixerr* fue la más cercana a la referencia. A pesar de que influyó de forma positiva en el resultado, no

significó que fuese la mejor opción. Combinar otros procesos de audio puede ayudar a moldear la respuesta en frecuencia de la mezcla con una curva referente.

En cuanto al diseño y ejecución de la prueba subjetiva, es necesario brindar a los sujetos con más criterios de evaluación que un simple «me gusta» como respuesta a la subjetividad de la calidad sonora. Debe ser posible poner en juicio otros criterios como el rango dinámico, amplitud estéreo, claridad en las fuentes sonoras, entre otros. De igual manera, nuestro sistema es capaz de generar muchas mezclas válidas. Para garantizar una mejor representación, es necesario realizar una evaluación colegiada previa de mezclas generadas por la IA y no como selección arbitraria de los autores.

Finalmente, se deja para futuras investigaciones evaluar el rendimiento del sistema para demostrar la versatilidad de ASP en la producción musical. Es decir, aclarar dudas sobre cuánto tiempo tarda en generar estas mezclas. Informalmente se puede decir que el sistema genera el primer resultado en cuestión de pocos segundos, aunque la mezcla óptima puede tardar algunos minutos.

7. Agradecimientos

Los autores agradecen a los ingenieros de mezcla Jay Castro, Cesar Arturo Díaz Ordaz, Sebastián Portillo y Guillermo Andrés Vega Castellanos, así como a las y los 18 profesionales de la industria que participaron en la evaluación subjetiva.

Referencias

- Balaganur, Sameer. «How Spotify's Algorithm Manages To Find Your Inner Groove». *Analytics India Mag*, 6 de enero de 2020. <https://analyticsindiamag.com/how-spotifys-algorithm-manages-to-find-your-inner-groove/>
- Baral, Chitta. *Knowledge Representation, Reasoning and Declarative Problem Solving*. Cambridge University Press, 2003.
- Bishop, C. M. y N. M. Nasrabadi. *Pattern Recognition and Machine Learning*. Nueva York: Springer, 2006.
- Case, A. *Mix smart: Professional techniques for the home studio*. Taylor & Francis, 2012.
- Coeckelbergh, M. *AI Ethics*. MIT Press, 2020.
- De Man, Bretch. «Audio effects 2.0: Rethinking the music production workflow (ADC'17)». Video en YouTube, 35:17. Noviembre 20 de 2017. <https://www.youtube.com/watch?v=yJcqNR7by9o>
- De Man, Bretch y Joshua D. Reiss. «A Knowledge-Engineered Autonomous Mixing System». *Audio Engineering Society Convention*, 135 (2013).
- De Man, Bretch, Ryan Stables y Joshua D. Reiss. *Intelligent Music Production*. Nueva York: Routledge, 2019.
- Erdem, Esra, Michael Gelfond y Nicola Leone. «Applications of Answer Set Programming». *AI Magazine*, vol. 37, n.º 3 (2016): 53–68.
- Everardo, Flavio. «Towards an Automated Multitrack Mixing Tool Using Answer Set Programming». Ponencia en 14th Sound and Music Computing Conference. Julio de 2017.
- . «Over a Decade of Producing Music with Answer Set Programming. A Survey». Charla en Trends and Applications of Answer Set Programming (TAASP) Workshop, Klagenfurt, Austria, noviembre 2020.
- Falkner, Andreas, Gerhard Friedrich, Konstantin Schekotihin, Richard Taupe y Erich C. Teppan. «Industrial Applications of Answer Set Programming». *KI-Künstliche Intelligenz*, vol. 32, n.º 2 (2018): 165–176.
- Fox, J. y S. Weisberg. «car: Companion to Applied Regression. [R package]». 2020. <https://cran.r-project.org/package=car>.
- Gebser, M., R. Kaminski, B. Kaufmann y T. Schaub. «Answer Set Solving in Practice». *Synthesis Lectures on Artificial Intelligence and Machine Learning*, vol. 6, n.º 3 (2012): 1–238.
- Gelfond, M. y V. Lifschitz. «The Stable Model Semantics for Logic Programming». *ICLP/SLP*, vol. 88 (agosto 1988): 1070–1080.
- Glasberg, Brian y Brian C. J. Moore. «A Model of Loudness Applicable to Time-Varying Sounds». *Journal of the Audio Engineering Society*, vol. 50, n.º 5 (2002): 331–342.
- Gunning, David, Mark Stefk, Jaesik Choi, Timothy Miller, Simone Stumpf y Guang-Zhong Yang. «XAI—Explainable artificial intelligence». *Science Robotics*, vol. 4, n.º 37 (2019).
- Lewis, M., Y. Liu, N. Goyal, M. Ghazvininejad, A. Mohamed, O. Levy y L. Zettlemoyer. *Bart: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension*. arXiv preprint arXiv:1910.13461. (2019).
- Manning, C. y H. Schütze. *Foundations of Statistical Natural Language Processing*. MIT Press, 1999.
- Miranda, Eduardo Reck (Ed.). *Handbook of Artificial Intelligence for Music: Foundations, Advanced Approaches, and Developments for Creativity*. Springer Nature, 2021.
- Moffat, David y Mark B. Sandler. «Approaches in Intelligent Music Production». *Arts*, vol. 8, n.º 125 (2019).
- Owsinski, Bobby. *The Mixing Engineer's Handbook (5th ed.)*. Bobby Owsinski Media Group, 2022.
- Palladini, A. «Intelligent Audio Machines». Charla magistral en 4th Workshop on Intelligent Music Production (WIMP-18), Huddersfield, Reino Unido (vol. 14), septiembre de 2018.
- Pfeiffer, Andreas. *Creativity and technology in the age of AI: Research Report (USA, Europe and Japan)*. Pfeiffer Report, 2018. <https://www.pfeifferreport.com/essays/creativity-and-technology-in-the-age-of-ai/>

- R Core Team. «R: A Language and environment for statistical computing. (Version 4.1) [Computer software]». Cran.r Project, 2021. <https://cran.r-project.org>. («R packages» extraídos de MRAN snapshot 2022-01-01).
- Russell, Stuart y Peter Norvig. *Artificial Intelligence: A Modern Approach*. Pearson, 2020.
- Sacristán-Ramírez, Carolina, Flavio Everardo, Yamil Burguete y Bretch De Man. «AI in Music: Implications and Consequences of Technology Supporting Creativity». En *What AI Can Do*. Edición de Manuel Cebal-Loureda, Elvira G. Rincón-Flores y Gildardo Sánchez-Ante, 233-252. Nueva York: Chapman and Hall/CRC.
- Schaub, Torsten y Stefan Woltran. «Answer Set Programming Unleashed!». *KI-Künstliche Intelligenz*, vol. 32, n.º 2 (2018): 105-108.
- Senior, M. *Mixing secrets for the small studio*. Routledge, 2018.
- Simpson, A. J., M. J. Terrell y Joshua D. Reiss. «A Practical Step-By-Step Guide to the Time-Varying Loudness Model of Moore, Glasberg, and Baer (1997; 2002)». *Audio Engineering Society Convention 134* (2013).
- The Jamovi Project. «jamovi (Versión 2.3) [Computer Software]». Jamovi, 2022. <https://www.jamovi.org>.
- Voss, Richard F. y John Clarke. «“1/f noise” in music and speech». *Nature*, 258 (1975): 317-318.
- . «“1/f noise” in music: Music from 1/f noise». *The Journal of the Acoustical Society of America*, vol. 63, n.º 1 (1978): 258-263.
- Völker, C., T. Bisitz, R. Huber, B. Kollmeier y S. M. Ernst. «Modifications of the Multi Stimulus Test with Hidden Reference and Anchor (MUSHRA) for Use in Audiology». *International Journal of Audiology*, vol. 57, n.º 3 (2018): S92-S104.

ANEXO 1

La programación de programas bajo el paradigma de Answer Set Programming sigue una metodología declarativa. Esto quiere decir, que, en lugar de programar la solución al problema, es necesario establecer las reglas y restricciones dentro del dominio de dicho problema para que el *solver* sea capaz decírnos si no existe solución alguna, si existe una o diversas. El objetivo de este artículo no es introducir el lenguaje ni sus diferentes constructores, sino ejemplificar la factibilidad de modelar un problema de manera declarativa en una forma compacta y concisa. Para mayor detalles sobre cómo modelar un programa, así como los fundamentos de ASP, puede revisar Gebser *et al.*³⁸

A continuación, se enlista el programa utilizado para proponer las mezclas usando la sintaxis del sistema *clingo*.

Las primeras dos líneas indicarán el número de pistas y si existe preferencia por alguna de estas pistas en particular como el *track* principal o *lead*. Las líneas 4 y 6 representan el conocimiento del número total de pistas en el *multitrack*, así como el dominio de decibeles (dBFS). La línea 8 es la responsable de asignar para cada *track* T, un valor DB en decibeles y asignarlo a *track_decibel*. En este momento, cada pista ya cuenta con un valor definido de decibeles para escalar su amplitud. La línea 10 es una restricción que indica que no puede haber una pista con una amplitud mayor que la denominada *lead_track*.

La línea 12 por medio de una función externa, convierte cada valor de decibel en amplitud lineal. Finalmente, se busca que la mezcla cuente con los valores de amplitud más elevados posibles, es por ello que la amplitud lineal de cada pista deberá estar entre -3 y 0 dBFS representados por 0.7 y 1.0 en amplitud lineal. Dado las restricciones del lenguaje, se manejan números enteros.

³⁸ M. Gebser et al., «Answer Set Solving in Practice», *Synthesis Lectures on Artificial Intelligence and Machine Learning*, vol. 6, n.º 3 (2012): 1-238.

```
1 |#const tracks = 2.
2 |#const lead_track_id = 1.
4 |track(1..tracks).
6 |decibel(-40..0).
8 |1 { track_decibel(T,DB) :
decibel(DB) } 1
:- track(T).
10|:- track_decibel(T,DB1),
T!= lead_track_id,
track_decibel(lead_track_id,DB
2), DB1>DB2.
12|track_amplitude(T,
@db2linear(DB)) :-  
track_decibel(T,DB).
14|:- 100 < #sum{A,T :
track_amplitude(T,A)}.
15|:- 70 > #sum{A,T :
track_amplitude(T,A)}.
```



Vol. 3 N.º 9 / Guayaquil
mayo 2024
ISSN 2953-657x

Estudio comparativo de enfoques probabilísticos para la generación de progresiones armónicas

Juan Camilo Núñez Jaramillo
Estudiante en Facultad de Ingenierías,
Universidad San Buenaventura
Medellín, Colombia
juan.nunez191@tau.usbmed.edu.co

Cristian Bajaña Díaz
Estudiante en Facultad de Ingenierías,
Universidad San Buenaventura
Medellín, Colombia
cristian.bajana191@tau.usbmed.edu.co

Mateo Yepes Díaz
Docente en Facultad de Ingenierías,
Universidad San Buenaventura
Medellín, Colombia
mateo.yepes@usbmed.edu.co

Daniel Gómez Marín
Music Technology Group, Universitat Pompeu Fabra, Barcelona, España
Facultad de Ingeniería, Universidad Icesi
Cali, Colombia
daniel.gomez@upf.edu
dgomez@icesi.edu.co

Resumen

En este artículo se presenta la comparativa de dos métodos para la armonización de melodías a través de modelos generativos de progresiones de acordes. La motivación de esta investigación surge de los avances logrados en algoritmos anteriores, específicamente los basados en modelos ocultos de Markov (HMM, por sus siglas en inglés), que han demostrado éxito en este campo y que han obtenido resultados prometedores al combinarse con modelos estocásticos de duración variable, o modelos bayesianos basados en funciones sintácticas tonales.

Recientemente, se han realizado avances en algoritmos para extraer datos cuantitativos de la estructura tonal en progresiones armónicas mediante espacios de tensión tonal y sus perfiles correspondientes. En este artículo se emplean los perfiles de tensión tonal para entrenar un HMM, independiente de las características melódicas, y, además, se entrena un segundo modelo utilizando las características tonales de la melodía. La validación de estos métodos se realiza a través de métricas objetivas y subjetivas, basadas en la comprensión del dominio musical.

El objetivo de este estudio es ampliar la exploración en el campo del aprendizaje automático musical, introduciendo nuevos métodos que permitan a músicos y productores interactuar con herramientas digitales y crear de manera más diversa composiciones musicales.

Palabras clave: Melodía, progresión de acordes, perfiles de tensión tonal, modelo oculto de markov

Abstract

In this paper, the comparison of two methods for harmonizing melodies through generative models of chord progressions is presented. The motivation for this research originates from the advances achieved in previous algorithms, specifically those based on Hidden Markov Models (HMM), which have demonstrated success in this field and have shown promising results when combined with stochastic models of variable duration or Bayesian models based on tonal syntactic functions.

Recently, there have been advances in algorithms for extracting quantitative data from the tonal structure in harmonic progressions through tonal tension spaces and their corresponding profiles. In this article, tonal tension profiles are used to train an HMM (Hidden Markov Model), independent of melodic features, and, in addition, a second model is trained using tonal characteristics of the melody at specific time points. The validation of these methods is carried out through objective and subjective metrics, based on the understanding of the musical domain.

The aim of this study is to broaden exploration in the field of musical machine learning, introducing new methods that enable musicians and producers to interact with digital tools and create musical compositions in a more diverse manner.

Keywords: Melody, chord progression, tonal tension profiles, hidden markov model

1. Introducción

Tradicionalmente, los procesos de aprendizaje y producción musical suelen ser lentos y requieren una gran inversión de recursos tanto para novatos como para expertos en música. Por ello, con el aprendizaje de máquina y sus aplicaciones en procesos de creación¹ se han desarrollado herramientas para la creación musical asistida.

La investigación en esta área ha permitido el desarrollo de diferentes sistemas de generación de elementos musicales, entre los que se encuentra la generación automática de acompañamientos melódicos y armónicos, mezcla y masterización. Estos sistemas son entrenados a partir de ejemplos musicales que los sistemas procesan, aprendiendo de ellos y conformando modelos estadísticos que son utilizados en diversos contextos y aplicaciones. Por ejemplo, el proyecto Flow Machines ha derivado en una herramienta de composición asistida denominada Flow Composer, con la cual se han asistido composiciones musicales y se han compuesto de álbumes como *Hello World*, por SKYGG.² También, investigaciones en ritmo han derivado en sistemas interactivos para el análisis y la generación de patrones de batería.³ Así mismo, técnicas derivadas del aprendizaje profundo o *deep learning*, como LSTM o *transformers*, han sido utilizadas en la creación sistemas generativos

de melodías musicales⁴, o incluso piezas completas.⁵

Las redes neuronales también han sido utilizadas para la armonización automática. Estas difieren de los modelos estadísticos clásicos en su capacidad de realizar predicciones más precisas, pero tienen la desventaja de que necesitan muchos más datos para lograr un buen rendimiento.

Un ejemplo de ello es el modelo bidireccional de redes de memoria a corto y largo plazo (LSTM por sus siglas en inglés) implementado por Lim, Rhyu y Lee⁶. Este modelo demostró un desempeño superior a un modelo oculto de Markov (HMM por sus siglas en inglés) que generaba progresiones armónicas a través de las notas de una melodía. Este desempeño se midió a partir de una métrica objetiva, que evaluaba la capacidad de predecir progresiones iguales al corpus de entrenamiento y una evaluación subjetiva con sujetos de prueba, que evaluaron la concordancia entre una melodía de una canción y la armonía generada de estos dos modelos.

Aun así, según Tsushima, Nakamura y Yoshii⁷, las redes neuronales tienen dificultades para representar la composición musical de manera similar a un humano debido a la dificultad de análisis e interpretabilidad de la estructura interna de las redes neuronales. En su investigación, se proponen métodos estadísticos de armonización que incorporan funciones sintácticas, estructuras de árboles y estructuras métricas de los acordes para generar progresiones.

Los HMM se han utilizado en áreas como el reconocimiento de habla,⁸ en fi-

1 Eduardo Reck Miranda, «Handbook of Artificial Intelligence for Music: Foundations, Advanced Approaches, and Developments for Creativity», en *Handbook of Artificial Intelligence for Music: Foundations, Advanced Approaches, and Developments for Creativity*, edición de Eduardo Reck Miranda (Springer International Publishing, 2021), <https://doi.org/10.1007/978-3-030-72116-9>

2 Shulei Ji, Jing Luo y Xinyu Yang, «A Comprehensive Survey on Deep Music Generation: Multi-level Representations, Algorithms, Evaluations, and Future Directions», *ArXiv* (2020), <http://arxiv.org/abs/2011.06801>

3 François Pachet, Pierre Roy y Benoit Carré, «Assisted Music Creation with Flow Machines: Towards New Categories of New», *ArXiv* (2020), <http://arxiv.org/abs/2006.09232>

4 Daniel Gómez-Marín, Sergi Jordà y Perfecto Herrera, «Drum Rhythm Spaces: From Polyphonic Similarity to Generative Maps», *J New Music Res* vol. 49, n.º 5 (2020): 438-456.

4 Kejun Zhang et al., «WuYun: Exploring Hierarchical Skeleton-guided Melody Generation Using Knowledge-enhanced Deep Learning», *ArXiv* (2023).

5 Andrea Agostinelli et al., «MusicLM: Generating Music From Text», *ArXiv* (2023), <http://arxiv.org/abs/2301.11325>

6 Hyungui Lim, Seungyeon Rhyu y Kyogu Lee, «Chord Generation From Symbolic Melody Using BLSTM Networks», *ArXiv* (2017).

7 Hiroaki Tsushima, Eita Nakamura y Kazuyoshi Yoshii, «Bayesian Melody Harmonization Based on a Tree-Structured Generative Model of Chord Sequences and Melodies», *IEEE/ACM Transactions on Audio, Speech, and Language Processing* (2020).

8 Lawrence R. Rabiner, «A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition», *Proceedings on the IEEE*, vol. 77, n.o 2 (1989).

nanzas y economía, para la predicción del precio de las acciones⁹ y también en el reconocimiento de patrones, que permiten identificar a diferentes hablantes basado en el timbre de sus voces.¹⁰ La generación automática de progresiones de acordes también ha sido objetivo de investigación para los HMM, utilizando métodos estadísticos para generar secuencias armónicas. Los HMM modelan procesos de Markov con dos variables de estado: los estados ocultos y los estados observables. Con estas variables un HMM utiliza la probabilidad para determinar la transición de un estado oculto hacia otro a través de un estado observable.¹¹

Aplicados al campo de la generación de acordes musicales, los HMM utilizan las notas de una melodía como estados observables y los acordes que las acompañan como estados ocultos, logrando generar una secuencia de acordes para una melodía introducida al modelo. Basados en esta metodología, se han desarrollado aplicaciones como MySong, la cual permite la armonización de una melodía vocal reconocida a través de detección de tono.¹² Otros HMM más complejos han sido diseñados utilizando modelos de grafos, a través de los cuales es posible capturar la estructura de los acordes y la interacción de estos basados en una melodía.¹³

Apesar de haber sido usados con éxito, una limitación de los HMM es que asumen que la probabilidad de cada estado oculto solo depende del estado anterior. Esto

restringe la capacidad de modelamiento de las distribuciones de probabilidad de los estados ocultos. En la generación musical esto hace que se asuma que un acorde solo está relacionado con una nota en una pieza musical sin tener en cuenta otras notas que puedan acompañar a ese acorde. Para abordar este desafío, se han implementado modelos semiocultos de Markov, los cuales tienen en cuenta las duraciones variables de los acordes. Estos modelos permiten modelar las características temporales de las secuencias entre notas y acordes.¹⁴

Los HMM mencionados anteriormente generan progresiones tomando variables musicales (por ejemplo, la armonía o la melodía de una pieza musical) como estados ocultos y como observaciones variables la secuencia de acordes, la melodía y/o la duración de los acordes. Aunque las dos últimas son las más relevantes para entrenar dichos modelos,¹⁵ es posible buscar otras variables que sirvan como estados para realizar armonización automática. Un ejemplo de esto¹⁶ combina modelos probabilísticos de diversas variables musicales utilizando métodos de interpolación logarítmica-lineal, a través de submodelos que representan las progresiones de acordes, la relación tonal y la relación melódica. También, Tsushima et al.¹⁷ utilizaron una combinación de diferentes modelos de Markov y gramática libre de contexto probabilístico (PCFG por sus siglas en inglés) para describir el ritmo, funciones armónicas y la estructura jerárquica de los acordes para generar progresiones. En esta investiga-

⁹ Xin An, «Prediction of Stock Price by Hidden Markov Model», *BCP Business & Management*, 38 (2023): 309–316.

¹⁰ Faisal Farhan, Ziaul Hoque y Provat Raihan, «Study of Pattern recognition system with Hidden Markov Model through the development of a raspberry pi based speaker recogniser» (tesis de grado, Ahsanullah University of Science & Technology, 2020), <https://www.researchgate.net/publication/348150123>

¹¹ Christopher M. Bishop, *Pattern Recognition and Machine Learning* (Springer, 2006).

¹² Ian Simon, Dan Morris y Sumit Basu, «MySong: Automatic Accompaniment Generation for Vocal Melodies», *CHI '08: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* Association for Computing Machinery (2008): 725–734.

¹³ Jean-François Paiement, Douglas Eck y Samy Bengio, «Probabilistic Melodic Harmonization», en *Advances in Artificial Intelligence. Canadian AI 2006. Lecture Notes in Computer Science*, edición de Lamontagne, L., Marchand, M., 218–229 (Berlín: Springer, 2006).

¹⁴ Ryan Groves, «Automatic Harmonization Using a Hidden Semi-Markov Model», *AIIDE Workshop Technical Report WS-13-22 (Musical Metacreation)*, vol. 9, n.º 5 (2013): 48–54.

¹⁵ Dan Morris, Ian Simon y Sumit Basu, «Exposing Parameters of a Trained Dynamic Model for Interactive Music Creation», *Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence*. (2008).

¹⁶ Stanislaw Raczyński, Satoru Fukayama, Emmanuel Vincent, «Melody Harmonisation with Interpolated Probabilistic Models» (reporte de investigación, RR-8110, INRIA, 2012), <https://hal.inria.fr/hal-00742957>

¹⁷ Hiroaki Tsushima, Eita Nakamura, Katsutoshi Itoya-ma, Kazuyoshi Yoshii, «Function-And Rhythm-Aware Melody Harmonization Based On Tree-Structured Parsing And Split-Merge Sampling Of Chord Sequences» (18th International Society for Music Information Retrieval Conference, Suzhou, China, 2017).

ción, se propone otra alternativa diferente basada en espacios tonales, en particular los perfiles de tensión tonal. Estos perfiles serán utilizados como los estados observables de un modelo oculto de Markov, es decir que este HMM genera progresiones armónicas basado en la información proporcionada por estos perfiles.

Los espacios tonales se han estudiado para describir de manera matemática y geométrica características que influyen en la armonía. Una de ellas es la tensión tonal, introducida primeramente por Lerdahl.¹⁸ A partir de esta introducción, se deriva la investigación de Randall y Khan¹⁹ que utilizaron esta teoría para formular una comparación con tres modelos empíricos para así comprobar si dicho modelo teórico se aproxima a los datos que modelan. La teoría desarrollada por Lerdahl también ha servido como base para implementar otros modelos, como el Espacio de Intervalo Tonal (TIS por sus siglas en inglés).²⁰ Este último modelo permite relacionar cada acorde en una progresión armónica específica con un valor numérico que describe un perfil de tensión tonal (PTT), permitiendo crear curvas de tensión tonal con dichos perfiles.

Los avances de Navarro y compañía en los TIS y PTT permiten cuantificar de manera sistemática y numérica aspectos relevantes de una pieza musical. Por un lado, los TIS dan la capacidad de encontrar la semejanza entre diferentes acordes y caracterizarlos, por medio de su respectiva posición dentro del espacio de intervalo tonal. Por el otro, los PTT ofrecen la facilidad matemática de cuantificar de una manera eficiente la tensión entre los acordes que

18 Fred Lerdahl, «Tonal Pitch Space», *Music Perception: An Interdisciplinary Journal*, vol. 5, n.º 3 (1988): 315-350, <https://www.researchgate.net/publication/257925384>

19 Richard R. Randall y Bilal Khan, «Lerdahl's Tonal Pitch Space Model and Associated Metric Spaces», *Journal of Mathematics and Music*, vol. 4, n.º 3 (2010): 121-131, <https://doi.org/10.1080/17459737.2010.529654>

20 Marcelo Caetano, María Navarro-Cáceres y Gilberto Bernardes, «Objective Evaluation of Tonal Fitness for Chord Progressions Using the Tonal Interval Space», *Artificial Intelligence in Music, Sound, Art and Design* (2020): 150-164.

componen una progresión alineada con la percepción del oyente.²¹

En la investigación presentada en este artículo se busca implementar un HMM a través de perfiles de tensión tonal y comparar dicho método con un HMM que crea progresiones a través de las notas de una melodía por medio de estudios objetivos y subjetivos que informen de su desempeño. Así se tendrá información de las ventajas y desventajas de ambos métodos a la hora de realizar la misma tarea. Esto con el fin de sentar las bases para un sistema de asistencia para la composición musical.

Para el entrenamiento de los modelos se hace uso de una base de datos de música electrónica EDM. Este género musical se elige por su gran popularidad en las últimas décadas y su influencia en personas de edades variadas, dado que el mismo ha sido tendencia desde los años 90. El estilo es denominado Electropop Fusion, y contiene diversas canciones de artistas que han estado en *rankings* de más escuchados de la última década.

El artículo está organizado de la siguiente manera: en la sección II se explica el preprocesamiento y entrenamiento de los modelos; en la sección III se explican los métodos de evaluación subjetiva y objetiva y se discuten los resultados obtenidos; luego, en la sección IV, se dan las conclusiones y el trabajo a futuro.

2. Metodología

Los algoritmos de los HMM diseñados en este trabajo buscan generar progresiones de acordes por medio de dos estados observables diferentes y un mismo estado oculto para los dos modelos. Uno de los modelos toma como estado observable los perfiles de tensión tonal (PTT) que describen las curvas de tensión encontradas en el corpus

21 María Navarro-Cáceres et al., «A Computational Model of Tonal Tension Profile of Chord Progressions in the Tonal Interval Space», *Entropy*, vol. 22, n.º 11 (2020): 1-30, <https://doi.org/10.3390/e2211291>

de la base de datos. Este modelo se denomina como PTT. El segundo modelo propone estados observables basados en las notas de la melodía de las canciones presentes en el corpus, y se le da el nombre de Melodía. Este modelo tomará una secuencia de notas y encontrará la secuencia de acordes más probable según cada nota.

La metodología propuesta para desarrollar este algoritmo se puede dividir en tres partes. Primero, se realiza un preprocesamiento de la base de datos para extraer información armónica y melódica (progresiones de acordes, notas, tonalidad y duraciones de nota y acordes). Segundo, se generan las matrices de transición, que representan la probabilidad de transición de un estado oculto a otro. Las matrices de emisión, que representan la probabilidad de tener un estado oculto dado un estado observable. Y de estado inicial, que representan la probabilidad inicial de un estado oculto. La figura 1 representa el diagrama de flujo para el entrenamiento de los dos modelos anteriormente mencionados. Finalmente, se procesan los datos del corpus musical con el propósito de entrenar cada modelo para generar progresiones armónicas y luego pasar a su debida evaluación subjetiva y objetiva.

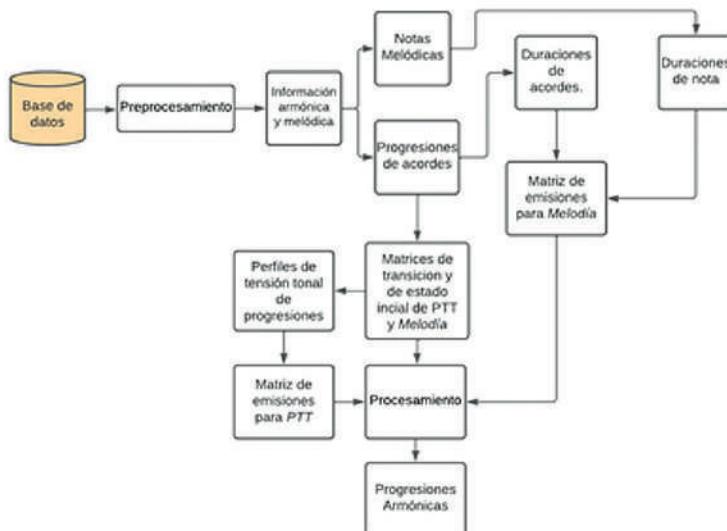


Figura 1: diagrama de flujo para el entrenamiento de los modelos.

2.1. Preprocesamiento

El corpus utilizado está compuesto por 83 fragmentos de cuatro compases extraídos de 25 canciones publicadas por artistas del género musical denominado EDM (entre ellos se encuentran artistas como Avicii, Calvin Harris, David Guetta, Dua Lipa, Marshmello). Estos fragmentos son elegidos porque cuentan con información tanto melódica como armónica en pistas MIDI independientes y por su reconocimiento en la última década. Los archivos MIDI fueron descargados de una base de datos pública disponible en línea.²²

De cada fragmento se extraen las siguientes características simbólicas musicales que representan la información extraída de armonía y melodía:

Para armonía son:

1. Progresión de acordes: representa la secuencia de acordes encontrada en cada fragmento.
2. Duración de acorde: indica el tiempo de duración de un acorde por medio de semicorcheas.
3. Perfil de tensión tonal: representan los cambios de tensión entre una progresión de acordes.
4. Tonalidad: indica la tónica o nota principal de cada fragmento del corpus musical.
5. BPM: indica la cantidad de pulsaciones por minuto de cada fragmento del corpus musical

Para la melodía son:

1. Tiempo inicial: indica el inicio de una nota melódica en un fragmento en semicorcheas.

22 Unison, <https://unison.audio/free-packs/>

2. Duración de cada nota: es la duración total de cada nota de un fragmento y está dada en semicorcheas.
3. MIDI pitch: indica el valor MIDI de cada nota de forma numérica a través de un número entero del 0 al 127.
4. Conjunto de tono de clase: es una representación de cada nota por medio de números enteros del 0 al 11 (cada uno representa una nota musical desde C a B).

Cinco algoritmos fueron utilizados para extraer la información armónica y melódica. Primero, para extraer la información de las progresiones de acordes de las canciones del corpus se implementó un algoritmo para reconocer la progresión de cada fragmento basado en la teoría propuesta por Pardo y Birmingham²³ y complementada por Scholz y Ramalho²⁴. El algoritmo divide cada fragmento en puntos de partición que marcan el inicio o final de una nota. Posteriormente, compara cada segmento con todas las posibles combinaciones de notas fundamentales y formas de acordes para identificar los acordes, las formas que compara el algoritmo son presentadas en la tabla I. Por último, utiliza un Grafo Acíclico Dirigido (DAG) para seleccionar los puntos de inicio y fin de los acordes. Se elige dicho algoritmo porque en una investigación actual que se encuentra en proceso este algoritmo obtuvo una eficiencia de reconocimiento del 79.77 %.

Tabla I
Forma y plantilla de acorde

Forma de acorde	Plantilla
M	[0, 4, 7, 0]
m	[0, 3, 7, 0]
o	[0, 3, 6, 0]
Aug	[0, 4, 8, 0]
Aug 7	[0, 4, 8, 10]
Aug Maj7	[0, 4, 8, 11]
7	[0, 4, 7, 10]
m7	[0, 3, 7, 10]
Maj 7	[0, 4, 7, 11]
m Maj7	[0, 3, 7, 11]
m7b5	[0, 3, 6, 10]
º7	[0, 3, 6, 9]
sus4	[0, 5, 7, 0]

Para extraer la información de la tonalidad de cada canción se utilizó un algoritmo basado en la investigación hecha por Madsen y Widmer²⁵. Se hace una normalización hacia C mayor o menor según corresponda el modo en el que se encuentra la canción, esto con el fin de normalizar todas las progresiones obtenidas y también las notas melódicas. Seguido a esto se le asigna a cada acorde su grado correspondiente en la tonalidad, por ejemplo, el acorde C M se clasifica como un primer grado I M.

Las progresiones normalizadas son utilizadas para obtener los perfiles de tensión tonal de los acordes. Los PTT representan los cambios de tensión que un oyente experimenta en una progresión de acordes, y se obtienen a través del cálculo de la tensión tonal instantánea en un intervalo de espacio tonal (TIS por sus siglas en inglés) para cada acorde, esta representa una combinación lineal de indicadores tonales proporcionados por Caetano, Navarro-Cáceres y Bernardes²⁶ descritos por la ecuación (1):

$$M(T_i, P) = \omega_1 \delta(T_i, P) + \omega_2 c(T_i) + \omega_3 M(T_i, P) + \omega_4 h(T_i, P) \quad (1)$$

²³ Bryan Pardo y William P. Birmingham, «Algorithms for Chordal Analysis», *Computer Music Journal*, vol. 26, n.º 2 (2002).

²⁴ Geber Ramalho y Ricardo Scholz, «COCHONUT: Recognizing Complex Chords from MIDI Guitar Sequences» (Conference: International Conferences on Music Information Retrieval, 2008), <https://www.researchgate.net/publication/269038716>

²⁵ Søren Tjagvad Madsen y Gerhard Widmer, «Key-Finding With Interval Profiles» (Proceedings of the International Computer Music Conference ICMC, Copenhague, Dinamarca, 2007).

²⁶ Caetano, Navarro-Cáceres y Bernardes, «Objective Evaluation of Tonal...».

Donde T_i representa el acorde i -ésimo de una progresión P , δ representa la distancia tonal, c disonancia, m conducción de voces, h tensión jerárquica y ω_j representa el peso de cada indicador tonal.

δ Representa una combinación lineal de tres descriptores de distancias tonales obtenidos a partir de un TIS que toma vectores croma de doce dimensiones y los convierte en un vector de intervalos tonales (TIV por sus siglas en inglés). Cada TIV puede representar un acorde, una tonalidad (C mayor o menor) y las tres funciones armónicas categóricas (tónica, subdominante y dominante). Con estos se calculan los indicadores de tensión de distancia entre acordes, distancia de cada acorde hacia la tonalidad y distancia de cada acorde hacia las funciones tonales.

Cada indicador de tensión tonal presenta un valor que se le asignará a un acorde en una progresión dada y de esta forma se obtienen las curvas de tensión necesarias para entrenar el modelo PTT.

Para extraer la información de duraciones de los acordes reconocidos y de las notas melódicas de cada canción se hace uso de la librería de Python Mido²⁷, donde se consigue información del BPM de las canciones por medio de los archivos MIDI. Para conocer la duración de cada acorde, se hace uso de los puntos de partición y el DAG obtenidos con el algoritmo de reconocimiento de acordes de Pardo y Birmingham²⁸ porque estos establecen el inicio y el final de cada acorde aproximado a la semicorchea más cercana.

Para la obtención de la duración de cada nota melódica, es necesario realizar la consecución del momento de inicio y final de esta. Sigue un procedimiento consistente para ambos datos: se selecciona la pista correspondiente a la melodía y, posterior-

mente, se recorre dicha pista en búsqueda de los mensajes Note-On y Note-Off del archivo MIDI. Estos mensajes representan los puntos temporales en los que inicia y termina una nota musical en el archivo. La librería Mido proporciona el tiempo inicial de cada nota con relación al último mensaje MIDI encontrado, por lo tanto, para calcular el tiempo inicial acumulativo de cada nota, se realiza una suma de los valores de tiempo de los mensajes Note-On y Note-Off encontrados por separado. Es importante tener en cuenta que los tiempos obtenidos inicialmente están dados en ticks MIDI, por lo que se debe realizar la respectiva conversión a semicorcheas. Esto se logra dividiendo el número de ticks por el equivalente en ticks a un dieciseisavo. Dada la elección de la unidad temporal elegida en la investigación, con los datos de inicio y final obtenidos, se calcula la diferencia entre el tiempo final y el tiempo inicial para determinar la duración de cada nota melódica.

Con la librería anteriormente descrita también se obtiene el descriptor de MIDI pitch de cada nota, el cual se pasa al descriptor de conjunto de tono de clase (PCS, por sus siglas en inglés) que representa la nota de la escala musical mediante un número de 0 a 11 donde 0 representa la nota C y 11 la nota B.

2.2. Modelos Ocultos de Markov

Diferentes autores²⁹ definen los HMM a través de las siguientes asunciones:

1. Una cadena de Markov finita de tiempo discreto X , que representa los estados ocultos.
2. Un proceso observable Y cuyas transiciones dependen de los estados X .

²⁷ Mido, <https://mido.readthedocs.io/en/stable/>
²⁸ Pardo y Birmingham, «Algorithms for Chordal Analysis».

²⁹ Jonh van der Hoek y Robert J. Elliot, *Introduction to Hidden Semi-Markov Models* (Cambridge University Press, 2018).

Lawrence R. Rabiner, «A Tutorial on Hidden...».

3. Una distribución de probabilidad de estados de transición, que representa la matriz de transición, $A = \{a_{ij}\}$ donde:

$$a_{ij} = P(X_{n+1} = e_j | X_n = e_i) \quad (2)$$

4. Una distribución de probabilidad inicial de estados que representa la matriz de estados iniciales $\pi = \{\pi_i\}$ donde:

$$\pi_i \quad P(X_0 = e_i) \quad (3)$$

5. Una distribución de probabilidad de observaciones Y_n en un estado X_n que representa la matriz de emisiones C_{ki}

$$C_{ki} = P(Y_n = f_k | X_n = e_i) \quad (4)$$

Con los descriptores de armonía y melodía ya obtenidos, se procede a generar las matrices de estado inicial, de transición y de emisión para los modelos a entrenar, el primer modelo es un HMM que utiliza como estado observable las curvas de tensión tonal generadas por los PTT de todas las progresiones de acordes existentes en la base de datos y el segundo es un modelo oculto convencional que utiliza como estado oculto las notas de una melodía.

Las curvas de PTT se obtienen a partir de cada progresión presente en el corpus y se toman como observaciones para el modelo PTT; el esquema de la figura 2 muestra la forma gráfica del HMM basado en PTT que es agnóstico a la melodía. Por otro lado, en el modelo Melodía las observaciones se obtienen a partir de las notas melódicas que se asocian con un acorde en un fragmento musical; el esquema de la figura 3 muestra la forma gráfica del HMM basado en Melodía.

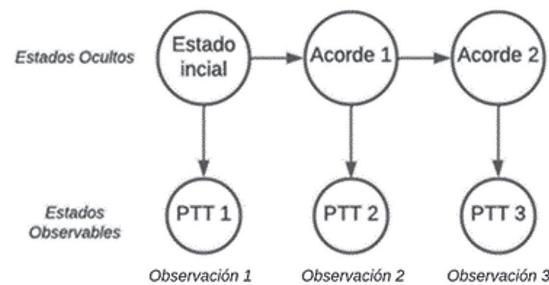


Figura 2: representación gráfica HMM a partir de perfiles de tensión tonal.



Figura 3: representación gráfica HMM a partir de notas de una melodía.

2.3. Procesamiento de datos del corpus musical

Una vez definidos los dos modelos ocultos de Markov, se procesa la información musical de la base de datos para obtener cada matriz descrita en la sección anterior.

Las progresiones obtenidas primero se dividen en tonalidad mayor y menor, ya que cada una cuenta con grados diferentes que pueden aparecer en una progresión; estos son representados por la tabla II. Luego de estar organizadas según su tonalidad se procede a obtener la matriz de estado inicial, con los grados iniciales en una progresión. De esta forma se obtiene la probabilidad de que una progresión inicie con un determinado grado inicial.

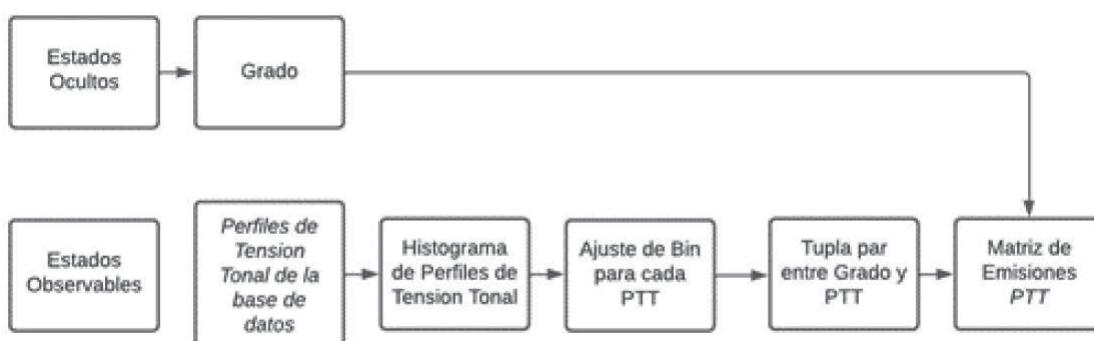
Tabla II
Grados según tonalidad

Tonalidad Mayor	Tonalidad Menor
I	I
I#/IIb	I#/IIb
II	II
II#/IIIb	III
III	III#/IVb
IV	IV
IV#/Vb	IV#/Vb
V	V
V#/VIb	VI
VI	VI#/VIIb
VI#/VIIb	VII
VII	VII#

Para generar la matriz de transición de acordes que representa la probabilidad de transición de un grado X_{i+1} dado otro grado X_i se toman todos los grados existentes en una progresión emparejándose de a dos, generando una tupla par que contiene un grado y su sucesor. Esto permite generar una cadena de Markov de orden uno y representar los estados ocultos X . Para determinar la probabilidad de transición se analiza un estado en el espacio, buscando en las tuplas que tengan en primera posición dicho estado y se realiza un conteo de

tonalidad de C menor). Dicha agrupación se realiza de manera empírica y corresponde a un ajuste del ancho de cada *bin* perteneciente al histograma generado. El objetivo de este ajuste es lograr que la cantidad de *bins* en la distribución sea proporcional a la cantidad de grados existentes en el corpus. De esta manera, se reduce la cantidad de estados ocultos asociados a un grado específico, lo que las hace más significativas para el modelo.

Seguido a la normalización, para generar la matriz de emisiones se analizan los posibles PTT generados por un grado en particular, generando ahora una tupla par de grado y perfil de tensión tonal. Como se indicó anteriormente un grado puede tener diferentes perfiles de tensión tonal por ello la probabilidad de cada observación está dada por $P(y_m|x_m)$ esta es obtenida contando la cantidad de veces que un grado x_m se asocia con un PTT y_m y luego se realiza una normalización sumando el total de perfiles asociados a un estado X . La figura 4 representa el diagrama de flujo de este proceso.



cuántas veces ese acorde pasa a otro en toda la base de datos, luego de dicho conteo se normaliza sumando la cantidad de posibles cambios para ese estado.

Antes de realizar la matriz de emisiones para el primer modelo, es relevante realizar una agrupación en intervalos de datos (también llamados *bin*), basados en los valores de PTT encontrados en el corpus (166 para tonalidad de C mayor y de 172 para

Figura 4: diagrama de flujo de obtención de la matriz de emisiones de PTT.

En contraste con las observaciones del modelo PTT, las de Melodía se derivan directamente de la melodía a la que se desea aplicar la armonización correspondiente. Es importante asegurar que la cantidad de acordes generados sea adecuada, por esto, tal como en el primer modelo, la longi-

tud del conjunto de observaciones debe ser igual a la longitud del conjunto de duraciones de acordes. Es común que una melodía contenga considerablemente más notas que acordes que la acompañen. Por lo tanto, al definir el conjunto de observaciones de notas, es necesario recortar las notas reconocidas originalmente del archivo MIDI ingresado. Para realizar este proceso se utilizan los descriptores de melodía mencionados en la etapa de preprocesamiento. Usando los puntos temporales de inicio y duración de las notas de la melodía, se establece una correspondencia temporal con la cantidad de duraciones de los acordes y los valores respectivos de esas duraciones. De esta manera, cada observación de la melodía se asocia con la primera nota reconocida que se encuentre dentro de la duración y el tiempo inicial de las duraciones de los acordes.

Con dichas notas se realiza un proceso similar al primer modelo para generar la matriz de emisiones del segundo modelo, como se observa en la figura 5; ahora la tupla par simboliza un grado y la nota con la que tiene correspondencia. Por ende, la probabilidad de cada observación se obtiene contando la cantidad de veces que un grado x_m se asocia con una nota y_m . Para encontrar la secuencia de estados más probable dado una cadena de observaciones de cada modelo, se utiliza el algoritmo de Viterbi.³⁰

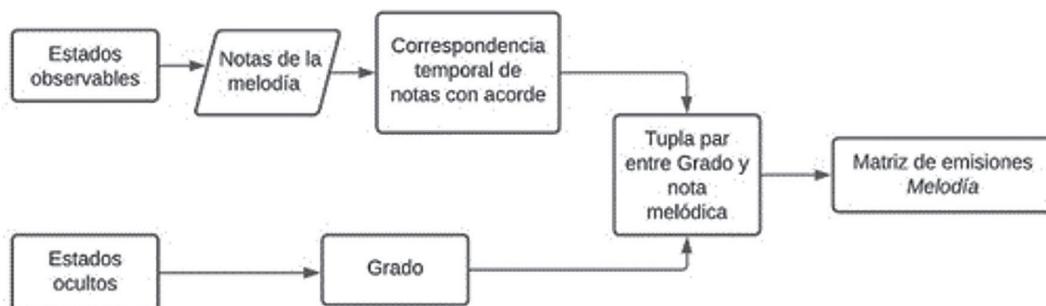


Figura 5: diagrama de flujo de obtención de la matriz de emisiones de Melodía.

Finalmente, para evaluar la eficiencia de los modelos creados se divide el corpus y se reentrenan los modelos, quedando así un 75 % de datos para el entrenamiento y un 25 % restante para la evaluación. Esta división es representada por el diagrama de flujo de la figura 6.

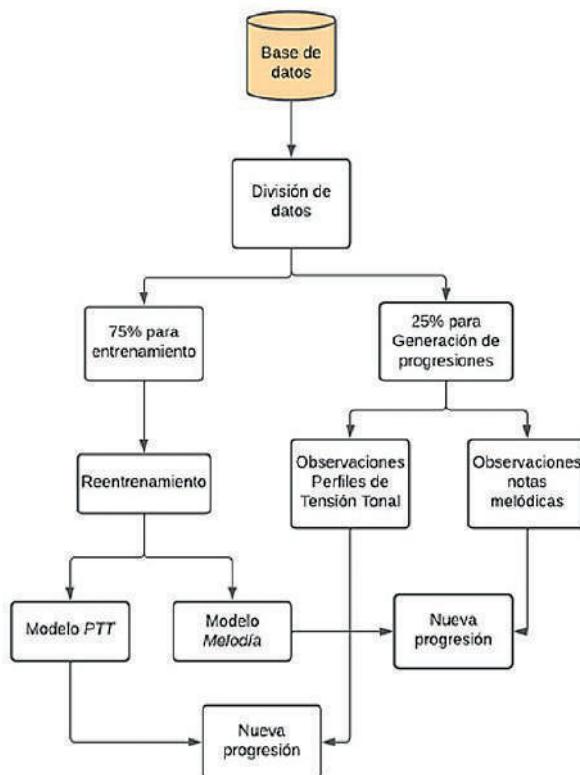


Figura 6: diagrama de flujo reentrenamiento de modelo PTT y Melodía.

³⁰ G. David Forney, «The Viterbi Algorithm», *Formal Structure of Electromagnetics*. Amsterdam vol. 61, n.o 3 (1973).

3. Resultados

Para la evaluación de la eficiencia de los modelos PTT y Melodía se realizan tres evaluaciones objetivas y una evaluación subjetiva con un grupo de personas. Para estas pruebas se toma aleatoriamente una progresión que pertenezca al 25 % de datos que sobran luego del reentrenamiento de los modelos que se indicó en la sección anterior. Seguido a esto se genera una progresión nueva con PTT, tomando la curva de tensión de la progresión elegida. Luego se crea una progresión nueva con Melodía, tomando las notas que acompañan la progresión elegida. Y por último se crea una progresión aleatoria con acordes pertenecientes al estilo musical, es decir que solo se toman en cuenta acordes que se reconocieron en el corpus musical. Este proceso se hace hasta obtener 21 progresiones para cada uno de los grupos que serán denominados como «Original» (grupo de verdad), «Modelo PTT», «Modelo Melodía» y «Aleatorio».

3.1. Experimento objetivo 1: evaluación de función de densidad de probabilidad

Con este primer experimento se busca comparar las generaciones obtenidas con las progresiones originales de la base de datos. Esto a través de una Estimación de Densidad de Kernel (EDK)³¹ y la Divergencia de Kullback-Leibler (DKL)³² y el Área de Superposición (AS). Una EDK permite estimar la función de densidad de probabilidad (FDP) de un conjunto de datos, de esta forma se obtiene una representación numérica de cómo se distribuyen los acordes de las

³¹ Murray Rosenblatt, «Remarks On Some Nonparametric Estimates Of A Density Function», *Annals of Mathematical Statistics* (1956).

³² I. Csiszar, « \mathcal{I} -Divergence Geometry Of Probability Distributions And Minimization Problems», *The Annals of Probability*, vol. 3, n.º 1 (1975): 146–158.

progresiones de cada grupo. Esto permite que se pueda comparar la FDP de un grupo respecto a la FDP del grupo «Original» que representa el grupo de verdad a través del cálculo de DKL y AS.

3.1.1. Procedimiento

Se crean entonces nuevas curvas de tensión con las progresiones generadas por los dos modelos y un tercer método denominado «Aleatorio». Para generar el EDK de cada grupo se calcula la distancia de Frechet³³ entre las curvas del grupo de verdad frente a cualquier otro grupo, es decir, se comparan las distancias del grupo «Original» con las del grupo «Modelo PTT».

Siguiendo la metodología de Yang y Lerch³⁴, se genera una EDK para obtener una FDP con cada grupo. De esta forma para hallar la similitud entre una FDP y la FDP del grupo de verdad se realiza la medida de DKL y AS que permite determinar qué tanto diverge cada distribución frente al grupo «Original».

3.1.2. Resultados y discusión

La figura 7 muestra las distribuciones obtenidas de cada grupo y en ella se puede observar qué tan similares son las progresiones generadas con respecto al grupo de verdad, en términos de sus perfiles de tensión tonal. Se puede ver que existe una mayor semejanza para los grupos «Modelo PTT» y «Aleatorio» con la FDP del grupo «Original» que representa el estilo musical del corpus. Algo que detallar en la similitud obtenida con el grupo anterior es que esta

³³ Helmut Alt y Michael Godau, «Computing The Fréchet Distance Between Two Polygonal Curves», *International Journal of Computational Geometry & Applications*, vol. 5, n.º 1 y 2 (1995): 75–91. <https://doi.org/10.1142/s0218195995000064>

³⁴ Li-Chia Yang y Alexander Lerch, «On the Evaluation of Generative Models in Music», *Neural Computing and Applications*, vol. 32, n.º 9 (2020): 4773–4784. <https://doi.org/10.1007/s00521-018-3849-7>

se da porque las progresiones contenidas en él se generan teniendo en cuenta los acordes que representan a toda la base de datos. Por otro lado, el modelo Melodía es el que menor influencia tiene del grupo de verdad en los PTT generados por este.

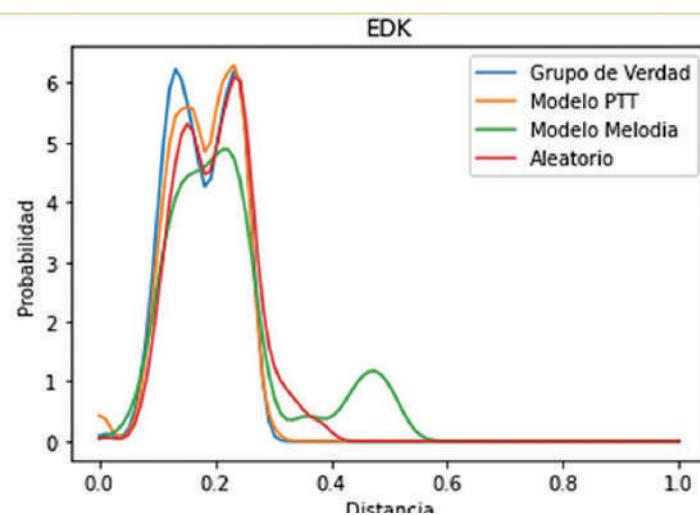


Figura 7: estimación de Densidad de Kernel (EDK) de cada grupo evaluado.

Tabla III
Divergencia de Kullback-Leibler (DLK)
y área de superposición

	DLK	AO
«Original» y «Modelo PTT»	1.578	0.95
«Original» y «Modelo Melodía»	19.698	0.80
«Original» y «Aleatorio»	9.903	0.89

Por otro lado, en Melodía se observa que este presenta la mayor divergencia y una menor área de superposición, indicando así que el modelo no es capaz de recrear de manera adecuada los PTT de las progresiones originales, incluso las progresiones generadas aleatoriamente lo superan. Un factor que afecta este resultado son las observaciones de este modelo porque se asume que un acorde solo puede estar ligado a una nota en una duración determinada, sin tener en cuenta que otras notas pueden acompañarlo.

3.2. Experimento objetivo 2: evaluación de descriptor de transición de acordes

Con el objetivo de evaluar la similitud en las transiciones de acordes generadas por cada grupo experimental en comparación con el grupo de referencia, se utiliza una métrica de evaluación que genera una matriz de transición entre acordes para cada uno de ellos. De esta forma se puede determinar cómo cada uno de los modelos define los patrones armónicos del estilo musical a través de las progresiones evaluadas.

3.2.1. Procedimiento

En primer lugar, se hace un proceso similar al señalado en la sección II donde se describe la obtención de la matriz de transición que representa una cadena de Markov de primer orden. Esto se aplica para esta prueba dividiendo las canciones en sus respectivas tonalidades (mayor y menor), teniendo en cuenta todos los acordes existentes en cada una de estas; estos acordes fueron descritos previamente en la tabla II. Luego se agrupan los acordes en tuplas pares para cada grupo y se analizan las transiciones de acordes existentes en cada progresión y finalmente se normalizan las matrices. Después de obtenerlas, se procede a realizar un cálculo de distancia de cada matriz respecto al grupo de verdad dependiendo la tonalidad, a partir de una norma matricial descrita por la ecuación (5).

$$d_2(A, B) = \sum_{i=1}^n \sum_{j=1}^n (a_{ij} - b_{ij})^2 \quad (5)$$

3.2.2. Resultados y discusión

La tabla IV representa las distancias descriptas por la ecuación 5 obtenidas para las progresiones con tonalidad mayor y menor respectivamente.

Tabla IV
Norma matricial d2

	Mayor	Menor
«Original» y «Modelo PTT»	0.636	0.0
«Original» y «Modelo Melodía»	2.741	3.647
«Original» y «Aleatorio»	3.449	4.573

En los resultados de la tabla IV se observa que el modelo PTT tiene un grado alto de similitud con los patrones armónicos que se encuentran en el estilo musical del grupo «Original». Además de esto, se nota que, para las progresiones evaluadas en tonalidad menor, este modelo genera patrones armónicos idénticos, porque la distancia obtenida es 0, y para la mayor genera patrones muy cercanos. Los resultados del segundo modelo Melodía indican que este tiene una menor capacidad para imitar los patrones armónicos, en concreto los patrones de la tonalidad menor con respecto a la mayor. Por último, las distancias obtenidas por el grupo «Aleatorio» son las más largas; esto es de esperar debido a que sus progresiones contienen patrones armónicos elegidos al azar.

3.3. Experimento objetivo 3: evaluación de similitud

Se establece una métrica de similitud de progresiones a partir de la similitud entre acordes³⁵ para evaluar los acordes de cada

³⁵ C. Wongsaroj, N. Prompoon y A. Surarerks, *Digital Information and Communication Technology and its Applications (DICTAP)*, 2014 Fourth International Conference (2014).

modelo con las progresiones evaluadas y determinar qué tan similar estas están frente al estilo musical elegido.

3.3.1. Procedimiento

La primera métrica es calculada el índice de Jaccard a través de la ecuación (6) por medio de cromas, en este caso entre dos cromas que representan los acordes a comparar. La similitud entre acordes representa una relación entre las notas compartidas entre ellos dividido por las notas de cada uno de ellos, en otras palabras, la intersección de sus tonos de clase frente a la unión de sus tonos de clase.

$$C_sim(C_1, C_2) = \frac{|P(C_1) \cap P(C_2)|}{|P(C_1) \cup P(C_2)|} \quad (6)$$

La segunda métrica deriva de la primera y se encarga de comparar la similitud entre acordes de dos progresiones diferentes representadas por vectores $a[i]$ y $b[j]$ cada uno con una longitud determinada. En este caso las longitudes de acordes se tomarán como iguales, lo que deriva a computar la suma de similitud entre cada acorde $a[j]$ con un acorde $b[j]$ y normalizar por la cantidad de acordes en cada progresión, la ecuación (7) describe este proceso.

$$C_sim(C_1, C_2) = \frac{\sum_{j=1}^n C_sim(a[j + i - 1], b[j])}{m} \quad (7)$$

3.3.2. Resultados y discusión

La figura 8 representa los resultados obtenidos de similitud de acordes de los grupos estudiados frente al grupo «Original» y la tabla V representa los valores obtenidos de promedio y desviación estándar.

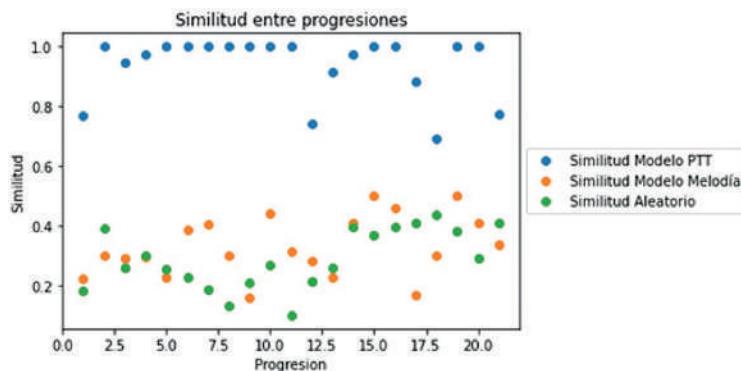


Figura 8: gráfica de similitud entre progresiones frente al grupo de verdad.

Tabla V
Resultados estadísticos

	Media	Desviación estándar
«Original» y «Modelo PTT»	93.6%	0.100
«Original» y «Modelo Melodía»	33%	0.099
«Original» y «Aleatorio»	28.9%	0.0976

Los resultados indican que el modelo PTT genera progresiones muy similares al corpus; esto significa que las secuencias armónicas generadas por este método tienen la capacidad de reproducir el estilo musical que representa el corpus. Por otro lado, el segundo modelo Melodía tiende a generar progresiones diferentes a las originales, lo que indica que su capacidad de imitación es inferior al primer modelo y que sus observaciones no permiten predecir el estilo musical subyacente en el corpus.

En general, el grupo «Modelo PTT» obtiene un promedio de similitud frente al grupo original del 93.6 %, superando así el promedio de «Modelo Melodía» de 33 % y el de «Aleatorio» de 28.9 %. Esto refuerza lo anteriormente dicho e indica que en promedio el modelo PTT captura en mejor medida las características y patrones de las progresiones de acordes de la base de datos original.

3.4. Experimento 4: prueba subjetiva

El experimento de prueba subjetiva se realiza con la finalidad de evaluar y comparar el rendimiento de los dos modelos HMM implementados, que mide qué tan bien se relacionan las armonías generadas con la melodía a la cual se le está aplicando la armonización. Además, se busca explorar la capacidad de los modelos desde otra dimensión teniendo en cuenta la subjetividad, factor importante que se debe tener presente en un arte como la música.

3.4.1. Preprocesamiento

Para realizar una evaluación del desempeño de los métodos implementados se eligen diez progresiones armónicas de las veintiuna generadas que se menciona al inicio de esta misma sección, obteniendo así cuatro muestras por canción correspondientes a los grupos de evaluación mencionados al inicio de esta sección (Original, PTT, Melodía y Aleatorio).

Se generan los archivos MIDI de cada progresión y se procede a formar los diez grupos de muestras, generando para cada canción escogida cuatro archivos de audio denominados «Fragmentos musicales» con melodía y armonía. Se usa un tempo de 120 BPM y todas las progresiones se establecen en la octava número cuatro del piano.

3.4.2. Procedimiento

En primera instancia se introduce el experimento y se explican las dos métricas correspondientes a la evaluación. Posterior a ello se realiza una sección de entrenamiento para que el sujeto comprenda el contexto

to y se familiarice con la prueba. Se busca poner dos fragmentos musicales que sean completamente opuestos en percepción de desempeño, es decir, uno correcto y otro que no lo es, teniendo en cuenta progresiones armónicas de la música popular occidental relacionadas al corpus utilizado en la investigación.³⁶

En cada iteración del experimento se dan al usuario cuatro fragmentos musicales para escuchar, todos ellos con la misma melodía y diferente armonía, cada armonía varía según los grupos de evaluación Original, PTT, Melodía y Aleatorio en la subsección 1 de este experimento. Después de escuchar cada «Fragmento musical», el sujeto procede a calificarlos en una escala de Likert de 0 a 5. Allí se encuentra con dos métricas³⁷ que, con sus respectivas escalas, son las siguientes:

- Fluidez de la armonización, en términos de transición de acordes: donde 0 es «la armonización presenta transiciones bruscas y abruptas entre acordes, dificultando la cohesión musical» y 5 corresponde a «la armonización muestra transiciones suaves y fluidas entre acordes, creando una progresión fluida y agradable».
- Enriquecimiento y transmisión emocional de la armonía a la melodía: donde 0 es «la armonía no realza ni complementa la expresión emocional de la melodía» y 5 corresponde a «la armonía enriquece y amplifica la expresión emocional de la melodía de manera notable».

La escala de calificación elegida tiene como objetivo evitar sesgos en el puntaje final

³⁶ Cuestionario de armonización de melodías: <https://form.jotform.com/231597775210056>

³⁷ Métricas que se basan en la investigación de Yin-Cheng Yeh et al., «Automatic Melody Harmonization with Triad Chords: A Comparative Study», *ArXiv* (2021).

de las métricas sin posibilidad de un punto medio en la calificación, permitiendo cuantificar el desempeño de forma tanto positiva como negativa. Se considera que un puntaje igual o mayor a 3 representa un desempeño positivo, mientras que aquellos inferiores a 3 se consideran negativos. Así, se plantean las siguientes hipótesis para el experimento:

- H1: se plantea que la métrica «Fluidez de la armonización» tenga una calificación promedio mayor o igual a 3 para el modelo PTT, lo que indica un desempeño positivo para dicho modelo.
- H2: la métrica «Enriquecimiento y transmisión emocional» tiene una calificación promedio mayor o igual a 3 para el modelo PTT, presentando un desempeño positivo para el modelo en ambas métricas evaluadas en el experimento.
- H3: la diferencia de medias entre el modelo Original y el modelo PTT es más baja que la diferencia de medias entre Original-Melodía y Original-Aleatorio.
- H4: con el objetivo de refutar la similitud observada en la figura 7, del modelo Aleatorio con el modelo Original, se plantea que la diferencia de medias sea la más alta para las dos métricas propuestas en la prueba.

3.4.3. Participantes

Para la consecución de los sujetos de prueba, se recopila para un total de 37 sujetos con experiencia académica específica en armonía. Divida en grupos de 1 a 3 años, 3 a 5 años y más de 5 años de experiencia.

3.4.4. Resultados y discusión

Se presentan en la tabla VI los estadísticos de media y desviación estándar para las

dos métricas evaluadas en el experimento. Además, se muestra de manera detallada la distribución de los datos en un diagrama de caja en las figuras 9 y 10 para las dos métricas evaluadas, respectivamente.

Tabla VI
Resultados estadísticos de los modelos

Métrica	Modelo	Media	Desviación estándar
Fluidez de la armonización	Original	3.72	1.12
	Melodía	2.92	1.28
	PTT	3.42	1.33
	Aleatorio	2.29	1.35
Enriquecimiento y transmisión emocional	Original	3.52	1.23
	Melodía	2.61	1.41
	PTT	3.18	1.42
	Aleatorio	1.89	1.36

Según se muestra en la tabla VI, las medias del puntaje del modelo PTT para las métricas «Fluidez de la armonización» y «Enriquecimiento y transmisión emocional» son 3.42 y 3.18, respectivamente. Estos resultados confirman las hipótesis H1 y H2, lo que indica un desempeño positivo del modelo PTT en cuanto a la generación de progresiones, tanto en la fluidez de la armonización como en el enriquecimiento y la transmisión emocional.

Es observable, además, e importante destacar que, aunque la media de puntaje del método Melodía no es mayor o igual a 3, se acerca por 0.08 al 3, lo que indica un desempeño que se acerca a un funcionamiento positivo, por lo que puede funcionar en casos donde se requiera acompañamientos armónicos menos técnicos. Esto sugiere que pueda ser valioso para la generación de progresiones armónicas menos predecibles o ceñidas a lo típico.

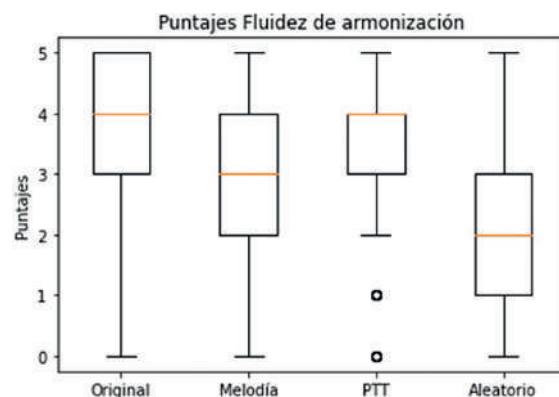


Figura 9: distribución de datos métrica «Fluidez de armonización».

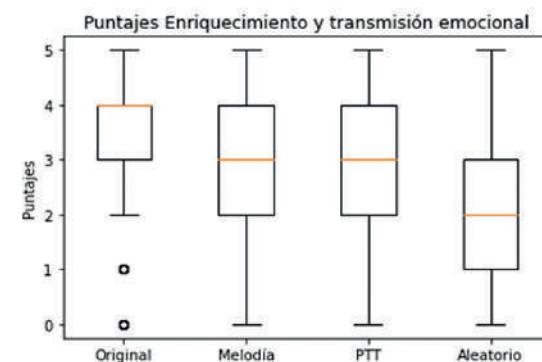


Figura 10: distribución de datos métrica «Enriquecimiento y transmisión emocional».

Se observa en la figura 9 que, aunque hay datos atípicos en el modelo PTT, en la métrica «Fluidez de armonización» comparte un intervalo en la distribución de la mayoría de los datos con el modelo Original siendo entre 3 y 5 para este, y entre 3 y 4 para el modelo PTT.

Tabla VII
Resultados de prueba anova métrica "fluidez de armonización"

Métrica	Valor P	Valor F	GL Num	GL Den
Fluidez de la armonización	1.14x10 ⁻⁵²	88.51	3	1476
Enriquecimiento y transmisión emocional	1.67x10 ⁻⁵⁹	101.06	3	1476

Se realiza un análisis de varianzas (ANOVA) para así determinar si hay diferencias significativas entre los grupos. Se observan los resultados del ANOVA en la tabla VII, con un valor de significancia $p < 0.05$ para ambas métricas, lo que indica que hay por lo menos un grupo con diferencias significativas a los demás en cuestiones de media esta-

dística. Con dicha información se procede a realizar una prueba de *post-hoc* de Tukey, que proporciona información más detallada acerca de las diferencias de medias entre grupos (resultados mostrados en la tabla VIII para la métrica «Fluidez de transición» y tabla IX para la métrica «Enriquecimiento y transmisión»).

Tabla VIII
Resultados de prueba *post-hoc tukey*
métrica «fluidez de armonización»

Grupo 1	Grupo 2	Diferencia de medias	Valor P
Aleatorio	Melodía	-0.3054	0.0066
Aleatorio	Original	-1.4324	0.0
Aleatorio	PTT	-0.8027	0.0
Melodía	Original	-1.127	0.0
Melodía	PTT	-0.4973	0.0
Original	PTT	0.6297	0.0

Tabla IX
Resultados de prueba *post-hoc tukey* para la
métrica «enriquecimiento y transmisión emocional»

Grupo 1	Grupo 2	Diferencia de medias	Valor P
Aleatorio	Melodía	-0.3378	0.0044
Aleatorio	Original	-1.6297	0.0
Aleatorio	PTT	-0.9081	0.0
Melodía	Original	-1.2919	0.0
Melodía	PTT	-0.5703	0.0
Original	PTT	0.7216	0.0

Se confirma la hipótesis H3 al observar los resultados en las tablas VIII y IX para ambas métricas evaluadas. La diferencia media entre el modelo Original y el modelo PTT es de 0.6297 en la primera métrica, un valor más bajo en comparación con los valores de 1.127 y 1.4324 obtenidos en las comparaciones entre el modelo Original y los modelos Melodía y Aleatorio, respectivamente. En la segunda métrica, se repite este patrón, con una diferencia media de 0.7216 entre el modelo Original y el modelo PTT, valor inferior a los 1.2919 y 1.6297 correspondientes a las comparaciones de medias entre el modelo Original y los otros dos modelos. Se observa, además, un valor *p* menor a 0.05 de las parejas de grupos, lo que significa una

diferencia estadística significativa entre ellos, lo cual no indica que haya un posible buen imitador del método Original con otro grupo, pero deriva en un análisis de la diferencia de medias para así determinar mayor o menor cercanía al grupo Original. Es importante destacar que los valores negativos de la diferencia medias observados en las tablas VIII y IX indican que la media del grupo 2 es mayor, mientras que un valor positivo indica que la media del grupo 1 es mayor. Sin embargo, la magnitud de este parámetro indica la diferencia de medias de igual manera.

Finalmente, se analiza las diferencias de medias del modelo Aleatorio con el Original y se observa que efectivamente muestra los valores más altos para ambas métricas. Estos valores son 1.4324 para la métrica «Fluidez de armonización» y 1.6297 para la métrica de «Enriquecimiento y transmisión emocional» por lo que se confirma la hipótesis H4 y se refuta una similitud del modelo Aleatorio con el modelo Original.

4. Conclusiones

En esta investigación se han desarrollado y comparado dos algoritmos de armonización de melodías basados en HMM. Uno de ellos utilizando notas presentes en la melodía y el otro utilizando perfiles de tensión tonal. A través de las pruebas objetivas y subjetivas con los grupos de evaluación «Original», «Aleatorio», «PTT» y «Melodía», se ha demostrado que el modelo PTT supera al modelo Melodía en la generación de patrones armónicos y progresiones similares al corpus con el que fueron entrenados.

Este descubrimiento representa un aporte en el campo de la generación automática de armonías musicales a través de HMM, especialmente al utilizar progresiones basadas en perfiles de tensión tonal

para un estilo musical particular.

Finalmente, se plantea la posibilidad de mejorar el modelo en el futuro mediante la eliminación de la dependencia de los puntos de inicio de los acordes y la reducción de los datos ingresados por el usuario.

5. Referencias

- Agostinelli, Andrea, Timo I. Denk, Zalán Boros, Jesse Engel, Mauro Verzetti, Antoine Caillon, Qingqing Huang, Aren Jansen, Adam Roberts, Marco Tagliasacchi, Matt Sharifi, Neil Zeghidour, Christian Frank. «MusicLM: Generating Music From Text». *ArXiv* (2023). <http://arxiv.org/abs/2301.11325>
- Alt, Helmut y Michael Godau. «Computing The Fréchet Distance Between Two Polygonal Curves». *International Journal of Computational Geometry & Applications*, vol. 5, n.º 1 y 2 (1995): 75–91. <https://doi.org/10.1142/S0218195995000064>
- An, Xin. «Prediction of Stock Price by Hidden Markov Model». *BCP Business & Management*, 38 (2023): 309–316.
- Bishop, Christopher M. *Pattern Recognition and Machine Learning*. Springer, 2006.
- Caetano, Marcelo, María Navarro-Cáceres y Gilberto Bernardes. «Objective Evaluation of Tonal Fitness for Chord Progressions Using the Tonal Interval Space». *Artificial Intelligence in Music, Sound, Art and Design* (2020): 150–164.
- Csiszar, I. «I-Divergence Geometry Of Probability Distributions And Minimization Problems». *The Annals of Probability*, vol. 3, n.º 1 (1975): 146–158.
- Farhan, Faisal, Ziaul Hoque y Provat Raihan. «Study of Pattern recognition system with Hidden Markov Model through the development of a raspberry pi based speaker recogniser». Tesis de grado, Ahsanullah University of Science & Technology, 2020. <https://www.researchgate.net/publication/348150123>
- Forney, G. David. «The Viterbi Algorithm».
- Formal Structure of Electromagnetics. Amsterdam vol. 61, n.º 3 (1973).
- Gómez-Marín, Daniel, Sergi Jordà y Perfecto Herrera. «Drum Rhythm Spaces: From Polyphonic Similarity to Generative Maps». *J New Music Res* vol. 49, n.º 5 (2020): 438–456.
- Groves, Ryan. «Automatic Harmonization Using a Hidden Semi-Markov Model». *AI-IDE Workshop Technical Report WS-13-22 (Musical Metacreation)*, vol. 9, n.º 5 (2013): 48–54.
- Ji, Shulei, Jing Luo y Xinyu Yang. «A Comprehensive Survey on Deep Music Generation: Multi-level Representations, Algorithms, Evaluations, and Future Directions». *ArXiv* (2020). <http://arxiv.org/abs/2011.06801>
- Lerdahl, Fred. «Tonal Pitch Space». *Music Perception: An Interdisciplinary Journal*, vol. 5, n.º 3 (1988): 315–350. <https://www.researchgate.net/publication/257925384>
- Lim, Hyungui, Seungyeon Rhyu y Kyogu Lee. «Chord Generation From Symbolic Melody Using BLSTM Networks». *ArXiv* (2017).
- Madsen, Søren Tjagvad y Gerhard Widmer. «Key-Finding With Interval Profiles». *Proceedings of the International Computer Music Conference (ICMC)*, Copenhague, Dinamarca, 2007.
- Miranda, Eduardo Reck. «Handbook of Artificial Intelligence for Music: Foundations, Advanced Approaches, and Developments for Creativity». En *Handbook of Artificial Intelligence for Music: Foundations, Advanced Approaches, and Developments for Creativity*. Edición de Eduardo Reck Miranda. Springer International Publishing, 2021. <https://doi.org/10.1007/978-3-030-72116-9>
- Morris, Dan, Ian Simon y Sumit Basu. «Exposing Parameters of a Trained Dynamic Model for Interactive Music Creation». *Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence* (2008).
- Navarro-Cáceres, María, Marcelo Caetano, Gilberto Bernardes, Mercedes Sánchez-Barba y Javier Merchán Sánchez-Jara. «A Computational Model of Tonal Tension Profile of Chord Progressions in the Tonal Interval Space». *Entropy*, vol. 22, n.º 11 (2020): 1–30. <https://doi.org/10.3390/e22111291>

- Pachet, François, Pierre Roy y Benoit Carré. «Assisted Music Creation with Flow Machines: Towards New Categories of New». *ArXiv* (2020). <http://arxiv.org/abs/2006.09232>
- Paiement, Jean-François, Douglas Eck y Samy Bengio. «Probabilistic Melodic Harmonization». En *Advances in Artificial Intelligence. Canadian AI 2006. Lecture Notes in Computer Science*. Edición de Lamontagne, L., Marchand, M., 218–229. Berlín: Springer, 2006.
- Pardo, Bryan y William P. Birmingham. «Algorithms for Chordal Analysis». *Computer Music Journal*, vol. 26, n.º 2 (2002).
- Rabiner, Lawrence R. «A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition». *Proceedings on the IEEE*, vol. 77, n.º 2 (1989).
- Raczynski, Stanislaw, Satoru Fukayama, Emmanuel Vincent. «Melody Harmonisation with Interpolated Probabilistic Models». Reporte de investigación, RR-8110, INRIA, 2012. <https://hal.inria.fr/hal-00742957>
- Ramalho, Geber y Ricardo Scholz. «COCHONUT: Recognizing Complex Chords from MIDI Guitar Sequences». Conference: International Conferences on Music Information Retrieval, 2008. <https://www.researchgate.net/publication/269038716>
- Randall, Richard R. y Bilal Khan. «Lerdahl's Tonal Pitch Space Model and Associated Metric Spaces». *Journal of Mathematics and Music*, vol. 4, n.º 3 (2010): 121–131. <https://doi.org/10.1080/17459737.2010.529654>
- Rosenblatt, Murray. «Remarks On Some Nonparametric Estimates Of A Density Function». *Annals of Mathematical Statistics* (1956).
- Simon, Ian, Dan Morris y Sumit Basu. «My-Song: Automatic Accompaniment Generation for Vocal Melodies». *CHI '08: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery (2008): 725–734.
- Tsushima, Hiroaki, Eita Nakamura, Katsutoshi Itoyama, Kazuyoshi Yoshii. «Function-And Rhythm-Aware Melody Harmonization Based On Tree-Structured Parsing And Split-Merge Sampling Of Chord Sequences». 18th International Society for Music Information Retrieval Conference, Suzhou, China, 2017.
- Tsushima, Hiroaki, Eita Nakamura y Kazuyoshi Yoshii. «Bayesian Melody Harmonization Based on a Tree-Structured Generative Model of Chord Sequences and Melodies». *IEEE/ACM Transactions on Audio, Speech, and Language Processing* (2020).
- Van der Hoek, Jonh y Robert J. Elliot. *Introduction to Hidden Semi-Markov Models*. Cambridge University Press, 2018.
- Williams, J. K. «An Interactive, Multimedia Environment for Exploring Tonal Pitch Space». Conference ICL2007, Septiembre 26 –28, 2007.
- Wongsaroj, C., N. Prompoon y A. Surarerks. *Digital Information and Communication Technology and it's Applications (DICTAP)*, 2014 Fourth International Conference. 2014.
- Yang, Li-Chia y Alexander Lerch. «On the Evaluation of Generative Models in Music». *Neural Computing and Applications*, vol. 32, n.º 9 (2020): 4773–4784. <https://doi.org/10.1007/s00521-018-3849-7>
- Yeh, Yin-Cheng, Wen-Yi Hsiao, Satoru Fukayama, Tetsuro Kitahara, Benjamin Genczel, Hao-Min Liu, Hao-Wen Dong, Yian Chen, Terence Leong, Yi-Hsuan Yang. «Automatic Melody Harmonization with Triad Chords: A Comparative Study». *ArXiv* (2021).
- Zhang, Kejun, Xinda Wu, Tieyao Zhang, Zhi-jie Huang, Xu Tan, Qihao Liang, Songruo-yao Wu, Lingyun Sun. «WuYun: Exploring Hierarchical Skeleton-guided Melody Generation Using Knowledge-enhanced Deep Learning». *ArXiv* (2023).



Vol. 3 N.º 9 / Guayaquil
mayo 2024
ISSN 2953-657x

Implementación de una aplicación *software* para la emulación de una consola mezcladora de audio como herramienta de aprendizaje significativo

Semillero Salud, Discapacidad e Ingeniería
Grupo de Investigación en Acústica Aplicada
Facultad de Ingeniería Universidad de San Buenaventura
Bogotá, Colombia

Michael Javier Acevedo Cutta
Estudiante en el programa de Ingeniería de Sonido
mjacevedoc@academia.usbbog.edu.co

Josué Felipe Ordóñez Castro
Estudiante en el programa de Ingeniería de Sonido
jfordonezc@academia.usbbog.edu.co

Juan Diego Prieto Pedraza
Estudiante en el programa de Ingeniería de Sonido
jdrietop@academia.usbbog.edu.co

Daniel Leonardo Layton Rodríguez

Estudiante en el programa de Ingeniería de Sonido

dllaytonr@academia.usbbog.edu.co

Eric Julián Rodríguez Babatiba

Estudiante en el programa de Ingeniería de Sonido

ejrodriguezb@academia.usbbog.edu.co

Daniel Jara-Hurtado

Tutor del proyecto, docente en el programa de Ingeniería de Sonido

jjara@unal.edu.co - jjara@usbbog.edu.co

Resumen

A raíz de la pandemia producto de la COVID-19, la educación tradicional se vio en la obligación de adoptar alternativas para continuar la formación de estudiantes. En el caso universitario, las clases con componentes prácticos debieron ser reestructuradas implementando herramientas digitales. Dados los cambios del modelo educativo y el continuo desarrollo de soluciones digitales, así como la necesidad de capacitar a quienes inician su proceso formativo en ingeniería de Sonido, es necesario desarrollar herramientas o métodos de enseñanza que tengan en cuenta las condiciones del aprendizaje. Por lo tanto, se propone la construcción de una aplicación software que emula una consola mezcladora de audio existente en el mercado, con sus componentes básicos por canal, efectos y suma de señales, como instrumento de enseñanza de audio en estudiantes del programa de ingeniería de Sonido de la Universidad de San Buenaventura. Al implementar y validar este material de apoyo, se llegará a una herramienta de aprendizaje significativo que se utilizará como recurso en semestres iniciales del programa y como plan de contingencia ante circunstancias que afecten el desarrollo de las clases prácticas, de modo que al manipular una consola real los estudiantes estén en la capacidad de operarla adecuadamente.

Palabras clave: Educación en ingeniería, enseñanza, sistemas de audio, programación de audio.

Abstract

As a result of the COVID-19 pandemic, traditional education was forced to adopt alternatives to continue students' education. Specifically, in the case of universities, classes with practical components had to be restructured by implementing digital tools. Given the changes in the educational model and the continuous development of digital solutions, as well as the need to train those starting their educational process in Sound Engineering, it is necessary to develop tools or teaching methods that consider the learning conditions. Therefore, the construction of a software application that emulates an audio mixing console available in the market is proposed. This application will include its basic components per channel, effects, and signal summation, serving as a teaching instrument for audio in students enrolled in the Sound Engineering program at the University of San Buenaventura. By implementing and validating this supportive material, a meaningful learning tool will be created, which will be used as a resource during the initial semesters of the program and as a contingency plan in case circumstances affect the development of practical classes. Consequently, students will be capable of operating a real console properly by manipulating the simulated application.

Keywords: Engineering education, teaching, audio systems, audio coding.

1. Introducción

El proceso de desarrollo de nuevas tecnologías ligadas al aprendizaje moderno ha sido evidente, y es necesario comprender que este es un factor clave en la enseñanza del siglo XXI.¹ Reconocer las fortalezas de los estudiantes que ingresan a un pregrado de universidad permite identificar sus capacidades. Además, tener en cuenta que las tecnologías de información y comunicación (TIC) se han convertido en una herramienta de uso cotidiano ayuda a mejorar significativamente los procesos de aprendizaje de las comunidades estudiantiles. Es claro que los jóvenes de hoy poseen las aptitudes y conocimientos que les posibilitan una destreza evidente en el uso de herramientas tecnológicas.² Ahora bien, así como se menciona que las herramientas asistidas por tecnología son de gran utilidad en la academia, también se debe considerar que para ello estos materiales necesitan ser enseñados de manera eficiente. Las estrategias de aprendizaje significativo pretenden, entre otras cosas, apelar a la motivación del estudiante para promover los procesos de apropiación del conocimiento.³ A partir de la idea anterior, se considera que el aspecto de la motivación es un factor de gran potencial para el correcto acoplamiento de los estudiantes con las herramientas que usan durante su proceso de formación.⁴

En otra línea, la transición hacia la universidad conlleva un proceso de transformaciones biológicas, psicológicas y comportamentales.⁵ Crear un ambiente apto para el estudiante permite cumplir ciertas condiciones elementales para el aprendizaje, entre ellas están la motivación, la práctica, la seguridad psicológica, la experimentación, la retroalimentación y la integración.⁶

Por otra parte, desde el ámbito pedagógico, con el propósito de empezar a focalizar todos los procesos de enseñanza a los estilos más comunes, es crucial identificar los tipos de aprendizajes que manifiesta el estudiantado. En contexto, para el caso de los estudiantes de nuevo ingreso al programa de ingeniería de Sonido de la Universidad de San Buenaventura, aunque no es requisito, sí es de gran utilidad tener conocimientos sobre las herramientas básicas que se usan diariamente en el mundo del audio, de acuerdo con la experiencia de los autores de este artículo.

Uno de los conocimientos comunes que no solo facilitan el avance, sino que son elementos transversales dentro de la malla curricular (ingeniería de Sonido, Universidad de San Buenaventura, 2013), es la operación de consolas mezcladoras de audio. En la línea de producción musical del programa, la etapa en la que se comienzan a usar esas herramientas inicia en cuarto semestre, por lo que hay una brecha temporal de año y medio, además de que en aquel lapso la mayoría de las materias corresponden al componente de ciencias básicas (Física y Cálculos). Esto puede presentar consecuencias como la falta de motivación y la pérdida del interés. Asimismo, es posible percibir que el desarrollo de herramientas relacio-

1 Luz Estela Gómez-Vahos, Luz Enid Muriel-Muñoz y David Alberto Londoño Vásquez, «El papel del docente para el logro de un aprendizaje significativo apoyando en las TIC», *Encuentros*, vol. 17, n.º 2 (2019).

2 Jesús Alberto Fuenmayor, «Características de la educación en el siglo XXI», Jesús Alberto Fuenmayor, 24 de septiembre de 2020, <https://jesusfuenmayor.com/2020/09/24/caracteristicas-de-la-educacion-en-el-siglo-xxi/>

3 Falta fuente: (Colby, 2017). También falta agregar en la bibliografía.

4 Alexis Morffe, «Las TIC como herramientas mediadoras del aprendizaje significativo en el pregrado: una experiencia con aplicaciones telemáticas gratuitas», *Revista de Artes y Humanidades UNICA*, vol. 11, n.º 1, (2010): 200–219.

5 Viviana Rodríguez Caicedo, Ana Ximena Fajardo, «La Llegada a la Universidad: un Desafío de Adapta-ción», *Revista Universitaria. Docencia, Investigación, Innovación*, vol. 2, n.º 1 (2013): 90–104.

6 José Manuel Sáenz López, *Estilos de aprendizaje y métodos de enseñanza* (Editorial UNED, 2018).

nadas con audio apoyadas con elementos tecnológicos tiene un efecto de aprendizaje significativo.⁷

De ese modo, se evidencia la necesidad de crear una estrategia educativa basada en una consola mezcladora, que permita al estudiante tener un acercamiento inicial al mundo del audio y temas afines como el flujo de señal. Esta consola podría fortalecer y dar una primera aproximación a los conocimientos que serán requeridos en el futuro de su carrera académica y laboral, por medio de aprendizaje mejorado por tecnología, activo y significativo.

2. Parte técnica

En el presente proyecto se ha propuesto la emulación de una consola mezcladora de audio con base en una existente en el mercado —la Yamaha MG10XU—, lo que se explica en la primera parte de la metodología que acompaña este documento. Luego, se plantea el proceso que se ha seguido para la evaluación del emulador. Se continúa con un análisis de estos resultados, revisando cada respuesta de las personas evaluadas.

2.1. Mezcladora

Haciendo uso de herramientas de programación, se hace la emulación de la consola Yamaha MG10XU sobre el lenguaje Python, implementando la librería pyo para procesamiento de señales de audio en tiempo real. Se utilizan los siguientes métodos:

- Lectura y reproducción de un archivo de audio

⁷ Samuel Z. Fisher, «Signal Flow Training with Virtual Simulations as a Co-Curricular Tool» (conferencia International Audio Education, Audio Engineering Society, julio 22–24 de 2021).

Ashell Fox, Jiayue Cecilia Wu «Teaching Modular Synth & Sound Design Online During COVID-19: Maximizing Learning Outcomes Through Open-source Software and Student-centered Pedagogy» (Convenção n.º 151, Audio Engineering Society, outubro de 2021).

- Procesos de frecuencia sobre una señal
- Procesos dinámicos
- Procesos de tiempo sobre un audio



Imagen 1: Yamaha MG10XU. Fuente: Yamaha.

Se propone la emulación de la consola Yamaha MG10XU (Imagen 1), para ello, se dividió el desarrollo del aplicativo en tres secciones:

- Sección del Channel Strip
- Sección de Efectos
- Sección Master

A modo de referencia, el resultado final se ve de esta manera:



Imagen 2: simulación Llama-ha MG10XU Beta versión 2.

Dentro de cada etapa se encuentran los distintos elementos (controles, diales y botones) que componen a cada sección del aplicativo construido. Sin embargo, antes de explicar las funcionalidades de la Llama-ha MG10XU Beta versión 2 —se ha hecho una adaptación al nombre— es necesario conocer los aspectos técnicos que la componen.

2.1.1. Aspectos técnicos

A continuación, se presenta una breve explicación del proceso de implementación de la Llama-ha Beta versión 2.

2.1.1.1. Interfaz gráfica

Para el diseño de la interfaz gráfica, se usaron dos herramientas:

- Qt designer: permite el desarrollo al momento de crear aplicaciones gráficas, además, ofrece el uso de *widgets* (elementos gráficos). Esta aplicación permite desarrollar una interfaz gráfica (con diales y botones) que se asemeje a la consola Yamaha MG10XU.
- La librería PyQt5: permite unir el trabajo desarrollado en QtDesigner con el código en Python, que, a su vez, permite manipular todos los parámetros de audio controlados por la interfaz gráfica.

2.1.1.2. Software

El programa en su totalidad fue desarrollado en el lenguaje de programación de Python y se usa la librería pyo que permite el procesamiento de audio. Ahora, es necesario mencionar que esta librería es un conjunto de funciones o recursos que permiten resolver tareas en un menor tiempo y en un

menor código.⁸ La librería de pyo permite procesar audio en tiempo real, aunque realmente tiene un tiempo tan corto de procesamiento de las señales que hace imperceptible ese retraso. El desarrollo de código se hizo bajo el paradigma de programación orientada a objetos (POO) que posibilita la reutilización de código y una óptima organización. Del mismo modo, durante el planteamiento del proyecto se crearon las siguientes clases:

- Yamaha: permite localizar el archivo .ui de la interfaz, controlar todos los eventos de la interfaz gráfica y unir estos con los objetos de la clase Canal.
- Canal: permite manipular a través de diferentes métodos los parámetros de audio de cada canal de la interfaz —*gain*, compresión, *EQ*, *FX* (qué tan ta señal que sale de la ecualización se envía como copia al efecto), paneo, *level* y *master*—. Esta clase permitió crear los siete canales; cada objeto que se crea es de la clase canal.

Por otro lado, la librería pyo permitió, mediante las siguientes clases, la construcción de la Llama-ha MG10XU Beta versión 2.

- SfPlayer (): permite importar el audio y manejar el nivel de intensidad por medio de un multiplicador.
- Disto (): permite generar una distorsión digital en caso de que lo que el multiplicador determinó (anterior) esté cerca al límite.
- Biquad (): para el filtro pasa altos [HPF].
- Compress (): permite comprimir la señal.
- EQ(): permite la ecualización del audio.

⁸ Ministerio para la Transformación Digital y de la Función Pública, «11 librerías para crear visualizaciones de datos», datos.gob.es, 3 de mayo de 2022, <https://datos.gob.es/es/blog/11-librerias-para-crear-visualizaciones-de-datos>

- Freeverb (): agrega el efecto de Reverb.
 - Delay () y SDelay (): retrasan la señal de audio.
 - Chorus (): agrega ocho señales retrasadas y moduladas.⁹
- b. Mid
 - c. Low (salida del canal)
 6. Fx (envío de una copia de la señal).

En el caso que el usuario esté interesado en conocer más sobre esta librería, en el siguiente enlace puede encontrar toda la información necesaria: <https://acortar.link/LM16iZ>.

2.1.2. Funcionalidad

Con base en lo anterior, la funcionalidad de cada parte del aplicativo Llama-ha MG10XU Beta versión 2 está conformado de la siguiente manera:

2.1.2.1. Channel Strip

En la imagen 2 hay siete canales. Los primeros dos, a diferencia de los canales del 3-7, tienen un control de compresión. Los canales del 1-4, en la parte de ecualización, tienen tres etapas (altos-medios-bajos). Los canales 5-7 solamente tienen dos etapas de ecualización (altos-bajos). Con esto presente, se plantea el flujo señal en cada canal:

1. Gain-Pad-Pan [LR]-Level-Master (todos estos conforman el multiplicador de la señal para cada canal).
2. Distorsión: en este paso del flujo de señal se simula esta saturación por exceso de ganancia y este depende del multiplicador del paso anterior.
3. HPF (High Pass Filter).
4. Compresor.
5. Ecualización:
 - a. High

⁹ Falta fuente para (Bélanger, 2021) y colocarla en bibliografía.

Se puede observar de mejor manera en la siguiente figura:

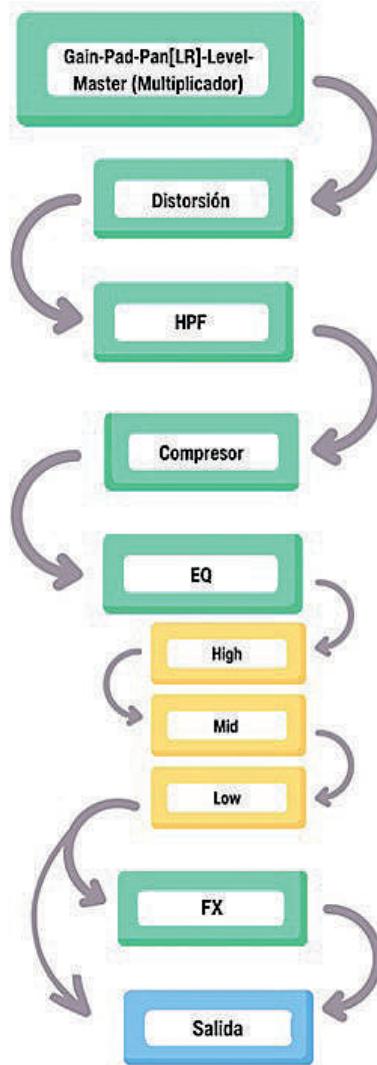


Figura 1: flujo de señal de la clase canal.

Desde el inicio del flujo de señal hasta atravesar la interfaz gráfica —que se constituye por medio de botones y diales donde se le indica al programa que cada vez que se genere un cambio (evento) cumpla con la función asignada según los parámetros que se les asignaron previamente—, de esta manera se obtiene el *channel strip* de la consola Llama-ha MG10XU Beta versión 2. Así se obtiene un flujo de señal claro con métodos eficientes para el óptimo funcionamiento de la aplicación.

Cabe destacar que, para el desarrollo del aplicativo y con el fin pedagógico planteado inicialmente, se desarrolló un manual de la aplicación para que los estudiantes puedan conocer las funcionalidades brindadas en el programa, utilizando un lenguaje sencillo y con varios ejemplos ilustrativos para algunos conceptos. Se lo puede conocer en el siguiente enlace: <https://acortar.link/9PbhIE>.

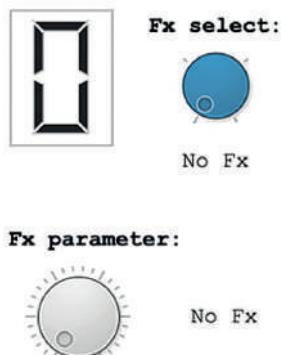


Imagen 3: sección de efectos de la Llama-ha Beta versión 2.

Esta sección recibe las copias enviadas de cada canal y permite por medio de la perilla azul, como se muestra en la imagen 3, escoger entre los tres efectos. Por su parte, la perilla blanca permite modificar un parámetro según el efecto seleccionado.

1. *Reverb (Room size)*
2. *Delay (tiempo de retraso en BPM)*
3. *Chorus (Depth)*

2.1.2.2. Master

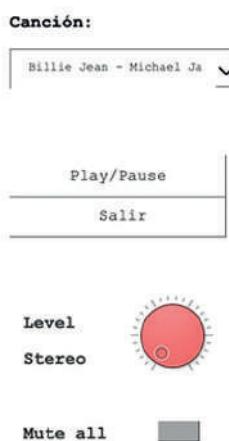


Imagen 4: sección Master de la Llama-ha Beta versión 2.

Esta sección permite controlar el nivel de intensidad de los siete canales. Además tiene un botón que permite silenciarlos a todos, pues este se encarga de detener el servidor de audio de cada canal. Como se explica en la sección del *Channel Strip*, el valor de la perilla del *Master* se multiplica junto con el *Gain-Pad-Pan-Level* de cada canal, para así permitirle controlar el nivel de intensidad a partir de la configuración individual de cada etapa.

Por último, en esta sección se encuentran el *ComboBox* —que permite escoger entre cuatro *multitracks* con instrumentos ya predeterminados para cada canal—, los botones de *play/pause* para controlar la reproducción de audio, y el botón de *salir* que permite detener el servidor del audio interno y cerrar el aplicativo.

2.2. Instrumento de evaluación

Con base en trabajos previamente desarrollados por este grupo,¹ se ha propuesto evaluar desde un paradigma cualitativo la utilidad y beneficio del emulador de consola como una herramienta que pueda aportar a procesos de aprendizaje significativo² en audio. Esto hará posible la implementación de este tipo de herramienta para procesos de formación educativa asistida por tecnología.³

1 M. Acevedo et al., «Exploración de una Propuesta Pedagógica para una Mezcladora de Audio dirigida a Estudiantes de Ingeniería» (Encuentro de Semilleros de Investigación 2022, Editorial Uniagustiniana, 2022).

J. Ordóñez et al., «Exploración de una propuesta didáctica para una consola mezcladora dirigida a estudiantes de ingeniería de Sonido», Jornadas de Acústica, Audio y Sonido UNTREF, 2022, https://usbogedu-my.sharepoint.com/:b/g/personal/jjara_usbbog_edu_co/EdNh8m3yYhBBqpg-VayYoEDwB9PxOSnjcQqh5mEpJwTvlw?e=HHal7j

2 Daniel Yasser Cárdenas Enríquez, Victoria Paula Murgado Hamann, Jesús Arturo Quispe Rodríguez y José Antonio Vargas Tour, «Uso de videotutoriales en la construcción de aprendizaje significativo en las clases de los estudiantes del 6º ciclo del curso de “Edición de Audio”, de la carrera de Ciencias de Comunicación de una universidad privada de Lima» (tesis de maestría, Repositorio Universidad Tecnológica del Perú, 2018).

3 Mercedes de la E. Inciarte Rodríguez, «Tecnologías de la información y la comunicación. Un eje transversal para el logro de aprendizajes significativos», REICE Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación, vol. 2, n.º 16 (2004).

Para el instrumento de evaluación, se establecen tres etapas o momentos que se realizan en un lapso de 40 minutos:

- a) Previo (diagnóstico de entrada o sondeo): 15 minutos
- b) Durante (enseñanza y exploración): 20 minutos
- c) Salida (de diagnóstico de salida o sondeo de salida): 5 minuto

Lo anterior se realiza con base en un proceso de enseñanza-aprendizaje¹⁰ que tenga en cuenta un planteamiento de reconocimiento, de transmisión de conocimiento y de verificación de aprendizaje del estudiante.

2.2.1. Previo (diagnóstico de entrada o sondeo)

En los primeros cinco minutos se realiza una introducción del ejercicio a los participantes. Luego se les pide responder una encuesta de diez minutos en la que se determinan o exploran sus conocimientos actuales. En este cuestionario se incluyen temas que tienen relación con el funcionamiento de una mezcladora de audio, como sus etapas, efectos o el flujo de señal en este elemento.

2.2.2. Durante (enseñanza y exploración)

Los evaluadores (los autores de este documento) imparten una lección magistral como método de enseñanza. Allí se explican las etapas de un *channel strip* y cómo se ve reflejado este flujo en una señal de audio. Luego se le permite a los participantes probar durante cinco minutos el aplicativo en modo libre, además de darles acceso al manual presentado en la sección 2.1.2.1. Terminado este tiempo, se presenta un caso

en el que los canales requieren parámetros y efectos específicos, haciendo que el participante utilice los conocimientos previamente adquiridos para llegar a cumplir lo propuesto en la prueba. Esto lleva a aplicar un aprendizaje basado en problemas (ABP), con el que se adquieren habilidades de resolución de problemas mientras se familiarizan con conocimientos básicos del flujo de señal en una mezcladora de audio.

2.2.3. Salida (diagnóstico o sondeo de salida)

Finalmente, el grupo evaluador revisa el progreso realizado en el aplicativo por cada individuo y establece una calificación utilizando una rúbrica analítica¹¹ en la que se desglosan todos los criterios que se solicitan en la prueba. Así, se logra obtener una imagen clara de qué aspectos fueron efectivos o deficientes en este proceso de aprendizaje.

En cada elemento se puede evidenciar la siguiente clasificación:

Tabla I. Clasificación de elementos de la rúbrica

< 3,0	Es necesario mejorar significativamente las habilidades y comprensión de este criterio, ya que no se realiza lo solicitado.
3,0 – 4,0	Se cumple con este criterio, no obstante falta precisión de acuerdo con lo solicitado, por lo que aún existen aspectos a mejorar y fortalecer en el manejo y comprensión de este punto.
4,0 – 5,0	Se demuestra comprensión y dominio de este criterio, donde sigue las instrucciones con precisión en el tiempo de la prueba.

Se califican los siguientes criterios:

- Selección de multitrack
- Ajuste de ganancia
- Ecualización de canal

¹¹ Falta fuente para (Cortés de las Heras, 2014) y poner en bibliografía.

- Aplicación de efectos
- Ubicación estéreo (paneo)
- Nivel de cada canal (volumen)
- Nivel de salida (*master*)

De ese modo, el evaluador puede brindar una retroalimentación en la que indica las fallas y los aciertos que tuvo cada participante.

Posteriormente, se realiza una prueba cualitativa de rendimiento en la que se presentan preguntas abiertas de tipo descriptivo. Allí se espera que el estudiante demuestre haber adquirido conocimientos.

3. Resultados

En esta sección se analizan el desempeño del aplicativo y los resultados del instrumento de evaluación.

3.1. Llama-ha MG10XU Beta versión 2

Al terminar el desarrollo del aplicativo, se tuvo un periodo de una semana para probar todas las funcionalidades de la Llama-ha MG10XU Beta versión 2; se identificaron errores en botones y en el ComboBox. Durante ese periodo, dichos *bugs* se lograron corregir, y de esta manera se obtuvo una aplicación eficiente en las funcionalidades planteadas, ya que las etapas de cada canal (ganancia, compresión, filtros, efectos, pano, volumen y *master*) se desempeñaron correctamente durante su uso por los estudiantes evaluados y en conjunto (los siete canales).

Para finalizar la ejecución de la aplicación, se debe oprimir el botón de *salir* (como se muestra en la imagen 4). Si esta acción no se realiza, puede ocasionar conflictos con el Kernel de Python, es decir, el responsable de que el código se ejecute.

Al momento de realizar la prueba con el primer grupo de estudiantes se produjo

una leve falla con el programa, ya que el servidor de pyo dejó de funcionar. Luego de hacer un análisis, se llegó a la conclusión de que el problema estaba ligado al procesador del computador, ya que paró su funcionamiento a los doce minutos luego de haber sido ejecutado. Por ese motivo, para una próxima evaluación de la aplicación, se espera corregir errores de rendimiento de la Llama-ha MG10XU Beta versión 2.

Por otro lado, se recomienda para una siguiente versión un control de reproducción de audio como el de Windows Media Player,¹² así como un medidor de los niveles de la señal, a manera de guía visual de la intensidad de la señal. Ambas alternativas permitirán un mayor control de la mezcla y una experiencia más didáctica para el usuario.¹³ También, para una próxima versión, se recomienda realizar el ejecutable del aplicativo para una mayor portabilidad del software.

3.2. Instrumento de evaluación

Luego de recopilar las respuestas de carácter cualitativo, se procede a realizar una síntesis en la que se presentan los resultados de manera cuantitativa. De esa forma, se obtiene un modelo numérico para entender de manera precisa los resultados. Para la recolección de datos se utilizó la herramienta de Microsoft Forms y, posteriormente, con la ayuda de Excel, se realizó la cuantificación de cada pregunta por estudiante. Allí se agregaron las tablas que muestran los resultados alcanzados.

En el primer momento se identifica si la población tiene experiencia dentro del mundo del audio. Se demostró que los participantes evaluados no tienen afinidad en esta área.

12 Ordóñez *et al.*, «Exploración de una propuesta...».
13 Ordóñez *et al.*, «Exploración de una propuesta...».

Tabla II
Primera pregunta de la encuesta

	SÍ	NO
¿Tiene experiencia en el mundo del audio?	0%	100%

Para catalogar las respuestas de la encuesta se utiliza la siguiente clasificación en porcentaje del total de la muestra:

Tabla III
Clasificación de las siguientes tablas

%	Responde correctamente
%	Responde con una noción básica
%	Respuesta errónea
%	No sabe / No responde

De tal manera que, en el diagnóstico de entrada, gran parte de los participantes respondieron «No sabe/No responde». Esta fue la respuesta más común a excepción de la pregunta acerca de efectos de audio, donde la mayoría responde con una noción básica.

Tabla IV
Resultados del diagnóstico de entrada

	%	%	%	%
¿Qué entiende por mezcladora de audio?	33	66	-	
¿Qué entiende por un <i>channel strip</i> ?	17	-	-	83
Mencione las etapas que este comprende:	-	17	-	83
Explique brevemente: ¿cómo considera usted que funciona el flujo de señal en una consola mezcladora?	-	-	50	50
¿Qué entiende por Pre-amplificación?	16,6	16,6	16,6	50
¿Qué entiende por efectos en audio?	-	83	17	-
¿Qué entiende por imagen panorámica (paneo) en audio?	16,6	16,6	16,6	50
¿Qué entiende por ecualización?	-	33	-	67

Después del espacio de enseñanza y exploración del aplicativo, se obtienen los

siguientes resultados de la actividad propuesta para evaluar el componente práctico, teniendo en cuenta la rúbrica planteada en la tabla 1. Se observa en la mayoría un valor sobresaliente (superior a 40), lo que permite evidenciar un comportamiento favorable al usar la herramienta Llama-ha MG10XU Beta versión 2.

Tabla V
Promedio de calificaciones de la actividad práctica

Estudiante	Nota
1	41
2	43
3	36
4	47
5	46
6	39
Promedio	42

En el diagnóstico de salida, se obtienen los siguientes resultados:

Tabla VI
Resultados del diagnóstico de salida

	%	%	%	%
Explique que es una mezcladora de audio.	-	67	33	-
Explique cómo se estructura un <i>channel strip</i>	17	33	33	17
Explique cómo funciona una etapa de preamplificación.	-	-	50	50
Explique cómo funciona un proceso de ecualización.	-	50	17	33
Explique cómo funciona el pano en una mezcladora de audio.	33	-	17	50
Explique cuál es la diferencia entre ganancia en el preamp y volumen en el fader.	-	16,5	16,5	67
Explique con sus palabras cómo funciona el flujo de señal en la mezcladora de audio.	17	-	50	33

Finalmente, se le presenta al participante una imagen de la mezcladora y debe reconocer qué etapas la componen. A continuación, se evidencia cuántas de las diez partes fueron acertadas:

Tabla VII
Resultados de la pregunta final

Relacione el número de la parte de la consola con el nombre de la etapa correspondiente.	>7	6-4	<=3
	50 %	17 %	33 %

4. Conclusiones

A partir del trabajo realizado, se llegaron a las siguientes conclusiones:

- La consola mezcladora Llama-ha MG10XU Beta versión 2 es viable para su uso como herramienta de práctica para nociones básicas de flujo de señal y clasificación de partes de una consola mezcladora (como la mostrada en la imagen 1).
- En temas relacionados con el software, el aplicativo Llama-ha MG10XU Beta versión 2 obtuvo los resultados esperados, ya que cumplió con todas las funcionalidades propuestas desde un inicio. Aparte del error del servidor de audio en la primera prueba, no presento ningún otro bug, cumpliendo así su propósito de manera óptima.
- Durante el uso del aplicativo por parte de los estudiantes se evidenció que el dispositivo donde se ejecuta la aplicación puede tener influencia en su correcta operación, por lo que se deben considerar elementos como CPU y RAM.
- La aplicación del proyecto a una población que posee nociones básicas sobre los elementos de la mezcladora permitió reforzar y aclarar los conceptos de cada componente durante la explicación.
- Los resultados del instrumento de evaluación demostraron, en primera instancia, que el aprendizaje observacional y de rutina son relevantes

para una buena apropiación de los conocimientos impartidos durante la actividad, debido a que la interacción con la aplicación permite una construcción de conocimiento a raíz de la práctica.

- A pesar de los resultados obtenidos, se debe considerar la revisión y validación de las preguntas que conforman el instrumento de evaluación con expertos en diseño y elaboración de rúbricas en específico.

5. Referencias

Acevedo, M., D. Layton, J. Ordóñez, J. Prieto y D. Jara-Hurtado. «Exploración de una Propuesta Pedagógica para una Mezcladora de Audio dirigida a Estudiantes de Ingeniería». Encuentro de Semilleros de Investigación 2022, Editorial Uniagustiniana, 2022.

Cárdenas Enríquez, Daniel Yasser, Victoria Paula Murgado Hamann, Jesús Arturo Quispe Rodríguez y José Antonio Vargas Tour. «Uso de videotutoriales en la construcción de aprendizaje significativo en las clases de los estudiantes del 6º ciclo del curso de “Edición de Audio”, de la carrera de Ciencias de Comunicación de una universidad privada de Lima». Tesis de maestría, Repositorio Universidad Tecnológica del Perú, 2018.

Fisher, Samuel Z. «Signal Flow Training with Virtual Simulations as a Co-Curricular Tool». Conferencia International Audio Education, Audio Engineering Society, julio 22-24 de 2021.

Fox, Ashell y Jiayue Cecilia Wu. «Teaching Modular Synth & Sound Design Online During COVID-19: Maximizing Learning Outcomes Through Open-source Software and Student-centered Pedagogy». Convención n.º 151, Audio Engineering Society, octubre de 2021.

- Fuenmayor, Jesús Alberto. «Características de la educación en el siglo XXI». *Jesús Alberto Fuenmayor*, 24 de septiembre de 2020. <https://jesusfuenmayor.com/2020/09/24/caracteristicas-de-la-educacion-en-el-siglo-xxi/>
- Gómez-Vahos, Luz Estela, Luz Enid Muriel-Muñoz y David Alberto Londoño Vásquez. «El papel del docente para el logro de un aprendizaje significativo apoyado en las TIC». *Encuentros*, vol. 17, n.º 2 (2019).
- Inciarte Rodríguez, Mercedes de la E. «Tecnologías de la información y la comunicación. Un eje transversal para el logro de aprendizajes significativos». *REICE Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, vol. 2, n.º 16 (2004).
- Ministerio para la Transformación Digital y de la Función Pública. «11 librerías para crear visualizaciones de datos». datos.gob.es, 3 de mayo de 2022. <https://datos.gob.es/es/blog/11-librerias-para-crear-visualizaciones-de-datos>
- Morffe, Alexis. «Las TIC como herramientas mediadoras del aprendizaje significativo en el pregrado: una experiencia con aplicaciones telemáticas gratuitas». *Revista de Artes y Humanidades UNICA*, vol. 11, n.º 1, (2010): 200–219.
- Ordóñez, J., M. Acevedo, D. Layton, J. Prieto y D. Jara-Hurtado. «Exploración de una propuesta didáctica para una consola mezcladora dirigida a estudiantes de ingeniería de Sonido». *Jornadas de Acústica, Audio y Sonido UNTREF*, 2022. https://usbboge-du-my.sharepoint.com/:b/g/personal/jjara_usbbog_edu_co/EdNh8m3yYhB-BqpgVayYoEDwB9PxoSnjicQqH5mEpJw-TvJw?e=HHaI7j
- Rodríguez Caicedo, Viviana, Ana Ximena Fajardo. «La Llegada a la Universidad: un Desafío de Adaptación». *Revista Universitaria. Docencia, Investigación, Innovación*, vol. 2, n.º 1 (2013): 90–104.
- Sáenz López, José Manuel. *Estilos de aprendizaje y métodos de enseñanza*. Editorial UNED, 2018.
- Webedia Brand Services. «Por qué la relación entre CPU, RAM y disco duro decide el rendimiento». Territorio Intel, 23 de octubre de 2019. <https://territoriointel.xataka.com/que-relacion-cpu-ram-almacenamiento-decide-rendimiento/>
- Yamaha. «MG10XU 10-Channel USB Mixing Desk with FX (Pre-Owned)». Kenny's Music. <https://www.kennysmusic.co.uk/yamaha-mg10xu-10-channel-usb-mixing-desk-with-fx-pre-owned-p5511>



Vol. 3 N.º 9 / Guayaquil
mayo 2024
ISSN 2953-657x

Software adaptivo para el entrenamiento auditivo técnico en las aulas de clases

Juan Pablo Corredor

Facultad de ingeniería, Pontificia Universidad Javeriana
Bogotá, Colombia
j-corredor@javeriana.edu.co

Resumen

Este *paper* estudia entrevistas e investigaciones realizadas alrededor de la docencia del entrenamiento auditivo técnico (o escucha crítica) en las universidades y escuelas. Busca posibles puntos por mejorar o nuevas vías en el desarrollo de *software* para minimizar la frustración y dificultad de los estudiantes, y optimizar la labor de los docentes.

Palabras clave: Entrenamiento auditivo, escucha crítica, pedagogía, programación, *software*

Abstract

This paper focuses on interviews and research carried around the teaching of technical ear training (also known as critical listening) in universities and colleges. We try to find potential improvements or new avenues for software development to minimize frustration in students and to optimize professors' duties.

Keywords: Ear training, critical listening, teaching, programming, software

1. Introducción

La enseñanza del entrenamiento auditivo técnico, también llamado ocasionalmente escucha crítica, es una práctica relativamente reciente en muchas escuelas de audio profesional y música. Se ha estado estudiando desde perspectivas pedagógicas cada vez más enfocadas hacia la eficiencia del aprendizaje y su facilitación.¹

No es una cuestión sorprendente, pues tanto los profesionales del audio como los mismos estudiantes son conscientes de que desarrollar la habilidad de asociar valores cuantitativos con impresiones aurales (como lo describe Jason Corey, experto en el tema) facilitará sus labores en el campo del audio como ingenieros de mezcla, de grabación o de sonido en vivo más adelante.

Sin embargo, el aprendizaje de la escucha crítica sigue siendo un campo limitado en cuanto a sus recursos disponibles, y en especial a las metodologías que se emplean para su enseñanza.² Muchas escuelas recurren a una metodología de repetición aleatoria, donde los estudiantes son expuestos a tipos de ejercicios (principalmente enfocados en el reconocimiento de cambios de intensidad en bandas de frecuencia) repetidamente hasta que logran crear las asociaciones necesarias para poder reconocerlos en contextos musicales o dialécticos.

Esto supone un problema considerable para los docentes, pues al ser un campo tan dependiente de las habilidades individuales de cada estudiante, y a la vez ser un proceso donde las dificultades específicas de cada uno pueden variar significativamente, se crean condiciones de desigualdad dentro de las aulas. Allí, los individuos que no encuentran afinidad con las metodolo-

gías o recursos utilizados tampoco tienen un análisis de resultados a gran escala que les permita llevar a cabo sesiones de estudio específicas y enfocadas, sino que deben cumplir con ejercicios uniformes para todos los estudiantes, e ir al mismo ritmo de alumnos sobresalientes, o incluso limitar el ritmo al que un aula puede avanzar.

Esta aproximación general al entrenamiento auditivo técnico (EAT) puede individualizarse con nuevos esquemas y modelos de software diseñados específicamente para la pedagogía universitaria. En este documento de investigación se presentarán algunas de las falencias que actualmente existen en las herramientas de EAT de acuerdo con docentes del área en la Pontificia Universidad Javeriana y la Virginia State University, así como publicaciones realizadas en el *Journal of the Association for Technology in Music Instruction* (JATMI).

2. Objetivos de la investigación

En este documento se proponen algunas características relevantes para el desarrollo de software pedagógico en el área de EAT utilizando entrevistas y estudios bibliográficos como referencia. Este artículo es una mirada al estado actual de las herramientas digitales, y algunos posibles caminos para su futuro desarrollo.

Se establecen dificultades pedagógicas halladas por docentes del EAT, y se habla de funcionalidad y modelos de software que puedan ayudar a facilitar el proceso de enseñanza y aprendizaje en aulas universitarias o escolares.

Asimismo, esta investigación permite construir una base teórica sólida para desarrollar a futuro un programa que cumpla con las características mencionadas.

También se presenta una relación entre aplicaciones de aprendizaje de otras

¹ Wellington Gordon, «Considering Best Practices in Technical Ear Training», *Journal of the Association for Technology in Music Instruction*, vol. 2, n.º 1 (2021).

² Jason Corey, «Technical Ear Training: Tools and Practical Methods», *Proceedings of Meetings on Acoustics*, vol. 19, n.º 1 (2013).

áreas como el estudio de los idiomas o la medicina (Duolingo, Anki, entre otras)³, y el aprendizaje de la escucha crítica, al proponer una aplicación práctica de los algoritmos de repetición espaciada en un contexto de audio profesional.

Por último, se propondrán algunas herramientas de programación (lenguajes, librerías y documentación) que pueden resultar particularmente efectivas para el adecuado desarrollo de una aplicación con estas características.

3. Aspectos pedagógicos y debilidades de las herramientas disponibles

A continuación, se presentarán algunos de los aspectos pedagógicos por mejorar que los mismos docentes de EAT entrevistados mencionaron.

3.1. Estandarización

Uno de los problemas más importantes que menciona el profesor Wellington Gordon de Virginia State University, tanto en nuestra entrevista como en su investigación del 2021, es la falta de estandarización pedagógica sobre el EAT. Por el momento no existe una claridad real en el mundo del audio sobre cómo se debería enseñar la escucha crítica ni cuáles son las competencias principales que se deberían entrenar en un futuro profesional del audio.⁴

Existe una fuerte tendencia hacia el estudio de los cambios en el espectro en frecuencia, que pretenden simular escenarios de ecualización o de reconocimiento tímbrico en el ejercicio profesional. No obstante, todas las otras competencias que se presentan en los cursos de entre-

namiento auditivo técnico varían considerablemente de escuela a escuela, incluso dentro de un mismo país o de una misma ciudad.

Esta falta de estandarización dificulta el desarrollo de prácticas y herramientas especializadas para la pedagogía del EAT, pues mientras que en la Facultad de Artes de la Universidad A pueden estar practicando tiempos de *delay* como una competencia importante, es posible que en el Departamento de Audio de la Universidad B nunca se hable del tema en clase, y más bien se opte por realizar ejercicios enfocados en las reverberaciones y los recintos.

Para llegar a un consenso profesional todavía faltan años de estudios, investigaciones y una labor pedagógica internacional, por lo que esta investigación no pretende solucionar el problema de la estandarización. Sin embargo, se ha utilizado lo más cercano que existe a un estándar en la pedagogía del EAT, es decir, el estudio de los cortes y aumentos de intensidad en bandas de frecuencia, para buscar soluciones a los otros problemas que se mencionarán en el documento.

3.2. Personalización

La personalización en la pedagogía es un tema que se ha desarrollado a la par con el *software*. Naturalmente, una de las grandes ventajas de trabajar el desarrollo de *software* como una herramienta pedagógica es que se pueden personalizar los cursos a un grado mucho mayor.

La personalización ha demostrado resultados positivos en grupos de prueba en el área del EAT, y se tiene evidencia de que una instrucción personalizada hace el proceso de aprendizaje más eficiente para los estudiantes. Sin embargo, hasta el momento, la educación personalizada ha estado considerablemente limitada y no ha visto

³ Anki, «Background», Anki, 2006, <https://docs.ankiweb.net/background.html>

⁴ Gordon, «Considering Best Practices...».

una adopción generalizada por parte de las escuelas y universidades.⁵

Este es posiblemente uno de los puntos cruciales en el área del EAT. Al tener como factor influyente la percepción auditiva y el entendimiento de ella de cada estudiante, la personalización puede convertir un arduo proceso grupal en un proceso clínico y enfocado para ayudar a hacer más eficiente el esfuerzo que los estudiantes le dedican a practicar.

La profesora Estefanía Toro de la Universidad Javeriana menciona en nuestra entrevista cómo las limitaciones de enseñar a un grupo completo con herramientas que presentan ejercicios aleatorios idénticos a todos los estudiantes hacen que aquellos con dificultades en su asignatura deban dedicarles muchísimo más tiempo a sus estudios. Naturalmente, el rendimiento individual siempre será un factor innegable en las evaluaciones académicas, pero una herramienta personalizada podría ayudar a reducir estas brechas en desempeño considerablemente.

3.3. Análisis estadístico

Otra dificultad con la que los docentes de EAT se han encontrado en su labor es la de monitorear los resultados y el desempeño de los estudiantes a gran escala. La mayoría de los docentes entrevistados estuvieron de acuerdo en que una vista general de sus aulas era necesaria, y habían optado por opciones proporcionadas por terceros.

Algunos utilizan encuestas o formularios de Google o Microsoft para poder tener un registro del desempeño académico de sus estudiantes en detalle, y otros lo mantenían manualmente en tablas de datos. Pero la desventaja de este procedimiento es que abre la puerta al error humano, y

⁵ Sungyoung Kim, «An Assessment of Individualized Technical Ear Training for Audio Production», *The Journal of the Acoustical Society of America*, vol. 138, n.º 1 (2015).

supone una mayor inversión de tiempo por parte de los docentes. Sin mencionar que tampoco provee un análisis estadístico real a los estudiantes, a menos que los docentes compartan en tiempo real los resultados con las clases.

Son demasiadas variables que pueden dificultar tener una perspectiva general del proceso del grupo en el tiempo. Esto también obscurece la información para todos los involucrados.

La falta de un análisis estadístico de fácil consumo y con criterios específicos hace que los mismos estudiantes muchas veces no sean conscientes de sus fortalezas y debilidades, y que, para tener una visión general de su aula, los maestros deban invertir más tiempo y esfuerzo del ideal.

3.4. Accesibilidad y difusión

El último gran problema que se ha encontrado en las metodologías actuales es el difícil acceso o conocimiento de muchas herramientas existentes para el estudio del EAT.

Para empezar, son escasas las herramientas que disponen de funciones para los docentes y la enseñanza en el aula de clases. La mayoría de programas disponibles se enfocan en el estudio individual y privado sobre un estudio social o supervisado como suele ocurrir en las aulas universitarias.

Incluso considerando tanto opciones comerciales como libres u *open source*, gran parte del software de fácil acceso existente fue desarrollado hace más de una década, y requiere descargas, instalaciones o suscripciones pagas. La Pontificia Universidad Javeriana en Bogotá utiliza el software TET (Technical Ear Training) desarrollado en McGill University⁶, así como una suite de aplicaciones desarrolladas por miembros de

⁶ René Quesnel, «A computer-assisted method for training and researching timbre memory and evaluation skills» (tesis de doctorado, Faculty of Music, McGill University, 2001).

la misma universidad en lenguajes gráficos como Pure Data⁷, entre otras herramientas comerciales como SoundGym⁸ o Pro Audio Essentials de iZotope⁹, mientras que el maestro Wellington Gordon de Virginia State University afirmó utilizar Auricula¹⁰ junto con una gran variedad de herramientas de manejo de información por internet.

Las herramientas que requieren una descarga e instalación en dispositivos de los estudiantes resultan inconvenientes y sufren de una menor capacidad de difusión. Sin embargo, las herramientas disponibles en línea no están diseñadas para asistir a los maestros con un análisis estadístico completo, y evidentemente van dirigidas más a un proceso de autoaprendizaje divorciado del contexto del aula universitaria.

4. Consideraciones para software de uso pedagógico en el EAT

Siendo este el caso, como desarrollador de software he pensado en algunos elementos a los que podrían aspirar nuevos proyectos que se enfoquen principalmente en ayudar a estudiantes y docentes en sus labores, que sean de más fácil acceso y que se ajusten a cierto estándar. Debido a la limitada investigación en otras habilidades de escucha crítica, el foco caerá principalmente sobre el estudio de reducciones y aumentos de intensidad en bandas de frecuencia, pues aún es necesario hacer juicios académicos sobre la efectividad y metodología de la enseñanza de reconocimiento de procesos como la compresión, reverberación, delay o cambios de fase.

⁷ Miller S. Puckette, «Pure Data», ICMC (septiembre de 1997).

⁸ SoundGym, «Audio ear training & learning center», SoundGym, 2016, <https://www.soundgym.co/site/about>

⁹ iZotope, «iZotope announces PAE», iZotope, 2016, <https://www.izotope.com/en/company/press/2016/izotope-announces-pae.html>

¹⁰ Jason Corey, *Audio Production and Critical Listening: Technical Ear Training* (Taylor & Francis, 2016).

4.1. Análisis estadístico y difusión a través de software web y manejo de bases de datos en la nube

La primera vía posible para el desarrollo de nuevo software en esta disciplina es utilizar un lenguaje de programación web como WebAssembly para poder crear una herramienta eficiente y funcional en el ámbito del audio.

WebAssembly funciona como un intérprete de bajo nivel de lenguajes de programación que permite escribir código en lenguajes comunes como C++/C, Java o Rust y convertirlos a aplicaciones web. Es ideal para aplicaciones de audio, que suelen ser exigentes en su desempeño en tiempo real, y son poco flexibles con las latencias o las imprecisiones. Usando WebAssembly es posible construir una estructura robusta y longeva que pueda correr sobre navegadores web. Esto facilitaría la difusión y accesibilidad de la herramienta dramáticamente, reduciendo la necesidad de realizar descargas o instalaciones adicionales.¹¹

Adicionalmente sería necesario contar con una base de datos que pueda almacenar usuarios y su desempeño personal. Esto permitiría eventualmente la creación de sistemas donde los docentes puedan administrar grupos para sus clases y observar el desempeño y progreso en tiempo real de sus estudiantes. Los estudiantes, por otra parte, podrían ver su desempeño en un período de tiempo específico, y ver una discriminación clara de bandas de frecuencia que les causan particulares problemas en el reconocimiento en diferentes contextos, o secuencias de respuestas que les resultan particularmente fáciles de responder. Así la misma plataforma podría ajustar sus algoritmos para agilizar el proceso de recono-

¹¹ Andreas Haas et al., «Bringing the web up to speed with WebAssembly», *Proceedings of the 38th ACM SIGPLAN Conference on Programming Language Design and Implementation* (2017): 185–200.

cimiento para cada estudiante de acuerdo a sus necesidades específicas.

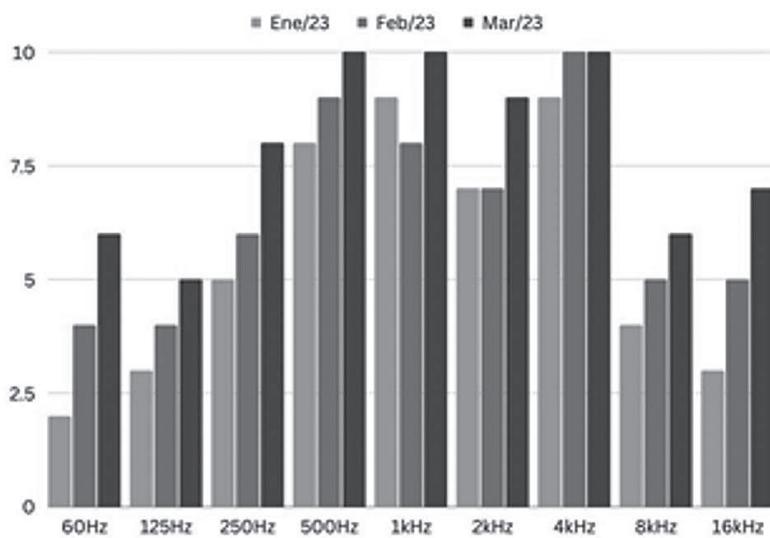


Figura 1: ejemplo de gráfico de barras con desempeño por mes en bandas de frecuencia.

Por último, se puede tomar como referencia el *software* existente para construir una herramienta que se vaya ajustando a un estándar: reconocimiento de frecuencias por octavas y tercios de octava, con factor de calidad y cambio en decibeles ajustables, con la posibilidad de reproducir un archivo de audio para evaluar cambios de frecuencia sobre música, voz humana o grabaciones de ambientes, así como un generador de ruido rosa para escuchar los cambios de frecuencia en un marco más abstracto.

4.2. Personalización a través de algoritmos de repetición espaciada

Idealmente, una vez el *software* esté en una etapa de prueba básica se probarían distintos algoritmos para la personalización de los ejercicios.

Una alternativa muy atractiva que ha recibido una recepción positiva de estudiantes de distintos campos son los algoritmos de repetición espaciada¹², donde una persona se

expone a un ejercicio, lo responde, compara su respuesta con la opción correcta, y evalúa la pregunta de acuerdo con la dificultad percibida. Si el estudiante cree que tuvo dificultad para responder la pregunta, la evalúa con un punaje bajo, mientras que, de responderla correctamente sin problemas, le asignaría un punaje alto.

En principio, este mecanismo funciona bien para ejercicios lingüísticos o académicos en los que los estudiantes

deben memorizar palabras o conceptos. Sin embargo, el audio es mucho más contextual, por lo que crear estas asociaciones resultará más difícil para un estudiante si no se considera también el contexto de las preguntas. Para esto, una aproximación interesante podría darse por medio de una modificación a los algoritmos de repetición espaciada, que evalúa también las preguntas anteriores a las que son expuestas los estudiantes. En este caso, el ejercicio que se registraría como un intento no consistiría exclusivamente en una respuesta aislada, sino en una posible secuencia de dos o hasta tres respuestas. Y cada grupo de respuestas se repetiría con la intensidad y consideraciones necesarias descritas por Tabibian *et al.*¹³

Esto sería un cambio sustancial de otros algoritmos de personalización propuestos en el pasado. Por ejemplo, en 2011, investigadores de la Universidad de Tsukuba en Japón plantearon un algoritmo de repetición balanceada específicamente para EAT que repetiría las preguntas dependien-

¹² David Ausubel y Mohamed Youssef, «The Effect of Spaced Repetition on Meaningful Retention». *The Journal of General Psychology*, vol. 73, n.º (1) (1965): 147-150.

¹³ Behzad Tabibian *et al.*, «Enhancing Human Learning Via Spaced Repetition Optimization», *Proceedings of the National Academy of Sciences*, vol. 116, n.º 10 (2019): 3988-3993.

do del desempeño de los estudiantes.¹⁴ Sin embargo, un factor importante a considerar es que las preguntas de los ejercicios de escucha crítica no ocurren aisladas, sino usualmente en sucesión. Para muchos estudiantes no es lo mismo reconocer un aumento con frecuencia central de 4 kHz luego de escuchar un aumento con frecuencia central de 2 kHz, comparado con 60 Hz o 16 kHz. Este tipo de consideraciones afectarían el algoritmo de detección.

Por ejemplo, si un estudiante tiene consistentemente respuestas incorrectas al reconocer aumentos en la banda de 4 kHz, no se repetiría incesantemente esta banda sin mayor consideración, sino que se evaluarían las bandas que el estudiante escuchó inmediatamente antes para buscar combinaciones sonoras que puedan causar mayor dificultad en el estudiante. Así, no solo se estaría ajustando constantemente la frecuencia con la que el usuario va a escuchar secuencias específicas de ejercicios, sino también se estarían identificando posibles patrones a mayor escala en grupos enteros de estudiantes.

Esto también facilita la implementación de algoritmos de repetición espaciada, que usualmente dependen de grandes grupos de información para memorizar y presentar al usuario.¹⁵ Incluso evaluando el reconocimiento de tercios de octava, el conjunto de preguntas posibles en los programas de EAT es bastante reducido. Pero al considerar distintas combinaciones de preguntas, el tamaño del conjunto crece exponencialmente, lo que permite una mayor flexibilidad a la hora de ajustar las preguntas en el tiempo.

Esta es la verdadera ventaja de buscar un desarrollo de software grupal para

las aulas de clases. También permite que, por ejemplo, si uno de los audios utilizados para los ejercicios resulta particularmente difícil en algunas bandas de frecuencia, se pueda normalizar el desempeño de la clase considerando los resultados obtenidos, y se puedan tomar decisiones futuras sobre los recursos a utilizar en las clases.

5. Descripción y uso del software

Luego de discutir dificultades encontradas en la enseñanza y aprendizaje del EAT, así como algunas opciones que la programación contemporánea ofrece para crear programas personalizables, de fácil acceso y con un buen manejo de la información para proveer retroalimentación inmediata a los estudiantes, a continuación, se describirá un programa prototípico que cumpla con estas características.

Se podrá acceder al programa directamente desde un navegador web, desde dispositivos de escritorio, así como desde dispositivos móviles sin la necesidad de instalar o descargar nada, únicamente con un URL. Por supuesto, la interfaz gráfica del programa también tendrá una versión de escritorio y una versión móvil.

El usuario podrá decidir reconocer aumentos y cortes por bandas de frecuencia sobre ruido rosa, o subir un archivo de audio (en algún formato sin pérdidas como .wav, .aiff o .flac) para realizar los ejercicios sobre música o voz. El programa también podrá ofrecer una selección de audios (libres de derechos de autor) escogidos especialmente por sus características espectrales. Estos audios serán debidamente consultados con docentes expertos en el EAT.

Antes de comenzar con su entrenamiento, el usuario podrá decidir un número máximo de ejercicios nuevos por día. Al igual que en otros programas como Anki¹⁶, este número determinará la cantidad de

14 Teruki Kaniwa *et al.*, «Towards a Personalized Technical Ear Training Program: An Investigation of the Effect of Adaptive Feedback», *Proceedings of the SMC* (2011): 439–443.

15 Hadi Amiri, Tim Miller y Guergana Savova, «Repeat Before Forgetting: Spaced Repetition for Efficient and Effective Training of Neural Networks», *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing* (2017): 2401–2410.

16 Anki, «Background».

nuevas preguntas que el estudiante deberá reconocer cada 24 horas. La facilidad o dificultad con la que el estudiante responda determinará la frecuencia y la cantidad de repeticiones.

Una vez comenzado el entrenamiento, el estudiante escuchará un audio con un aumento o corte de frecuencia y deberá indicar la banda en la que se encuentra la afectación. Los docentes podrán decidir si sus clases tendrán disponible la opción de *bypass* para realizar una comparación A/B, o si los estudiantes primero escucharán la versión filtrada y luego el audio limpio. La misma aplicación ofrecerá flexibilidad para los docentes y para sus clases.

En el momento del entrenamiento, el estudiante deberá responder una secuencia de preguntas. Sin embargo, la aplicación en el fondo no solo estará evaluando las preguntas individuales, sino que también estará revisando el desempeño del estudiante en cada banda de frecuencia y considerando su contexto, es decir, las preguntas previas. De esta manera, a medida que avanza el entrenamiento, el algoritmo de repetición espaciada va definiendo cuáles secuencias de bandas de frecuencia le generan más problemas al estudiante, y se las va presentando con cierta regularidad para poder realizar un análisis más efectivo.

Al finalizar la prueba, el estudiante podrá ver sus resultados, así como un análisis comparativo de sus resultados a través del tiempo, información discriminada por bandas de frecuencia sobre su desempeño, y secuencias de bandas en las que tuvo un mayor porcentaje de error.

Por último, esta información se actualizará automáticamente en la base de datos en línea, para que los docentes puedan consultar el desempeño de cada estudiante, así como el desempeño general de su clase, y puedan enfocarse en las preguntas que más problemas les generen a los estudiantes de la clase.

Con este uso, los estudiantes podrán tener acceso a una herramienta más personalizada y clara y monitorear de cerca su progreso, sus fortalezas y sus debilidades. Por su parte, los docentes podrán evaluar el progreso de sus estudiantes con mayor facilidad, calificar a sus estudiantes de acuerdo con su progreso individual, y tener mayor control sobre las metodologías de estudio que sus estudiantes utilizan, para asegurarse de que todos estén entrenando con una metodología unificada.

6. Conclusiones

Estas son solo algunas de las ideas fundamentales que se podrían explorar en un futuro *software* para el EAT en las escuelas y universidades.

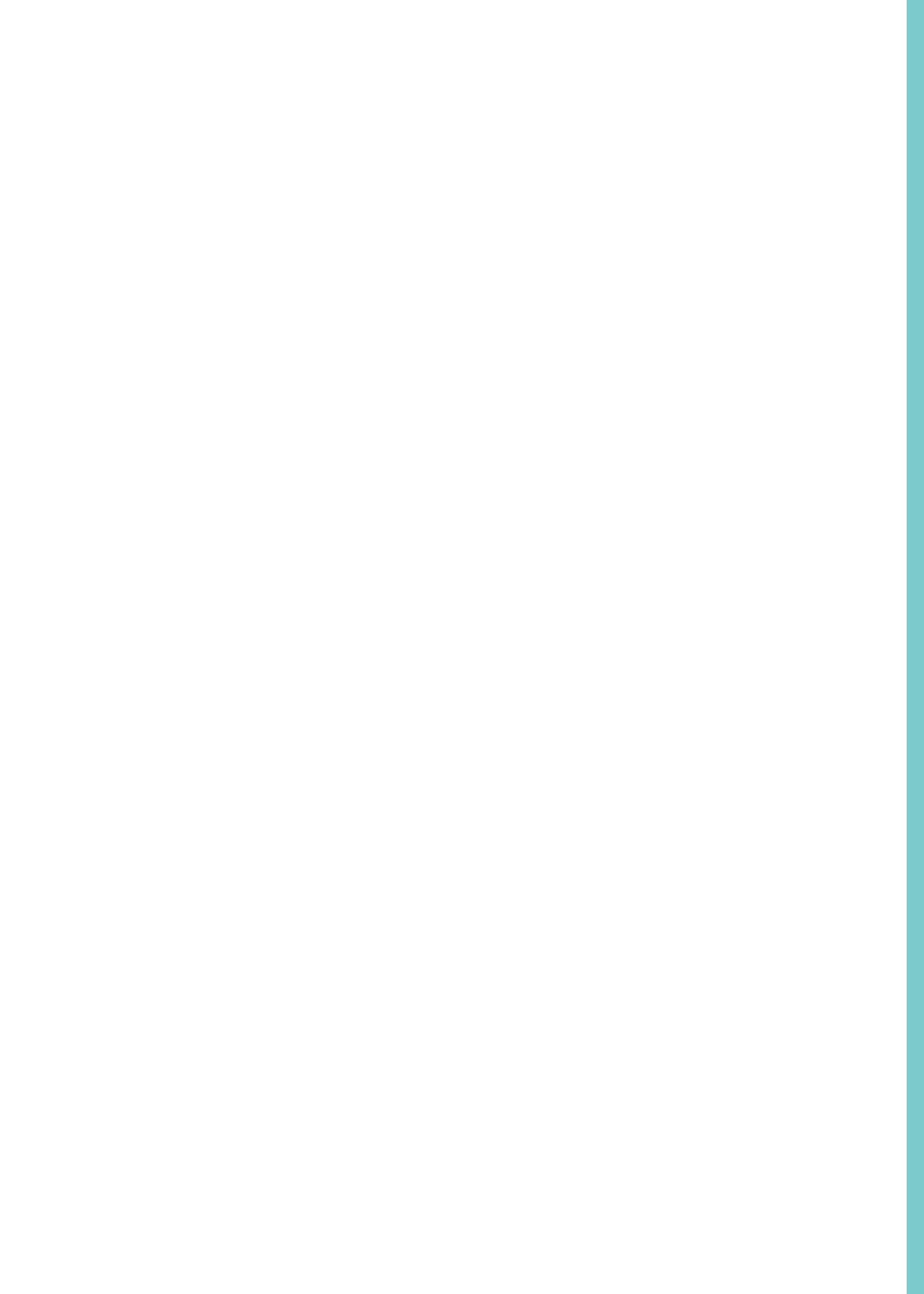
Se necesita priorizar la eficiencia y personalización en las herramientas, la mejor comunicación de información y análisis estadístico, la estandarización en los contenidos de las asignaturas, y los algoritmos de selección de ejercicios para minimizar la frustración y dificultades en los estudiantes, y facilitar el proceso de evaluación y planeación para los docentes.

Muchas de estas dificultades pueden abordarse a través de mejores herramientas diseñadas específicamente para asistir a los docentes y personalizar el proceso de los estudiantes.

En investigaciones futuras se esperan hacer pruebas reales con grupos de estudiantes y *software* desarrollado siguiendo estos parámetros para evaluar distintos desempeños con distintos algoritmos, así como otras opciones para seguir optimizando y estandarizando la enseñanza de la escucha crítica.

Referencias

- Amiri, Hadi, Tim Miller y Guergana Savova. «Repeat Before Forgetting: Spaced Repetition for Efficient and Effective Training of Neural Networks». *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing* (2017): 2401-2410.
- Anki. «Background». Anki, 2006. <https://docs.ankiweb.net/background.html>
- Ausubel, David y Mohamed Youssef. «The Effect of Spaced Repetition on Meaningful Retention». *The Journal of General Psychology*, vol. 73, n.º 1 (1965): 147-150.
- Corey, Jason. «Technical Ear Training: Tools and Practical Methods». *Proceedings of Meetings on Acoustics*, vol. 19, n.º 1 (2013).
- . *Audio Production and Critical Listening: Technical Ear Training*. Taylor & Francis, 2016.
- Gordon, Wellington. «Considering Best Practices in Technical Ear Training». *Journal of the Association for Technology in Music Instruction*, vol. 2, n.º 1, (2021).
- Haas, Andreas, Andreas Rossberg, Derek L. Schuff, Ben L. Titzer, Michael Holman, Dan Gohman, Luke Wagner, Alon Zakai y J. F. Bastien. «Bringing the web up to speed with WebAssembly». *Proceedings of the 38th ACM SIGPLAN Conference on Programming Language Design and Implementation* (2017): 185-200.
- iZotope. «iZotope announces PAE». iZotope, 2016. <https://www.izotope.com/en/company/press/2016/izotope-announces-pae.html>
- Kaniwa, Teruki, Sungyoung Kim, Hiroko Terasawa, Masahiro Ikeda, Takeshi Yamada y Shoji Makino. «Towards a Personalized Technical Ear Training Program: An Investigation of the Effect of Adaptive Feedback». *Proceedings of the SMC* (2011): 439-443.
- Kim, Sungyoung. «An Assessment of Individualized Technical Ear Training for Audio Production». *The Journal of the Acoustical Society of America*, vol. 138, n.º 1 (2015).
- Puckette, Miller S. «Pure Data». ICMC (septiembre de 1997).
- Quesnel, René. «A computer-assisted method for training and researching timbre memory and evaluation skills». Tesis de doctorado, Faculty of Music, McGill University, 2001.
- SoundGym. «Audio ear training & learning center». SoundGym, 2016. <https://www.soundgym.co/site/about>
- Tabibian, Behzad, Utkarsh Upadhyay, Abir De, Ali Zarezade, Bernhard Schölkopf y Manuel Gómez-Rodríguez. «Enhancing Human Learning Via Spaced Repetition Optimization». *Proceedings of the National Academy of Sciences*, vol. 116, n.º 10 (2019): 3988-3993.
- Ye, Junyao, Jingyong Su y Yilong Cao. «A Stochastic Shortest Path Algorithm for Optimizing Spaced Repetition Scheduling». *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining* (2022): 4381-4390.





Vol. 3 N.º 9 / Guayaquil
mayo 2024
ISSN 2953-657x

La sonificación de datos como herramienta tecnológica para la creación artística

Diego Alejandro Molina Quintero

Facultad de Artes y Humanidades
Instituto Tecnológico Metropolitano
Medellín, Colombia
diegomolina4406@correo.itm.edu.co

Isabella Álvarez Franco

Facultad de Artes y Humanidades
Instituto Tecnológico Metropolitano
Medellín, Colombia
isabellaalvarez251817@correo.itm.edu.co

Lina Alejandra González Ramírez

Facultad de Artes y Humanidades
Instituto Tecnológico Metropolitano
Medellín, Colombia
linagonzalez305094@correo.itm.edu.co

Jamir Mauricio Moreno Espinal

Facultad de Artes y Humanidades
Instituto Tecnológico Metropolitano
Medellín, Colombia
mauriciomoreno@itm.edu.co

Resumen

La sonificación de datos (en adelante SD) es una técnica cada vez más utilizada en diversos campos para convertir bases de datos en sonidos y permitir su interpretación auditiva. Para tal propósito, en este estudio se utilizaron los softwares TwoTone, SuperCollider y Sonic Visualiser. De esa forma, el artista puede elegir la herramienta de creación más adecuada a sus intereses y capacidades para lograr las experiencias sonoras deseadas.

Esta investigación aplicó los métodos analítico y comparativo, ya que examina en detalle la forma de uso de las herramientas seleccionadas, su proceso de instalación y ejecución, y explora la personalización de los sonidos y su integración con dispositivos externos como, por ejemplo, placas de desarrollo similares a Arduino o Raspberry Pi. Adicionalmente, realiza un análisis comparativo de las estadísticas de las contribuciones realizadas por la comunidad en GitHub para cada una de estas herramientas.

Los resultados obtenidos pueden tener aplicaciones en campos como la música, la ciencia de datos y la investigación en sonido. Además, pueden ser de gran utilidad para los usuarios interesados en realizar sonificaciones de manera creativa y personalizada.

Es importante destacar que esta investigación se centró en explorar la generación de sonidos simples y básicos en cada plataforma evaluada. Sin embargo, se debe recordar que es posible que la elección de la herramienta de software para la sonificación de datos dependa en gran medida del nivel de conocimiento del usuario en áreas como la programación, la música y el manejo de datos. Estos factores influyen en la capacidad del usuario para aprovechar al máximo las herramientas disponibles y personalizar los resultados según sus necesidades específicas. Para aquellos con un nivel bajo de conocimiento, TwoTone y Sonic Visualiser pueden ser una buena opción, debido a su intuitiva interfaz de usuario. Por su parte, SuperCollider es una herramienta poderosa para usuarios con conocimientos medios y altos en programación, pero su curva de aprendizaje es empinada. Respecto a la personalización, SuperCollider y TwoTone ofrecen una amplia variedad de opciones, mientras que Sonic Visualiser se enfoca principalmente en el análisis y la visualización de datos de audio. Cada herramienta tiene, pues, sus limitaciones y fortalezas en relación con la generación y la personalización de sonidos.

Palabras clave: sonificación, programación creativa, TwoTone, SuperCollider, Sonic Visualiser

Abstract

Data sonification, hereafter referred to as SD, is an increasingly utilized technique in various fields to convert databases into sounds and enable auditory interpretation. In this study, the software tools TwoTone, SuperCollider, and Sonic Visualiser were employed for this purpose. In this way, artists can choose the creation tool most appropriate to their interests and abilities to achieve the desired sound experiences.

This research employed analytical and comparative methods to examine in detail the usage of each selected tool, including their installation and execution processes, as well as exploring sound customization and integration with external devices such as Arduino or Raspberry Pi development boards. Additionally, a comparative analysis of contribution statistics from the GitHub community was conducted for each of these tools.

The results have potential applications in fields like music, data science, and sound research, offering significant value to users interested in creatively and personally generating sonifications.

The choice of software tool for sonifications depends on the user's proficiency in programming, music, and data. For those with limited knowledge, TwoTone and Sonic Visualiser are recommended due to their intuitive user interface. On the other hand, SuperCollider is a powerful tool suitable for users with intermediate to advanced programming skills, despite its steep learning curve. In terms of customization, SuperCollider and TwoTone provide a wide range of options, while Sonic Visualiser primarily focuses on audio data analysis and visualization. Therefore, each tool has its own limitations and strengths concerning sound generation and customization.

Keywords: sonification, creative programming, TwoTone, SuperCollider, Sonic Visualiser

1. Introducción

La sonificación es una técnica que permite transformar información en sonidos.¹ A pesar de que no es tan conocida ni se practica tanto como la visualización gráfica de datos, el sonido, en todo caso, es un medio naturalmente afectivo, estético y cultural.² Aunque parezca una idea moderna, la humanidad lo ha utilizado desde siglos atrás para interpretar información; un ejemplo de ello son las campanas de las iglesias para marcar el tiempo y convocar a la congregación, sin dejar a un lado su papel para señalar el horario de las actividades seculares.³

El sonido también se ha usado para representar datos en campos como la detección de la radiación, la distancia de los objetos en el mar, la altura de los aviones o las afectaciones en la salud. Es así como el contador Geiger, un dispositivo que utiliza un tubo Geiger-Müller para medir la cantidad de partículas ionizantes en el aire⁴ produce, según su cantidad, un sonido que advierte que su concentración es alta y puede ser peligrosa para el ser humano.

Otro ejemplo destacado del uso del sonido para representar datos es el optófono, creado en 1913 para facilitarles la lectura a las personas invidentes. Actualmente, la tecnología ha permitido el desarrollo de múltiples dispositivos de este tipo para individuos con discapacidad visual, a través

1 David Rodríguez Mateos y Alicia Tapia López, «Sonificación y periodismo: la representación de datos mediante sonidos», *Revista de Comunicación*, vol. 22, n.º 1 (2023): 413–433, <http://dx.doi.org/10.26441/rc22.1-2023-3022>

2 Paul Vickers y Bennet Hogg, «Sonification abstraite / Sonification concrète: An aesthetic perspective space for classifying auditory displays in the Ars Musica domain», en *Proceedings of the 12th International Conference on Auditory Display*, (210–216) (Londres, 2006). <https://typeset.io/pdf/sonification-abstraite-sonification-concrete-an-aesthetic-1urt1pe0ql.pdf>

3 Anne Staples, «El abuso de las campanas en el siglo pasado», *Historia Mexicana*, vol. 27, n.º 2 (1977): 177–194, <https://www.jstor.org/stable/25135599>

4 El tubo Geiger-Müller es un tubo con gas que contiene un electrodo positivo y uno negativo. Leonardo Ramírez Arias y Julián Uribe Cortés, «El contador Geiger como herramienta en la enseñanza del concepto de radiación», *Góndola, Enseñanza y Aprendizaje de las Ciencias*, vol. 2, n.º 1 (2007): 33–36, <https://revistas.udistrital.edu.co/index.php/GDLA/article/download/5304/6935/0>

de tecnologías como la OCR (*optical character recognition*) y el TTS (*text-to-speech*).⁵

Entre 1929 y 1930 surgieron herramientas como el radioaltímetro y el sonar. El primero utiliza el radar para emitir pulsos de ondas hacia el suelo,⁶ y el segundo se usa como un medio de localización acústica de forma similar al radar. Sin embargo, este último, en lugar de emitir ondas electromagnéticas, emplea impulsos sonoros.⁷ Cabe anotar que, anterior a la aparición del GPS (Global Positioning System), la localización acústica se había usado en el aire para la investigación atmosférica⁸ a través de la aplicación del Sodar (Sonic Detection and Ranging, exploración vertical aérea con el sonar). Con esta tecnología, la interpretación de los datos verticales podía guiar a las aeronaves emitiendo un sonido de alerta si frente a ellas había montañas u obstáculos amenazadores de su seguridad.

En 1940 se desarrolló el ecógrafo, que utiliza el ultrasonido para visualizar estructuras tan pequeñas que escapan a la detección del ojo humano.⁹ Este dispositivo ha sido utilizado en la medicina durante muchos años, facilitando la auscultación de los ruidos cardíacos o respiratorios, así como para objetivar ruidos intestinales o soplos por flujos anómalos sanguíneos en venas y arterias.¹⁰

5 I. A. J. Gárate Piñones, «Doroti» (tesis doctoral, Universidad Andrés Bello, 2020).

6 Bryan Marcelo Peralta Rocha, «Inspección y pintura del avión Escuela Hawker Siddeley 125-400, mediante información y datos técnicos para la unidad de gestión de tecnologías UGT-ESPE» (tesis de pregrado, Universidad de las Fuerzas Armadas, Sangolquí, Ecuador, 2018), <https://repositorio.espe.edu.ec/bitstream/21000/27745/1/M-ESPEL-CMA-0478.pdf>

7 «Sonar», Wikipedia, 21 de marzo de 2019, <https://es.wikipedia.org/wiki/Sonar>

8 Andrea Vicenta Remiro Gargallo, «Técnicas numéricas aplicadas a la propagación de sonido en el medio marino» (tesis doctoral, Universitat Politècnica de València, España, 2019), <http://hdl.handle.net/10251/128636>

9 Andrea del Pilar Sánchez Corzo, «Sonorización de señales de EEG basada en estructuras musicales» (tesis de pregrado, Universidad de los Andes, Bogotá, 2012), <https://repositorio.uniandes.edu.co/bitstream/handle/1992/25182/u627972.pdf?sequence=1&isAllowed=y>

10 Fernando Saraví, «Laennec, el método anamoclinico y la invención del estetoscopio. Parte 1: De Bretaña a París, hasta la gran invención», *Revista Médica Universitaria*, vol. 8, n.º 2 (2012), https://bdigital.uncu.edu.ar/objetos_digitales/4769/rmuhistoriasdelamedicina.pdf

Por otro lado, el estetoscopio es un dispositivo que ha ayudado a la medicina por largo tiempo, facilitando la escucha interna del cuerpo.¹¹ Este instrumento se utiliza de la misma manera que el ecógrafo.¹²

Finalmente, los detectores de metales, que emiten sonidos cuando localizan metales a diferentes profundidades, tienen una gran variedad de aplicaciones en la industria actual tanto en materia de seguridad como en el ocio.¹³

En resumen, la sonificación, una técnica utilizada por la humanidad desde hace siglos, permite transformar información en sonidos. A lo largo de la historia se han desarrollado dispositivos que utilizan el sonido para representar datos en campos como la medicina, la detección de la radiación o la exploración aérea o marina. En la actualidad, gracias a la tecnología, se han desarrollado múltiples dispositivos que utilizan el sonido para facilitar la vida de las personas.

Este documento profundiza en aspectos técnicos de las herramientas de software evaluadas, como la capacidad de SuperCollider para la síntesis de sonidos complejos y su curva de aprendizaje más empinada, o las interfaces intuitivas y fáciles de usar de TwoTone y Sonic Visualiser, que las hacen más accesibles para aquellos usuarios con menores conocimientos en programación y música. Además, se pueden explorar las limitaciones y fortalezas de cada herramienta en términos de su procesamiento de las señales y el análisis y la visualización de los datos de audio, lo que puede afectar la calidad y la complejidad del sonido generado.

2. Perspectiva teórica

La importancia del sonido para el ser humano es indiscutible. Más allá de la vista, el

11 «Estetoscopio», Quimica.es, <https://www.quimica.es/enciclopedia/Estetoscopio.html>

12 Saraví, «Laennec, el método...».

13 Carlos Martínez Cantón, «Diseño y construcción de un detector de metales basado en un microcontrolador» (tesis de pregrado, Universitat Politècnica de Catalunya, España, 2016), <https://upcommons.upc.edu/handle/2117/98926?locale=attribute=es>

olfato, el tacto y el gusto, la audición es, sin duda, una de las funciones vitales para sobrevivir e interactuar con el mundo exterior. El oído tiene la capacidad de transformar las ondas físicas en impulsos nerviosos que llegan al cerebro, permitiéndole a las personas interactuar con el mundo desde temprana edad y desarrollar diferentes habilidades. El oído es esencial, ya que la humanidad se desenvuelve en un paisaje sonoro que le permite desarrollar diferentes tipos de habilidades.¹⁴

Los sonidos tienen un papel clave en la experiencia sensorial y emocional del mundo que rodea al ser humano, permitiéndole distinguir diferentes timbres y procesarlos de forma fluida. De hecho, los sonidos tienen un rango de intensidades mucho mayor que las detectadas por la vista humana para percibir diferentes niveles de luz¹⁵ y, por tanto, lo influyen constantemente y generan diferentes reacciones en su cerebro.

A través del oído, el humano es capaz de procesar diferentes tipos de sonidos, interpretarlos y reaccionar ante ellos, lo que le brinda una amplia variedad de habilidades y le permite tener una experiencia sensorial y emocional más completa. Además, la SD le permite traducir la información en sonidos para una mejor comprensión auditiva que pueda aplicarse en áreas como la seguridad, la monitorización y el arte, esta última, en particular, por su valor estético y artístico. La importancia del sonido en la vida de la humanidad es indudable, por lo que es fundamental seguir valorando y explorando esta capacidad sensorial para mejorar la comprensión y la experiencia del mundo que la rodea.

A continuación, se presenta un listado de los programas que permiten la SD.

14 Pierre Schaeffer, Christine North y John Dack, *Treatise on Musical Objects: An Essay Across Disciplines* (California: University of California Press, 2017), <https://www.jstor.org/stable/10.1525/j.ctt1qv5pb>

15 Alina Castillo Vázquez, Diedmar Céspedes Cabrera, Mercedes Jiménez Betancourt y Gustavo Rodríguez Rosabal, «Alternativa didáctica para el estudio de los órganos sensitivos en la enseñanza de la Biología 3», *Revista Dilemas Contemporáneos: Educación, Política y Valores*, vol. 8, n.º 1 (2020), <https://dilemascontemporaneoseducacionpoliticaayvalores.com/index.php/dilemas/article/view/2386>

- TwoTone: *software* gratuito y de código abierto que se ejecuta totalmente en la red y no requiere ser descargado. TwoTone permite convertir datos en sonidos y música a través del proceso de sonificación, y es compatible con computadoras de escritorio, tabletas y teléfonos móviles. Además, es una herramienta útil tanto para la comprensión de datos como para la creación de música, sin necesidad de conocimientos musicales o tecnológicos previos.¹⁶
- Pure Data (Pd): lenguaje de programación gráfico desarrollado por Miller Puckette para la creación de música interactiva y obras multimedia. Pd es un entorno de programación visual de código abierto que se puede ejecutar en una variedad de dispositivos, desde computadoras personales hasta unidades integradas y teléfonos inteligentes. Este *software* permite a músicos, artistas visuales, investigadores y desarrolladores crear *software* gráficamente sin tener que escribir líneas de código. Pd se puede utilizar para procesar y generar sonido, video, gráficos en 2D y 3D, sensores de interfaz, dispositivos de entrada y MIDI (Musical Instrument Digital Interface).¹⁷
- SuperCollider: *software* libre desarrollado por James McCartney en 1996, que presenta tres componentes principales: un servidor de audio en tiempo real, un lenguaje de programación interpretado y un editor de scLang con un sistema de ayuda integrado. SuperCollider es adecuado para aprender métodos básicos de procesamiento multimedia y programación visual, así como para realizar sistemas complejos para proyectos a gran escala.¹⁸
- Dear Diary ai: sitio web que permite a los usuarios traducir lo que se escribe en sonidos. Es una herramienta útil para las personas que deseen experimentar con sonidos a partir de textos.¹⁹
- Sonic Visualiser: *software* libre y de código abierto que permite a los usuarios la visualización y el análisis de archivos de sonido. Adicionalmente, ofrece funciones de sonificación que hacen posible convertir los datos de un archivo de sonido en un sonido audible.²⁰
- Sonification Sandbox: plataforma en línea que permite a los usuarios cargar y sonificar datos de diferentes tipos. Adicionalmente, cuenta con herramientas de análisis de sonido y facilita la descarga de archivos de audio.²¹
- AudioSpray: *software* que permite a los usuarios la creación de sonidos y música a partir de imágenes. Es una herramienta útil para la creación de paisajes sonoros y efectos de sonido.²²
- Sonic Pi: *software* libre y gratuito que permite a los usuarios crear música electrónica mediante programación. El programa cuenta con un entorno de desarrollo interactivo en el que pueden escribir código en lenguaje Ruby para crear y manipular sonidos en tiempo real. Además,

16 Gonzalo de la Vega, Leonardo Martín Exequiel Domínguez, Johanna Casado, Beatriz García, «SonoUno Web: An innovative user centred web interface», *Arxiv* (2022), <https://arxiv.org/pdf/2302.00081.pdf>

17 Francisco Javier Ruiz-del-Olmo y José Vertedor, «Procesos, herramientas y prácticas de la sonificación», Málaga, España, Universidad de Málaga, 2016, <http://hdl.handle.net/10630/12429>

18 Scott Wilson, David Cottle y Nick Collins, *The SuperCollider Book* (Cambridge, MA: The MIT Press, 2011).

19 Hengbo Wang, «Dear diary: an investigation of emotionally engaging computer-game narratives» (tesis de maestría, Auckland University of Technology, Australia, 2015), <http://hdl.handle.net/10292/9137>

20 Chris Canam, Christian Landone, Mark Sandler y Juan Pablo Bello, «The sonic visualiser: a visualisation platform for semantic descriptors from musical signals», *ISMIR 2006 7th International Conference on Music Information Retrieval*, 324–327 (Victoria, BC, Canadá: 2006), <https://nyuscholars.nyu.edu/en/publications/the-sonic-visualiser-a-visualisation-platform-for-semantic-descri>

21 Bruce N. Walker y Joshua T. Cothran, «Sonification sandbox: a graphical toolkit for auditory graphs», *Proceedings of the 2003 International Conference on Auditory Display* (Atlanta, GA: Georgia Institute of Technology, 2003).

22 Richard Garrett, «AudioSprayGun0.8-The generation of large sound-groups and their use in three-dimensional spatialisation», *ICMC* (2015), <https://quod.lib.umich.edu/cgi/p/pod/dod-idx/audio-spray-gun-08-the-generationof-large-sound-groupspdf?c=icmc;idno=-bbp2372.2015.071;format=pdf>

incluye una amplia variedad de muestras y sonidos predefinidos para facilitar el proceso de creación musical.²³

- Max/MSP: entorno de programación visual para la creación de música, sonido y multimedia. Max/MSP es ampliamente utilizado por músicos, artistas visuales, diseñadores de sonido y desarrolladores de software para crear instalaciones interactivas, performances y composiciones multimedia. El programa permite a los usuarios crear sus propios objetos y algoritmos personalizados para procesar y generar sonidos, video, gráficos y otros tipos de datos.²⁴

- Csound: lenguaje de programación para la síntesis de sonido y el procesamiento de señales de audio. Csound ha sido utilizado por músicos, compositores y artistas sonoros para crear una amplia variedad de obras musicales y experimentales. El programa permite a los usuarios definir y manipular múltiples canales de audio, crear efectos de procesamiento de señales y controlar dispositivos externos a través de MIDI.²⁵

23 Aaron, S., «Sonic Pi-The live coding music synth for everyone», Sonic Pi, 2018, <https://sonic-pi.net/>

24 Daniel Edward Barsetti-Nerland, «Creating a responsive visualization that reacts with music in real time: integrating Ableton Live 9 and Cycling'74 max for live into a musical performance» (tesis de pregrado, University of Arizona, Tucson, 2016), <https://repository.arizona.edu/handle/10150/612546>

25 B. Vercoe et al., «The canonical Csound reference manual», Csound, 2007, <https://csound.com/docs/manual/index.html>

- ChucK: lenguaje de programación de audio desarrollado por Ge Wang y Perry Cook para la creación de música en tiempo real y multimedia interactiva. ChucK es un lenguaje de programación textual que permite a los usuarios crear y manipular sonidos utilizando objetos y sintaxis especializados. Además, cuenta con herramientas de análisis y síntesis de sonido que facultan la sonificación de datos y la generación de diferentes tipos de sonidos y efectos.²⁶

La tabla 1 muestra un resumen de los programas SD expuestos.

Tabla I
Programas de sonificación de datos

Software	Sistemas operativos	Github	Creador
TwoTone	Windows, Mac, Linux	https://github.com/sonifydata/twotone	Paul Batchelor
Pure Data (Pd)	Windows, Mac, Linux	https://github.com/pure-data/pure-data	Miller Puckette
Super Collider	Windows, Mac, Linux	https://github.com/supercollider/super-collider	James McCartney
Dear Diary ai	Windows, Mac, Linux	https://github.com/frederick0329/Dear-diary-ai	frederick 0329 (usuario de GitHub)
Sonic Visualiser	Windows, Mac, Linux	https://github.com/svi-project/sonic-visualiser	Chris Cannam
Sonification Sandbox	Windows, Mac	https://github.com/Ching-Chieh/Sonification-Sandbox	Ching-Chieh Cheng
SoniPy	Windows, Mac, Linux	https://github.com/carlthome/python-sonify	Carl Thomé
Audio Spray	Windows	https://github.com/flexmonkey/audiospray	Matt Bellingham
Sonic Pi	Windows, Mac, Linux	https://github.com/samaaron/sonic-pi	Sam Aaron
Max/MSP	Windows, Mac	https://github.com/Cycling74/max-sdk	Miller Puckette
Csound	Windows, Mac, Linux	https://github.com/csound/csound	Barry Vercoe
ChucK	Windows, Mac, Linux	https://github.com/ccrma/chuck	Ge Wang, Perry Cook et al.

Fuente: elaboración de los autores.

26 Ge Wang, Dan Trueman, Scott Smallwood y Perry R. Cook, «The laptop orchestra as classroom», *Computer Music Journal*, vol. 32, n.º 1 (2008): 26–37, <https://www.jstor.org/stable/40072662>

3. Perspectiva metodológica

La metodología empleada en el artículo proporciona una descripción detallada de los procedimientos y pasos seguidos para evaluar y comparar tres herramientas de sonificación de datos: TwoTone, SuperCollider y Sonic Visualiser.

3.1. Fases de la investigación

La investigación se llevó a cabo en varias fases claramente definidas:

- Selección del *software*: se identificaron y seleccionaron las herramientas de sonificación de datos que se estudiarían. Se describieron criterios específicos para la elección de estas herramientas, como la disponibilidad en múltiples sistemas operativos y la accesibilidad del código fuente en GitHub.
- Instalación y configuración: se detallaron los procedimientos de instalación de cada herramienta en un entorno operativo basado en Ubuntu 22.04.2 LTS x86_64. Esto incluyó los comandos y pasos necesarios para la instalación de cada herramienta.
- Caso de uso y personalización: se proporcionó un ejemplo práctico de cómo utilizar cada herramienta para generar sonidos a partir de una base de datos específica. Se explicó cómo personalizar y modificar los sonidos según las necesidades del usuario.

Esta sección presenta la descripción detallada de la implementación de tres herramientas de SD: TwoTone, SuperCollider y Sonic Visualiser.

Dichas herramientas fueron instaladas en un entorno operativo basado en Ubuntu 22.04.2 LTS x86_64, ejecutado

en un equipo modelo Katana GF76 12UC REV:1.0.

El equipo cuenta con las siguientes especificaciones técnicas: procesador de 12.^a generación Intel i7-12700H y una unidad de procesamiento gráfico NVIDIA GeForce RTX 3050 Mobile, y opera con el Kernel 5.19.0-35-generic y la shell Zsh 5.8.1. Adicionalmente, se llevó a cabo una revisión exhaustiva de las principales características y funcionalidades de cada plataforma, usando el método analítico combinado con el método comparativo.

3.1.1. Selección del software

La selección de las herramientas de sonificación de datos desempeña un papel fundamental en cualquier investigación que involucra la creación de sonidos a partir de datos. Para garantizar la idoneidad de las herramientas elegidas y la validez de las evaluaciones realizadas, se establecieron criterios específicos de selección. Estos criterios se basaron en consideraciones clave que afectarían tanto la accesibilidad como la calidad del análisis. Los criterios utilizados para seleccionar las herramientas incluyeron:

- Disponibilidad en múltiples sistemas operativos: se priorizó la selección de herramientas que estuvieran disponibles en una variedad de sistemas operativos ampliamente utilizados, como Windows, Mac y Linux. Esto se hizo con el objetivo de garantizar que las herramientas fueran accesibles para un público diverso y no limitar su utilidad a un sistema operativo específico.
- Accesibilidad del código fuente: se consideró esencial que las herramientas seleccionadas fueran de código abierto y que su código fuente estuviera disponible públicamente a

- través de plataformas como GitHub. Esto facilitaría la transparencia y la capacidad de examinar y evaluar el funcionamiento interno de las herramientas, lo que es esencial para la investigación y la evaluación objetiva.
- Comunidad y documentación activa: se evaluó la presencia de una comunidad activa de usuarios y desarrolladores en torno a cada herramienta. Una comunidad activa es indicativa de un soporte continuo, corrección de errores y desarrollo en curso. Además, se consideró la disponibilidad de documentación detallada y recursos de aprendizaje que facilitaran la comprensión y el uso de las herramientas.
 - Estabilidad y madurez: se prestó atención a la estabilidad y madurez de las herramientas, lo que incluyó la evaluación de la frecuencia de actualizaciones, la ausencia de problemas críticos conocidos y la duración de su existencia en el ámbito de la sonificación de datos.
 - Funcionalidad relevante: se tuvo en cuenta la funcionalidad específica de sonificación de datos de cada herramienta. Esto incluyó la capacidad de generar sonidos a partir de datos, así como la flexibilidad y versatilidad en la manipulación de estos sonidos para adaptarse a diferentes contextos de investigación.
 - Licencia de software: se verificó que las herramientas seleccionadas utilizaran licencias de software compatibles con los objetivos de la investigación y la publicación de resultados. Las licencias de código abierto o de uso gratuito fueron preferidas para promover la accesibilidad y el intercambio de conocimientos.

- Aplicabilidad a la investigación: se consideró si las herramientas eran adecuadas para los objetivos específicos de la investigación, lo que incluyó la capacidad de trabajar con los tipos de datos y bases de datos utilizados en el estudio.

La aplicación de estos criterios aseguró que las herramientas seleccionadas fueran apropiadas y justas para la comparación y evaluación en el contexto de la sonificación de datos. Además, proporcionó una base sólida para la objetividad y la replicabilidad de los resultados de la investigación.

3.1.2. Instalación

TwoTone

Cabe recordar que esta aplicación puede usarse en línea desde cualquier navegador (<https://twotone.io/>), aunque también permite descargar el código y emplearlo *offline* de la siguiente forma:

- Clonar el proyecto en un directorio de la computadora: git clone <https://github.com/datavized/twotone.git>
- Ingresar a la carpeta en la que se acaba de clonar desde la terminal usando el siguiente comando: cd twotone
- Instalar las dependencias necesarias usando el siguiente comando: npm instal.
- Para poder ejecutarlo en modo de desarrollo, se necesita configurar un servidor web local en la computadora. Esto permitirá visualizar la aplicación en el navegador web del usuario sin necesidad de publicarla en un servidor remoto. Además, cuando se realicen cambios en los archivos de código fuente, la aplicación se reconstruirá

- automáticamente para reflejarlos, sin que haya que reiniciar manualmente el servidor; el comando es el siguiente: npm run star
- Para compilar una versión de producción del proyecto, se puede utilizar el siguiente comando: npm run build

SuperCollider

Para la instalación estándar en Ubuntu, se deben ejecutar los siguientes comandos de la terminal, una línea a la vez:

```
sudo add-apt-repository ppa:super-
collider/ppa
sudo apt-get update
sudo apt-get install supercollider
```

Para enviar los datos que se van a procesar y enviarlos a SuperCollider, se utiliza Python. Es importante tener en cuenta las siguientes consideraciones al utilizar Python en Linux:

- Crear un entorno virtual. Un entorno virtual es un espacio donde se pueden instalar paquetes específicos para un proyecto. Esto significa que es posible tener un conjunto de paquetes/librerías aislados de la instalación principal de Python en el sistema para un proyecto en particular.
- Abrir la terminal en Ubuntu Linux e instalar pip (el administrador de paquetes de Python) en el sistema. Si aún no lo está, puede hacerse con el siguiente comando: sudo apt-get install python3-pip
- Instalar virtualenv, una herramienta para crear entornos virtuales de Python, a través del siguiente comando: sudo pip3 install virtualenv
- Crear un nuevo directorio para el entorno virtual, que será su directorio

raíz. Se puede nombrar de cualquier forma, por ejemplo: «mi_entorno_virtual», a través del siguiente comando: mkdir mi_entorno_virtual

- Entrar en el directorio recién creado utilizando el siguiente comando: cd mi_entorno_virtual
- Crear un entorno virtual dentro de este directorio. Se puede nombrar de cualquier forma, por ejemplo: «venv», a través del siguiente comando: virtualenv venv
- Activar el entorno virtual utilizando el siguiente comando: source venv/bin/activate

Ahora, el usuario ya está dentro del entorno virtual y debe instalar el paquete de Python que va a usar utilizando pip a través del siguiente comando: pip install numpy y pip install python-osc

Una vez que SuperCollider esté instalado, puede ejecutarse en la terminal ingresando el siguiente comando: sclang. Esta acción iniciará el intérprete de SuperCollider en la terminal.

Sonic Visualiser

1. Abrir la terminal de Ubuntu y actualizar la lista de paquetes con el siguiente comando: sudo apt-get update
2. Instalar Sonic Visualiser con el siguiente comando: sudo apt-get install sonic-visualiser
3. Esperar a que el proceso de instalación termine. Si se solicita confirmación, presionar «y» y «Enter» para continuar. Una vez que la instalación esté completa, el programa puede abrirse desde el menú de aplicaciones de Ubuntu o desde la terminal, a través del siguiente comando: sonic-visualiser

Tabla II
Aspectos técnicos de instalación, lenguaje, navegabilidad y tipo de licencia de las herramientas seleccionadas

Aspecto	TwoTone	SuperCollider	Sonic Visualiser
Sistema operativo	Windows, Mac, Linux	Windows, Mac, Linux	Windows, Mac, Linux
Lenguajes de programación	JavaScript	C++, Super Collider language	C++, Python
Interfaz gráfica	Sí	No (interfaz de línea de comandos)	Sí
Navegabilidad web	Sí (se puede ejecutar en un navegador web)	No	No
Tipo de licencia	Licencia MIT (software libre y de código abierto)	Licencia GPLv2 (software libre y de código abierto)	Licencia GPL (software libre y de código abierto)

Fuente: elaboración de los autores.

4. Resultados y discusión

TwoTone

El resultado puede escucharse en el siguiente enlace: <https://soundcloud.com/acorde-semillero/twotone>

SuperCollider

El resultado puede escucharse en el siguiente enlace: <https://soundcloud.com/acorde-semillero/supercollider>

Sonic Visualiser

El resultado puede escucharse en el siguiente enlace: <https://soundcloud.com/acorde-semillero/sonicvisualiser>

Esta sección desempeña un papel crucial en el análisis y comprensión de las herramientas seleccionadas para la sonificación de datos. Aquí presentamos un estudio de caso detallado para cada una de estas herramientas, proporcionando una visión práctica y detallada de su aplicación en este campo.

La inclusión de los casos de uso es esencial, ya que proporciona una comprensión completa y profunda de cómo estas herramientas se aplican en la práctica. Esto establece las bases para una evalua-

ción y una comparación significativa de sus capacidades y limitaciones en el contexto especializado de la sonificación de datos. Además, en nuestro enfoque multidisciplinario, analizamos las estadísticas de contribuciones de la comunidad en GitHub para cada herramienta, lo que arroja luz sobre el nivel de apoyo y el grado de desarrollo colaborativo impulsado por la comunidad de usuarios y desarrolladores en constante evolución.

4.1. Caso de uso

A continuación, se presenta un ejemplo práctico de cómo utilizar las herramientas seleccionadas. En este caso, se analizó minuciosamente la sintaxis y la estructura utilizadas para generar los sonidos a partir de una base de datos específica.

Para llevar a cabo las pruebas, se utilizó la base de datos denominada «informacion_lotes.csv», que contiene información detallada sobre la construcción de lotes en la ciudad de Medellín durante 2019. Esta base de datos proporciona los datos necesarios para generar el sonido inicial en las plataformas.

Este ejemplo práctico permite demostrar el potencial y la versatilidad de las herramientas investigadas en la creación y manipulación eficiente y efectiva de sonidos. Al utilizar la sintaxis y estructura adecuadas, se logró generar el primer sonido de manera simple y sin personalizaciones profundas. Esta acción resalta la capacidad de tales herramientas para realizar tareas básicas y sentar las bases para desarrollos más complejos en el ámbito del sonido en las plataformas.

- TwoTone

Para utilizar esta herramienta de forma web, se puede acceder a través de las siguientes direcciones URL: <https://twotone.>

io/ o <https://twotone-midiout-beta.netlify.app/>. En caso de que se haya instalado localmente, se debe ejecutar el comando npm start y acceder a localhost:9000 en el navegador.

Una vez adentro, es necesario hacer clic en «Get started» y seleccionar la base de datos informacion_lotes.csv (Figura 1). Seguidamente se pueden elegir las columnas que se desean sonificar y el sonido correspondiente para cada una de ellas. Además, es posible modificar parámetros como el volumen, la altura tonal (la afinación), la repetición del sonido y la velocidad en la que se leen los datos.

Es importante destacar que la herramienta cuenta con conexión MIDI, lo que permite enlazarla con sintetizadores, controladores y otros dispositivos que utilicen este protocolo. De este modo, tal acción ofrece la posibilidad de generar interacciones y una interpretación de los datos más completa. También es posible crear filtros mediante la comparación de las columnas seleccionadas con otras de la misma base de datos, para obtener resultados más precisos y ajustados a las necesidades del usuario.

La figura 1 muestra una vista general de la herramienta y sus opciones de configuración.

El resultado puede escucharse en el siguiente enlace: <https://soundcloud.com/acorde-semillero/twotone>



Figura 1: visualización de TwoTone con la base de datos elegida para la comparación.
Nota: la imagen incluye algunos de los instrumentos y la cantidad de datos utilizados.

Fuente: elaboración de los autores.

• SuperCollider

SuperCollider es una plataforma que cuenta con una *library* (biblioteca) que permite leer directamente la base de datos. Sin embargo, para recrear un proceso similar al de TwoTone, que implica cargar un audio y modificar la velocidad de lectura de los datos, se implementó un *script* en Python. Este *script* fue diseñado para leer la base de datos y enviarlos de las columnas seleccionadas a SuperCollider mediante el protocolo OSC (Open Sound Control).

En virtud de su flexibilidad y eficiencia, el protocolo OSC permitió enviar los datos con una velocidad personalizada hacia SuperCollider, donde se realizaron las modificaciones necesarias para generar el sonido deseado. Esta combinación de herramientas y protocolos resultó efectiva para lograr el objetivo de modificar, en tiempo real, el sonido basado en los datos de la base de datos.

El código de Python que envía los datos a SuperCollider es el siguiente:

```
import csv
import time
import numpy as np
from pythonosc import osc_message_builder
from pythonosc import udp_client

def csv_generator(filename):
    with open(filename) as csvfile:
        reader = csv.reader(csvfile, delimiter=";")
        next(reader) # Omitir la primera fila (encabezados)
        print(reader)
        for row in reader:
            print(row)
            print(row[0])
            # Convertir los valores de la fila a los tipos de
            # datos adecuados
            codComuna = np.float64(row[1])
```

```

AreaLote = np.float64(row[3])
AreaCons = np.float64(row[4])
Zona = np.float64(row[5])
numPisos = np.float64(row[7])
yield (codComuna, AreaLote, AreaCons, Zona,
numPisos)
def send_messages(messages, client, address, time_delay):
for message in messages:
msg = osc_message_builder.OscMessageBuilder(address=address)
msg.add_arg(message[0], 'f')
msg.add_arg(message[1], 'f')
msg.add_arg(message[2], 'f')
msg.add_arg(message[3], 'f')
msg.add_arg(message[4], 'f')
msg = msg.build()

client.send(msg)

time.sleep(time_delay)
print(time_delay)

if __name__ == '__main__':
ip = "127.0.0.1"
port = 8888
client = udp_client.SimpleUDPClient(ip, port)

filename = "informacion_lotes.csv"
generator = csv_generator(filename)

address = "/csvSynth"
time_delay = 0.175 # tiempo de espera en segundos entre cada mensaje similar a TwoTone

# Enviar los mensajes en lotes para reducir la cantidad de tiempo de espera
batch_size = 100
messages = [message for message in generator]
for i in range(0, len(messages), batch_size):
send_messages(messages[i:i+batch_size], client, address, time_delay)
print(messages[i:i+batch_size], client, address, time_delay)

```

Una vez ejecutado este script, se procede a abrir SuperCollider para manipular el audio.

```

// Crear un servidor OSC en el puerto 57120
s.boot;

// Crear un cliente de OSC
n = NetAddr.new("127.0.0.1", 57120); // Reemplaza la dirección IP por la de tu computadora si es necesario
(
~codComuna = 0;
~areaLote = 0;
~areaCons = 0;
~zona = 0;
~numPisos = 0;
)

// Crear una función para procesar los mensajes OSC recibidos
(
OSCFunc({ |msg|
postln(msg[0].asString=="/csvSynth");
if(msg[0].asString == "/csvSynth") {
~codComuna = msg[1].asFloat;
~areaLote = msg[2].asFloat;
~areaCons = msg[3].asFloat;
~zona = msg[4].asFloat;
~numPisos = msg[5].asFloat;
}

// Imprimir los valores recibidos
postln("codComuna:" + ~codComuna);
postln("AreaLote:" + ~areaLote);
postln("AreaCons:" + ~areaCons);
postln("Zona:" + ~zona);
postln("numPisos:" + ~numPisos);
}
}, '/csvSynth').add;
)
(
SynthDef("sinteguan",{|codComuna, areaLote, areaCons, zona, numPisos|
var onda, envo, onda2, onda3, filtro;
onda = Saw.ar(zona);

```

```
onda2 = SinOsc.ar(areaCons);
onda3 = SinOsc.ar(areaLote);
filtro = LPF.ar(onda*onda2*onda3,numPisos*300+200);
envo = EnvGen.kr(Env.perc(0.05,0.8),1,doneAction:2);
Out.ar(0,filtro*envo);
}).add
)
(
Tdef(\sabor,{

inf.do{

Synth(\sinteguan,[\zona,\~zona,\areaCons,\~areaCons.midicps,\areaLote,\~areaLote.midicps,\numPisos,\~numPisos]);
0.25.wait;
}

});
)
Tdef(\sabor).play
```

El resultado puede escucharse en el siguiente enlace: <https://soundcloud.com/acorde-semillero/supercollider>

- **Sonic Visualiser**

Para generar el sonido más básico utilizando este programa en Ubuntu, primero debe abrirse desde el panel de aplicaciones. Una vez abierto, se selecciona la opción «Convertir audio a partir de datos» en el menú «Archivo». A continuación, se elige la base de datos que se desea utilizar para sonorizar. Automáticamente se abrirá una ventana donde se pueden seleccionar las columnas específicas que se vayan a emplear (las mismas usadas en TwoTone y SuperCollider). Por último, solo es necesario hacer clic en el botón «Reproducir» para escuchar el sonido resultante (Figura 2).

El resultado puede escucharse en el siguiente enlace: <https://soundcloud.com/acorde-semillero/sonicVisualiser>

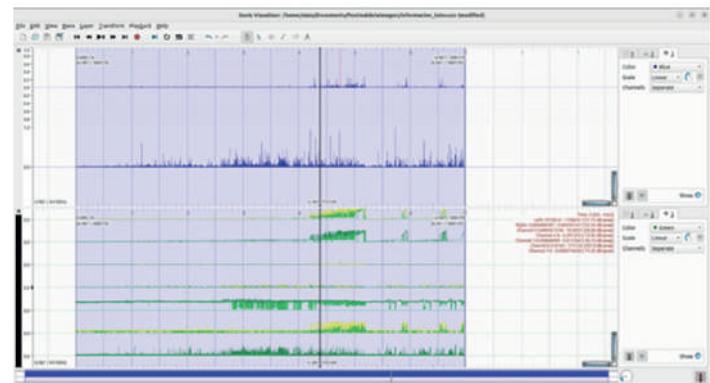


Figura 2: visualización de Sonic Visualiser con la base de datos elegida para la comparación
Nota: la imagen muestra las frecuencias que el programa eligió al convertir los datos.

Fuente: elaboración de los autores.

4.2. Discusión

4.2.1. Comparación entre TwoTone y SuperCollider

En términos de instalación, TwoTone es más fácil de instalar y configurar. Por otro lado, la instalación de SuperCollider puede ser más complicada, ya que requiere de una comprensión básica de programación y de los procesos de compilación. Sin embargo, una vez configurada, SuperCollider es una herramienta muy poderosa y flexible para la síntesis de sonido y el procesamiento de señales de audio.

En relación con la modificación, ambos softwares son altamente personalizables, pero SuperCollider es mucho más versátil y ofrece una amplia variedad de herramientas para la programación en tiempo real.

En términos de integración con Arduino y otros elementos externos, ambos softwares ofrecen una amplia variedad de herramientas para la comunicación con hardware externo. No obstante, la integración de SuperCollider con Arduino es mucho más robusta, ya que TwoTone solo

puede integrarse por medio del protocolo de comunicación MIDI.

4.2.2. Comparación entre TwoTone y Sonic Visualiser

En términos de instalación, Sonic Visualiser es más fácil de instalar y configurar. Sin embargo, su uso se limita principalmente a la visualización y el análisis de archivos de audio, lo que lo hace menos versátil que TwoTone.

En relación con la modificación, Sonic Visualiser es bastante limitado en términos de personalización, ya que no permite cambiar de instrumentos en las diferentes columnas de las bases de datos.

En términos de integración con Arduino y otros elementos externos, Sonic Visualiser no ofrece ninguna opción de integración, mientras que TwoTone brinda algunas opciones de integración básicas como MIDI.

4.2.3. Comparación entre SuperCollider y Sonic Visualiser

En términos de instalación, SuperCollider puede ser más complicado de instalar que Sonic Visualiser.

En relación con la modificación, SuperCollider es mucho más personalizable que Sonic Visualiser, ya que ofrece una amplia variedad de herramientas para la programación en tiempo real y la síntesis de sonido.

En términos de integración con Arduino y otros elementos externos, SuperCollider es mucho más robusto y ofrece más opciones de personalización que Sonic Visualiser.

5. Conclusiones

Elección de la herramienta de software: la elección de la herramienta de software para la sonificación es crucial y debe basarse en el nivel de conocimiento del usuario en programación, datos y música. Herramientas como TwoTone y Sonic Visualiser son ideales para aquellos con un conocimiento técnico limitado, ya que ofrecen interfaces gráficas intuitivas que no requieren una programación extensa.

TwoTone y Sonic Visualiser para usuarios con conocimiento limitado: estas herramientas son especialmente adecuadas para principiantes y personas sin experiencia previa en programación o síntesis de sonido. TwoTone se destaca por su accesibilidad en la conversión de datos en sonidos, mientras que Sonic Visualiser facilita la visualización y el análisis de datos de audio, lo que las hace aptas incluso para usuarios no técnicos.

SuperCollider para usuarios con conocimientos avanzados en programación: SuperCollider es una herramienta poderosa que permite una programación en tiempo real y una personalización avanzada de sonidos. Sin embargo, su uso eficaz requiere un nivel medio o alto de conocimiento en programación. Es la elección preferida para quienes buscan un control total sobre la síntesis de sonido y la manipulación de datos.

Personalización de sonidos: tanto SuperCollider como TwoTone lideran en la personalización de sonidos. SuperCollider brinda un control detallado sobre la síntesis y el procesamiento de señales, ideal para usuarios que desean crear sonidos únicos y experimentar con algoritmos de sonido personalizados. TwoTone, aunque más sencillo, ofrece una interfaz gráfica fácil de usar, pero con menos opciones de personalización en comparación con SuperCollider.

Limitaciones de las herramientas: cada herramienta tiene sus propias limitaciones. TwoTone se enfoca en la creación de sonidos a partir de datos, pero puede tener limitaciones en la síntesis de sonido más avanzada. SuperCollider, por otro lado, puede verse limitado por la habilidad de programación del usuario y los recursos de hardware disponibles, ya que puede ser intensivo en CPU.

Posibilidades creativas: SuperCollider es conocido por su versatilidad y puede utilizarse para crear una amplia variedad de sonidos, desde oscilaciones simples hasta texturas sonoras complejas. Esto lo hace ideal para proyectos artísticos y experimentales. TwoTone, aunque más sencillo, permite a usuarios con diversos niveles de conocimiento musical crear composiciones de manera intuitiva. Sonic Visualiser, en contraste, está diseñado principalmente para análisis científico y carece de las capacidades artísticas complejas de las otras herramientas.

Percepción auditiva y resultados artísticos: los resultados de Sonic Visualiser tienden a ser más científicos y analíticos, lo que puede resultar en la falta de elementos artísticos, como sonidos melódicos o de larga duración. SuperCollider destaca en términos de creatividad, pero requiere un conocimiento previo para lograr resultados artísticos complejos. TwoTone se encuentra en un punto intermedio, ofreciendo una forma accesible de crear música a partir de datos, pero con ciertas limitaciones en la complejidad musical.

Futuros trabajos: se sugiere la exploración de aspectos estéticos y filosóficos en la sonificación de datos para la creación de obras sonoras más reflexivas y profundas en el futuro. Además, la colaboración interdisciplinaria y la investigación de técnicas avanzadas pueden enriquecer la representación de datos y mejorar la calidad artística de las obras sonoras resultantes.

5.1. Limitaciones

La investigación se centró en generar el sonido más sencillo y básico de cada plataforma, evitando entrar en discusiones filosóficas sobre su creación o en reflexiones estéticas. Así, se buscó brindar un primer acercamiento sonoro en cada plataforma y comparar las dificultades encontradas en el proceso.

Como resultado, es importante destacar que los sonoros pueden ser notablemente monótonos y planos si se comparan con otros géneros musicales más complejos y elaborados en términos de arreglos musicales, instrumentación y estructuras melódicas. No obstante, también es cierto que algunos artistas y compositores de música experimental pueden buscar intencionalmente una estética similar a la de los sonoros, explorando la simplicidad y la repetición como elementos expresivos en su música.

En esta investigación, el enfoque se centró en generar los sonidos más simples posibles utilizando los recursos disponibles en cada plataforma. Esto implicó la utilización de osciladores fundamentales, formas de onda básicas y tonos puros, sin agregar elementos adicionales como modulaciones, efectos o procesamiento sonoro complejo. Se abordó cada plataforma desde una perspectiva técnica y objetiva, evitando cualquier interpretación subjetiva o intenciones artísticas. El propósito principal era explorar las capacidades más básicas en la generación de sonido de cada plataforma y evaluar las limitaciones y desafíos encontrados en el proceso. Sin embargo, es esencial recordar que la elección de la herramienta de *software* para la sonificación de datos puede depender en gran medida del nivel de conocimiento del usuario en áreas como la programación, la mu-

sica y la gestión de datos. Estos factores influyen en la capacidad del usuario para aprovechar al máximo las herramientas disponibles y personalizar los resultados según sus necesidades específicas.

En muchas plataformas, la generación de un sonido básico resultó más complicada de lo esperado, debido a limitaciones técnicas, falta de opciones directas para generar ondas puras o una interfaz poco intuitiva. Además, algunos sistemas presentaron una calidad sonora inferior o

una respuesta limitada en relación con la variación sonora.

Es importante resaltar que estos resultados no deben ser considerados como un juicio absoluto sobre las capacidades de cada plataforma. El objetivo de este estudio fue puramente investigativo y no se tuvieron en cuenta otros aspectos que podrían influir en la percepción de la calidad sonora como la capacidad de síntesis, la complejidad de los algoritmos o las opciones de personalización.

Referencias

- Aaron, S. «Sonic Pi-The live coding music synth for everyone». Sonic Pi, 2018. <https://sonic-pi.net/>
- Barsetti-Nerland, Daniel Edward. «Creating a responsive visualization that reacts with music in real time: integrating Ableton Live 9 and Cycling'74 max for live into a musical performance». Tesis de pregrado, University of Arizona, Tucson, 2016. <https://repository.arizona.edu/handle/10150/612546>
- Cannam, Chris, Christian Landone, Mark Sandler y Juan Pablo Bello. «The sonic visualiser: a visualisation platform for semantic descriptors from musical signals». *ISMIR 2006 7th International Conference on Music Information Retrieval*, 324–327. Victoria, BC, Canadá: 2006. <https://nyuscholars.nyu.edu/en/publications/the-sonic-visualiser-a-visualisation-platform-for-semantic-descri>
- Castillo Vázquez, Alina, Diedmar Céspedes Cabrera, Mercedes Jiménez Betancourt y Gustavo Rodríguez Rosabal. «Alternativa didáctica para el estudio de los órganos sensitivos en la enseñanza de la Biología 3». *Revista Dilemas Contemporáneos: Educación, Política y Valores*, vol. 8, n.º 1 (2020). <https://dilemascontemporaneoseduacionpoliticaayvalores.com/index.php/dilemas/article/view/2386>
- De la Vega, Gonzalo, Leonardo Martín Exequiel Domínguez, Johanna Casado, Beatriz García. «SonoUno Web: An innovative user centred web interface». Arxiv (2022). <https://arxiv.org/pdf/2302.00081.pdf>
- Gárate Piñones, I. A. J. «Doroti». Tesis doctoral, Universidad Andrés Bello, 2020.
- Garrett, Richard. «Audio Spray Gun 0.8-The generation of large sound-groups and their use in three-dimensional spatialisation». *ICMC* (2015). <https://quod.lib.umich.edu/cgi/p/pod/dod-idx/audio-spray-gun-08-the-generation-of-large-sound-groups.pdf?c=icmc;idno=bbp2372.2015.071;format=pdf>
- Martínez Cantón, Carlos. «Diseño y construcción de un detector de metales basado en un microcontrolador». Tesis de pregrado, Universitat Politècnica de Catalunya, España, 2016. <https://upcommons.upc.edu/handle/2117/98926?locale-attribute=es>
- Peralta Rocha, Bryan Marcelo. «Inspección y pintura del avión Escuela Hawker Siddeley 125-400, mediante información y datos técnicos para la unidad de gestión de tecnologías UGT-ESPE». Tesis de pregrado, Universidad de las Fuerzas Armadas, Sangolquí, Ecuador, 2018. <https://repositorio.espe.edu.ec/bitstream/21000/27745/1/M-ESPEL-CMA-0478.pdf>
- Quimica.es. «Estetoscopio». Quimica.es. <https://www.quimica.es/enciclopedia/Estetoscopio.html>

- Ramírez Arias, Leonardo y Julián Uribe Cortés. «El contador Geiger como herramienta en la enseñanza del concepto de radiación». *Góndola, Enseñanza y Aprendizaje de las Ciencias*, vol. 2, n.º 1 (2007): 33-36. <https://revistas.udistrital.edu.co/index.php/GDLA/article/download/5304/6935/0>
- Remiro Gargallo, Andrea Vicenta. «Técnicas numéricas aplicadas a la propagación de sonido en el medio marino». Tesis doctoral, Universitat Politècnica de València, España, 2019. <http://hdl.handle.net/10251/128636>
- Rodríguez Mateos, David y Alicia Tapia López. «Sonificación y periodismo: la representación de datos mediante sonidos». *Revista de Comunicación*, vol. 22, n.º 1 (2023): 413-433. <http://dx.doi.org/10.26441/rc22.1-2023-3022>
- Ruiz-del-Olmo, Francisco Javier y José Vertedor. «Procesos, herramientas y prácticas de la sonificación». Málaga, España, Universidad de Málaga, 2016. <http://hdl.handle.net/10630/12429>
- Sánchez Corzo, Andrea del Pilar. «Sonorización de señales de EEG basada en estructuras musicales». Tesis de pregrado, Universidad de los Andes, Bogotá, 2012. <https://repositorio.uniandes.edu.co/bitstream/handle/1992/25182/u627972.pdf?sequence=1&isAllowed=y>
- Saraví, Fernando. «Laennec, el método anatomooclínico y la invención del estetoscopio. Parte 1: De Bretaña a París, hasta la gran invención». *Revista Médica Universitaria*, vol. 8, n.º 2 (2012). https://bdigital.unc.edu.ar/objetos_digitales/4769/rmuhistoriasdelamedicina.pdf
- Schaeffer, Pierre, Christine North y John Dack. *Treatise on Musical Objects: An Essay Across Disciplines*. California: University of California Press, 2017. <https://www.jstor.org/stable/10.1525/j.ctt1qv5pqb>
- Staples, Anne. «El abuso de las campanas en el siglo pasado». *Historia Mexicana*, vol. 27, n.º 2 (1977): 177-194. <https://www.jstor.org/stable/25135599>
- Vercoe, B., J. Fitch, J. Piche, P. Nix, R. Boulan-ger, R., Ekman y A. Cabrea. «The canonical Csound reference manual». Csound, 2007. <https://csound.com/docs/manual/index.html>
- Vickers, Paul y Bennet Hogg. «Sonification abstraite / Sonification concrète: An aesthetic perspective space for classifying auditory displays in the Ars Musica domain». En *Proceedings of the 12th International Conference on Auditory Display* (210-216). Londres, 2006. <https://typeset.io/pdf/sonification-abstraite-sonification-concrete-an-aesthetic-1urtp0ql.pdf>
- Walker, Bruce N. y Joshua T. Cothran. «Sonification sandbox: a graphical toolkit for auditory graphs». *Proceedings of the 2003 International Conference on Auditory Display*. Atlanta, GA: Georgia Institute of Technology, 2003.
- Wang, Ge, Dan Trueman, Scott Smallwood y Perry R. Cook. «The laptop orchestra as classroom». *Computer Music Journal*, vol. 32, n.º 1 (2008): 26-37. <https://www.jstor.org/stable/40072662>
- Wang, Hengbo. «Dear diary: an investigation of emotionally engaging computer-game narratives». Tesis de maestría, Auckland University of Technology, Australia, 2015. <http://hdl.handle.net/10292/9137>
- Wikipedia. «Sonar». Wikipedia, 21 de marzo de 2019. <https://es.wikipedia.org/wiki/Sonar>
- Wilson, Scott, David Cottle y Nick Collins. *The SuperCollider Book*. Cambridge, MA: The MIT Press, 2011.





Vol. 3 N.º 9 / Guayaquil
mayo 2024
ISSN 2953-657x

Pódcast inmersivo para la divulgación de historias de niños, niñas y adolescentes víctimas del conflicto armado colombiano

María Manuelita Castañeda Quintero
Estudiante en el programa de Ingeniería de Sonido
Universidad de San Buenaventura, sede Bogotá
mmcatanedaq@academia.usbbog.edu.co

Julia Mercedes Cubides Cardeña
Estudiante en el programa de Ingeniería de Sonido
Universidad de San Buenaventura, sede Bogotá
jmcubidesc@academia.usbbog.edu.co

Laura Sofía Pineda Gómez
Estudiante en el programa de Ingeniería de Sonido
Universidad de San Buenaventura, sede Bogotá
lspinedag@academia.usbbog.edu.co

Resumen

En este documento se expone un estudio de caso de un pódcast con tecnología inmersiva en el que se divulga la historia de una niña víctima del conflicto armado, con el propósito de recrear de forma más realista los hechos y generar una experiencia sensorial más profunda. Para su creación se usa el modelo Design Thinking, el cual consiste de cinco etapas: empatizar, definir, idear, prototipar y testear. En las etapas finales, se realiza el diseño sonoro del relato con el uso del plugin DearVR PRO en formato binaural y finalmente se llevan a cabo quince pruebas subjetivas sobre la inmersividad del pódcast. Se determina que con tecnología inmersiva los oyentes generan interés en este tipo de contenido, lo cual da paso a la posibilidad de divulgar estas historias con un formato innovador que ofrece una mayor claridad de los hechos.

Palabras clave: audio inmersivo; experiencia inmersiva; pódcast

Abstract

The following document presents a case study of a podcast with immersive technology in which the story of a little girl victim of the armed conflict is disclosed, with the purpose of recreating the events in a more realistic way and generating a deeper sensory experience. For its creation, the Design Thinking model is used which consists of five stages: empathize, define, ideate, prototype, test and implement. In the final stages the sound design of the story is done with the use of the DearVR PRO plugin in binaural format and finally fifteen subjective tests are carried out on the immersiveness of the podcast. It is determined that with immersive technology the listeners generate interest in this type of content, which leads to the possibility of disseminating these stories in an innovative format that offers more clarity over these events.

Keywords: immersive audio; immersive experience; podcast

1. Introducción

El esclarecimiento y la divulgación de los sucesos que acontecieron en el marco del conflicto armado colombiano se presenta como un tema de gran importancia para una sociedad que añora la paz, puesto que el reconocimiento de lo sucedido permite sentar las bases para la no repetición y la construcción de un país en paz. Según el Centro Nacional de Memoria Histórica, «la memoria histórica es un vehículo para el esclarecimiento de los hechos violentos, la dignificación de las voces de las víctimas y la construcción de una paz sostenible en los territorios»¹. De ese modo, hacer memoria permite preservar la identidad, identificar y saber quiénes somos, de dónde venimos y cómo hemos evolucionado con el tiempo. La memoria histórica es un proceso colec-

tivo en el que se crea un lenguaje que da un significado a las memorias de los miembros de una sociedad, generando así un sentido compartido de los eventos vividos.² Esto es muy importante hacerlo ya que «la exposición pública y oficial de la verdad facilita el perdón y olvido de las víctimas»³. Por tanto, la creación de contenido digital como pódcasts, documentales, artículos, videos y demás que traten este tema se convierte en un aspecto fundamental para el cumplimiento de este objetivo. Asimismo, el uso de herramientas que proporcionan un acercamiento más real a los consumidores ayuda a que el mensaje se transmita de forma certera, por esta razón, el uso de audio inmersivo se hace sumamente crucial para lograr esta meta; un ejemplo de esto es la serie radial de quince capítulos *Luz de Noche*, un rela-

¹ Centro Nacional de Memoria Histórica, «Caminos para la Memoria», Centro Nacional de Memoria Histórica, 2015, <https://centredememorihistorica.gov.co/micrositios/caminosParaLaMemoria/>

² Ángela Aguirre, «¿Por qué es importante la memoria histórica en Colombia?», *Revista del Colegio Mayor de Nuestra Señora del Rosario*, vol. 1, n.º 3 (2015).

³ Yeni Loaiza, «La verdad como criterio fundamental de reparación a las víctimas en la justicia transicional colombiana», *Revista Lasallista de Investigación*, vol. 16 n.º 2 (2019).

to colectivo sobre la Comisión de la Verdad que busca relatar el dolor y a su vez la resistencia y resiliencia de un pueblo que a pesar de todo busca la paz y la esperanza del perdón.⁴ ¿Por qué un pódcast inmersivo? El pódcast, en comparación con otros medios digitales, se considera como un medio íntimo que, a través de su narrativa y contenido, puede generar una sensación de inmersión en los oyentes. En palabras de los creadores de este tipo de contenido, se han escogido géneros que les permiten incluir diversos elementos sonoros que generan una mayor sensación de inmersión en el oyente, con lo que es posible visualizar los elementos representados.⁵ Por otro lado, el audio inmersivo es un término que ha tomado gran relevancia en los últimos tiempos, esto gracias a que provee una experiencia más real y personal a los oyentes haciendo uso de sistemas que permiten el posicionamiento de fuentes puntuales de sonido en diferentes ejes del espacio, lo que simula una sensación realista de encontrarse dentro de un campo sonoro.⁶ Un ejemplo de lo que audio inmersivo puede generar en una narración se muestra en un proyecto de Sadhbh Bashford Rance⁷ en el que se busca explorar la respuesta emocional y la sensación de inmersión que tienen los oyentes al escuchar una narración con formato inmersivo; se determinó que el uso de música, Foley, paisajes detallados y el audio binaural permite que los oyentes empaticen con las narraciones que escuchan. Del mismo modo, en

un artículo de Euritt y McMurtry⁸ se explora como el pódcast inmersivo en narrativas del género de terror le da una experiencia sensorial al oyente que le permite comprender mejor el horror, además de posicionarlo en medio de los mundos narrativos que crean los productores. Con esto, se determina que el pódcast en formato inmersivo es un medio digital que no solo permite divulgar estos relatos, sino que también permite generar un acercamiento a los oyentes con la narrativa. Por lo tanto, en el presente trabajo se busca crear un pódcast inmersivo que se componga de un capítulo en el que se relaten historias, documentadas por la Comisión de la Verdad, de niños, niñas y adolescentes víctimas del conflicto armado.

2. Parte técnica del artículo

Se utiliza la metodología de investigación Design Thinking cuyo objetivo es permitir la innovación en la creación de productos y servicios. Esta consta de cinco etapas que impulsan el desarrollo de soluciones centradas en el usuario.⁹ A continuación, se describen las etapas y el proceso que se lleva a cabo en cada una.

2.1. Empatizar

En esta etapa se hace el mapeo de actores, esto quiere decir que se identifica el público al que va dirigido el pódcast. Para esto se realiza una revisión de distintas fuentes bibliográficas relacionadas con los procesos de paz, como la Comisión de la Verdad¹⁰ y el Centro Nacional de Memoria Histórica¹¹,

4 Comisión de la Verdad, *Luz de Noche, una serie de la Comisión de la Verdad inspirada en las víctimas del conflicto armado*, 2021.

5 Sneha Gore Mehendale, «Immersion Is “Hear”: A Practitioners’ Perspective on Immersive Podcasts», *Journal of Content, Community & Communication*, vol. 9 (2019).

6 Alba Pelegrin, «El audio Inmersivo y cómo está cambiando la industria musical», *La Rosa*, 27 de septiembre de 2022, <https://www.larrosa.pro/post/el-audio-inmersivo-y-c%C3%B3mo-est%C3%A1-cambiando-la-industria-musical>

7 Sadhbh Bashford Rance, «An Exploratory Study on How Binaural Sound Can Improve the Immersive Quality of Audio Storytelling» (tesis de grado, Institute of Art, Design and Technology IADT, Irlanda, 2022).

8 Alyn Euritt y Leslie McMurtry, «Immersive night: audio horror in radio and podcasting», *Refractory: a Journal of Entertainment Media*, vol. 35 (2021).

9 Luciano Augusto Toledo, Marcos Fernando Garber y Adriana Beatriz Madeira, «Consideraciones acerca del Design Thinking y Procesos», *Revista Gestão & Tecnologia*, vol. 17, n.º 3 (2017): 284–304.

10 Comisión de la Verdad, «Comisión de la Verdad», 2022, <https://www.comisiondelaverdad.co/>

11 Centro Nacional de Memoria Histórica, «Caminos para la Memoria».

con el objetivo de comprender las necesidades en este tipo de procesos y contribuir en la concientización y en la construcción de paz en el país. En este caso, se observa una notable falta de difusión de estos acontecimientos, ya que para generar memoria histórica en la población se requiere de medios digitales que produzcan sentimientos de solidaridad hacia estos sucesos. Por lo tanto, se espera que los principales oyentes sean los ciudadanos colombianos.

2.2. Definir

En esta etapa se define el problema que se quiere solucionar. Teniendo en cuenta la necesidad descrita en la anterior etapa, se define que la problemática que se va a abordar es la falta de divulgación de los hechos ocurridos en el marco del conflicto armado colombiano.

2.3. Idear

A partir del problema planteado se generan ideas creativas y diversas. Entre diferentes medios digitales como documentales, artículos, y videos se escoge finalmente el pódcast, ya que es un formato que permite llevar al oyente a una experiencia sonora más personal, además de ser un contenido que puede ser consumido a cualquier hora y en cualquier lugar, lo que permite que llegue a más personas.¹² Asimismo, el formato inmersivo es considerado en este pódcast, pues, como se muestra en los textos de Sadhbh Bashford Rance¹³ y Euritt y McMurtry¹⁴, el audio inmersivo permite que los oyentes tengan un acercamiento más empático con las narraciones que se muestran.

¹² José González, «Innovación en contenidos para públicos distintos. El auge del pódcast como nueva narrativa periodística», *Cuadernos del Periodista*, n.º 37 (2019).

¹³ Bashford Rance, «An Exploratory Study...».

¹⁴ Euritt y McMurtry, «Immersive night...».

2.4. Prototipar

Se propone la creación de un capítulo del pódcast narrando el fragmento «Yo sentía la muerte: los primeros recuerdos de Olga María», tomado del volumen «No es un mal menor» del último informe de la Comisión de la Verdad del 2022.

Para la realización del pódcast se planifica un *storyboard* donde se organizan las escenas, comenzando con la descripción. Allí se definen los personajes, el tipo de narración y el libreto de cada uno. Luego, en la sección de tiempo se calcula desde y hasta qué parte se extiende cada escena según el tiempo de narración. Por otro lado, se utilizan los «procesos dinámicos», que son aquellos diferentes procesos por los que pasan los clips de audio del pódcast para mejorar la calidad, control y énfasis que requiere.

La siguiente sección la componen los aportes del sonido inmersivo que detalla cada parte del diseño sonoro, como los efectos, ambientes y pasos. Estos, por medio del plugin DearVR PRO, dan la sensación inmersiva.

Por último está la sección de panorama que describe en qué zona se encuentra cada parte del diseño sonoro. A continuación, en la tabla 1, se muestra un fragmento del *storyboard* de la primera escena.

Tabla I
Esquema del *storyboard*

# Escena	Descripción	Tiempo	Procesos Dinámicos	Sonido Inmersivo	Panorama
1	«Era de noche y se escuchaban perros aullando...	00:00 – 00:10	Voz íntima. Aullidos de perros con filtro...	Ambiente de noche. Voces y pasos alrededor.	Río en la izquierda. Perro ladrando en la derecha...

Asimismo, como se muestra en las figuras 1 y 2, se ha desarrollado un gráfico visual ini-

cial de cómo están ubicados los personajes y los demás elementos en el escenario, que en este caso son la casa de Olga María y el comedor.

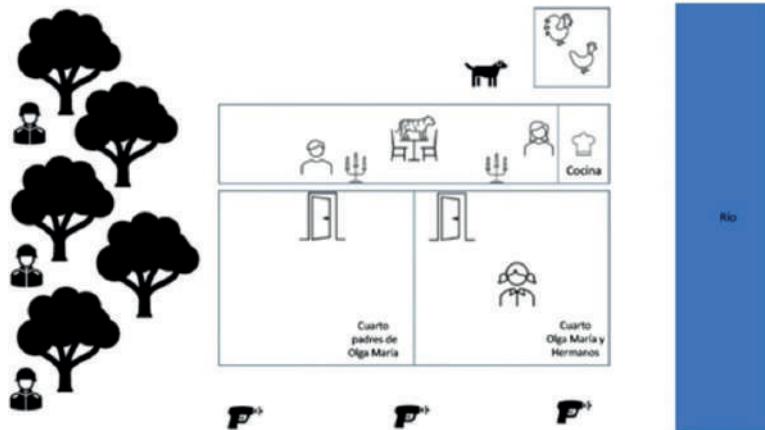


Figura 1: gráfico de posición de personajes en la casa

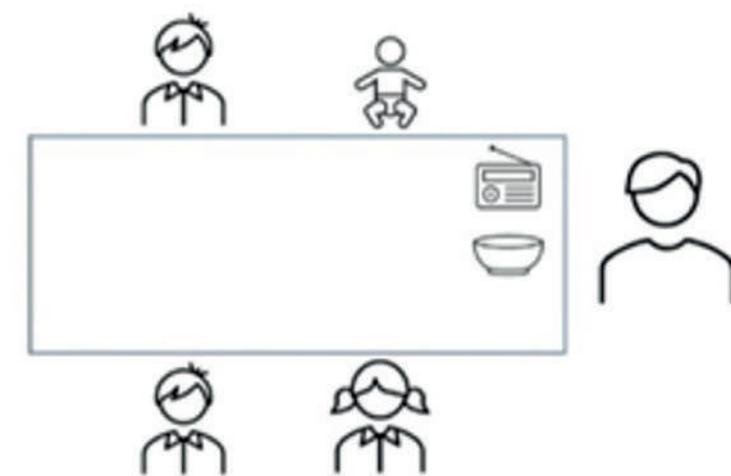


Figura 2: gráfico de posición de personajes en el comedor

Para el diseño sonoro se determinan los ambientes y elementos sonoros que acompañan el relato, obtenidos a través de bancos de sonidos o grabados por medio de Foley. Un ejemplo de esto son el sonido de los pasos al caminar, algunos roces de ropa y golpes. Estos elementos se dispondrán a la etapa de postproducción en la que se realizará, primero, un proceso de edición y selección, seguido de un proceso de reducción de ruido y sonidos no deseados.

Posteriormente, teniendo en cuenta la narración, se ha desarrollado una secuen-

cia que represente auditivamente lo que se narra a través del diseño sonoro para luego realizar la mezcla. En esta se busca obtener un balance de todos los elementos sonoros

y, dependiendo de lo que se requiera, su respectivo proceso de ecualización, compresión y efectos.

Después, con el plugin DearVR PRO —que usa la tecnología de audio 3D binaural para simular la audición espacial humana a través de auriculares— se le da a la mezcla una imagen inmersiva. De este plugin, se usa principalmente la posición, la elevación, la dirección y la distancia; también se usa el pad gráfico XYZ, que, por medio de su interfaz (como se muestra en la figura 3), permite colocar las fuentes de sonido en cualquier cuadrante del escenario para hacer un panorama de 360°. Además, se usa la reverberación del plugin para dar la sensación de entorno acústico adecuado. Finalmente, con la automatización se les da a ciertos sonidos la sensación de movimiento.



Figura 3: interfaz gráfica de posición, Pad Gráfico XYZ DearVR PRO.

2.5. Testear

Con el propósito de determinar el nivel de inmersividad que siente le oyente al escuchar el pódcast en formato binaural, se ha desarrollado un testeo enfocado en tres adjetivos calificativos: distancia y profundidad; claridad; e impresión espacial.¹⁵ En la evaluación subjetiva se opta por un espacio sonoro controlado¹⁶ y como sistema de reproducción se hace uso de los audífonos Audio Technica ATH-M50x.

La técnica de muestreo empleada es no probabilística accidental.¹⁷ Se reclutan quince personas que, en su mayoría, son estudiantes de quinto semestre de Ingeniería de Sonido que pasan por el estudio de grabación con el objetivo de tener una perspectiva general de cómo la sensación de inmersividad es percibida. El tamaño de la muestra se determina considerando las limitaciones de tiempo y se decide dejar el número de participantes que accedieron a realizar la prueba. Además, teniendo en cuenta que las tecnologías inmersivas intentan representar la forma natural en la que escuchamos los seres humanos, no se requiere un conocimiento amplio en la temática, puesto que es algo que se experimenta naturalmente.

En la prueba realizada se hicieron tres preguntas, con cuyas respuestas —utilizando los tres adjetivos descritos anteriormente— se define el nivel de inmersividad que sintió el oyente al escuchar el pódcast. Esta evaluación se hizo por medio de la escala de Likert, con la cual se mide la confor-

¹⁵ Nick Zacharov, Toben Holm Pedersen y Chris Pike, «A common lexicon for spatial sound quality assessment—latest developments», 2016 Eighth International Conference on Quality of Multimedia Experience (QoMEX) (Lisboa: 2016).

¹⁶ Estudio híbrido de la Universidad de San Buenaventura.

¹⁷ Tamara Otzen y Carlos Manterola, «Técnicas de muestreo sobre una población a estudio», *International Journal of Morphology*, vol. 35, n.º 1 (2017): 227-232.

midad que presenta el oyente con respecto a las afirmaciones presentadas.¹⁸

En las interrogaciones se establece un rango de respuesta de 1 a 5, siendo 1 totalmente en desacuerdo y 5 totalmente de acuerdo. Las preguntas son las siguientes:

- a. ¿Fue posible distinguir la cercanía y lejanía de los diferentes elementos sonoros del pódcast?
- b. ¿Hubo claridad al momento de distinguir los diferentes elementos que se presentaban a lo largo del pódcast?
- c. ¿Al momento de escuchar el pódcast, tuvo la sensación de encontrarse en un entorno caracterizado (entorno caracterizado se refiere a los diferentes espacios que se representan de la historia narrada)?

3. Resultados

Los resultados y gráficos obtenidos a través de la prueba subjetiva se presentan a continuación:

La primera pregunta hace referencia a la posibilidad de distinguir la distancia en la que se ubican los diversos elementos del diseño sonoro. Como se observa en la figura 4, el 40 % de los participantes estuvo de acuerdo, el 53,3 % totalmente de acuerdo y el 6,7 % tomó una posición neutral con respecto a esta afirmación.



Figura 4: resultados estadísticos de la primera pregunta de la prueba realizada.

¹⁸ Andrés Muguira, «¿Qué es la escala de Likert y cómo utilizarla?», QuestionPro, <https://www.questionpro.com/blog/es/que-es-la-escala-de-likert-y-como-utilizarla/>

La segunda pregunta indaga sobre la capacidad del oyente de distinguir los diversos elementos presentados a través de la historia de forma clara. Como se observa en la figura 5, un 46,7 % de los participantes se encuentra de acuerdo, mientras que el 53,3 % está totalmente de acuerdo.



Figura 5: resultados estadísticos de la segunda pregunta de la prueba realizada.

La tercera pregunta trata sobre la inmersión experimentada por el oyente y la posibilidad de sentirse parte del entorno creado mediante el proceso de producción. En este caso, en la figura 6 se observa que el 46,7 % de los oyentes estuvieron totalmente de acuerdo, el 40 % estuvieron de acuerdo y el 13,3 % tomó una posición neutral con respecto a esta afirmación.



Figura 6: resultados estadísticos de la tercera pregunta de la prueba realizada.

4. Análisis de resultados

El pódcast se desarrolla como un prototipo que es probado solo una vez y para ello se

consideran los atributos perceptuales anteriormente mencionados. Los resultados obtenidos de las pruebas subjetivas son los siguientes:

Con respecto a la profundidad de los elementos sonoros, un 93,3 % de los participantes expresó acuerdo o total acuerdo en que fue posible distinguir la distancia en la que se ubican los elementos. Esto quiere decir que la mayoría de los oyentes percibieron claramente la ubicación de los componentes de la narración.

En cuanto a la claridad, se puede decir que todos los oyentes estuvieron de acuerdo con respecto a la capacidad de distinguir los elementos presentados a través de la historia de forma clara. Esto indica que los oyentes fueron capaces de distinguir los elementos sonoros representados en cada escena.

Además, un 86,7 % de los oyentes estuvieron de acuerdo con respecto a la afirmación sobre la inmersión experimentada. Sin embargo, un 13,3 % tomó una posición neutral, indicando que su sensación de inmersión no fue tan significativa como para la mayoría de los participantes.

Con estos resultados, se evidencia que el producto creado permite tener un grado de inmersividad considerable, ya que en cada atributo perceptual evaluado se obtuvo una respuesta, en su mayoría, positiva. Esto indica que estas características sonoras fueron percibidas en el pódcast. No obstante, la razón por la que algunos oyentes no tuvieron una sensación de inmersividad posiblemente se deba a las diferencias con respecto a la función de transferencia con la cabeza (HRTF) de cada persona, ya que el tamaño, la forma de la cabeza, las orejas y demás son factores que influyen la forma en la que se perciben este tipo productos en formato binaural, por tanto, es normal que en algunos casos la inmersión no fuera efectiva.

5. Conclusiones

Con respecto a los resultados obtenidos en las pruebas subjetivas se puede evidenciar que, en general, los adjetivos calificativos fueron distinguibles para los oyentes. Con esto, se determina que cada elemento sonoro empleado influyó significativamente en la narración y posibilitó la creación de un ambiente envolvente que transporta a los oyentes a los escenarios de la historia. Según la prueba, los oyentes pudieron distinguir los elementos sonoros de manera clara, es decir que la mezcla y producción del pódcast fueron efectivas, ya que permitieron tener claridad de los detalles importantes que hacen apreciar la narrativa de forma más completa.

Por otro lado, es importante destacar el uso de un *storyboard* bien estructurado, pues esto permite una buena organización y producción del pódcast. Tener planificada cada escena hace que haya una lógica secuencial junto a una visión clara de cómo será la integración de todos los elementos en la narración. Asimismo, una descripción detallada de cada elemento del diseño sonoro permite organizar las ideas principales con el fin de comunicar la historia de un modo más claro y conciso. También, facili-

ta entender la estructura y el contenido del pódcast y, si se trabaja en equipo, permite la colaboración de cada miembro.

Para concluir, estos resultados demuestran que el pódcast inmersivo logró generar una experiencia sensorial en los oyentes y así se logró que tuvieran una mayor compresión y acercamiento al relato. Entonces, se determina que el uso del pódcast como medio digital para la divulgación de este tipo de relatos es efectivo en lo que respecta a dar una sensación más íntima de la narración. Además, es un medio que permite que las personas visualicen los elementos representados, llevando a que tengan una mayor inmersión. Finalmente, se debe resaltar que, tras la prueba realizada, los participantes mostraron gran interés por el producto, destacando que el formato usado hace que estas historias se sientan más cercanas y presentes, un factor importante al momento de captar la atención de quien escucha para transmitir el mensaje deseado. Así, se abren las puertas a la posibilidad de una mayor difusión de los hechos sucedidos bajo el marco del conflicto armado en Colombia por medio de la creación de contenido de calidad y con alta capacidad de envolver al oyente.

Referencias

- Aguirre, Ángela. «¿Por qué es importante la memoria histórica en Colombia?». *Revista del Colegio Mayor de Nuestra Señora del Rosario*, vol. 1, n.º 3 (2015).
- Audio Engineering Society sección Colombia. *Avances del audio en Latinoamérica*. 2016.
- Bashford Rance, Sadhbh. «An Exploratory Study on How Binaural Sound Can Improve the Immersive Quality of Audio Storytelling». Tesis de grado, Institute of Art, Design and Technology (IADT), Irlanda, 2022.
- Centro Nacional de Memoria Histórica. «Caminos para la Memoria». Centro Nacional de Memoria Histórica, 2015. <https://centrodememoriahistorica.gov.co/micrositios/caminosParaLaMemoria/>
- Comisión de la Verdad. *Luz de Noche, una serie de la Comisión de la Verdad inspirada en las víctimas del conflicto armado*. 2021.
- . «Comisión de la Verdad». 2022. <https://www.comisiondelaverdad.co/>
- Euritt, Alyn y Leslie McMurtry. «Immersive night: audio horror in radio and podcasting». *Refractory: a Journal of Entertainment Media*, vol. 35 (2021).
- González, José. «Innovación en contenidos para públicos distintos. El auge del pódcast como nueva narrativa periodística». *Cuadernos del Periodista*, n.º 37 (2019).
- Leslie Gaston-Bird, R. M. (2021). Inclusivity in Immersive Audio: Current Participation and Barriers to Entry.
- Loaiza, Yeni. «La verdad como criterio fundamental de reparación a las víctimas en la justicia transicional colombiana». *Revista Lasallista de Investigación*, vol. 16 n.º 2 (2019).
- Mäkivirta, Aki. «Audio Inmersivo: Las soluciones del mañana, hoy». Audio Technica Disstributed Brands. 28 de noviembre de 2017. <https://distribution.audio-technica.eu/es-es/noticias/audio-inmersivo-las-soluciones-manana-hoy/>
- Mehendale, Sneha Gore. «Immersion Is “Hear”: A Practitioners’ Perspective on Immersive Podcasts». *Journal of Content, Community & Communication*, vol. 9 (2019).
- Muguira, Andrés. «¿Qué es la escala de Likert y cómo utilizarla?». QuestionPro. <https://www.questionpro.com/blog/es/que-es-la-escala-de-likert-y-como-utilizarla/>
- Otzen, Tamara y Carlos Manterola. «Técnicas de muestreo sobre una población a estudio». *International Journal of Morphology*, vol. 35, n.º 1 (2017): 227-232.
- PCQUEST. «The demand for enhanced audio experiences to continue to grow in 2022». PCQUEST. 10 de enero de 2022. <https://www.pcquest.com/demand-enhanced-audio-experiences-continue-grow-2022/>
- Pelegrin, Alba. «El audio Inmersivo y cómo está cambiando la industria musical». La Rosa. 27 de septiembre de 2022. <https://www.larrosa.pro/post/el-audio-inmersivo-y-c%C3%B3mo-est%C3%A1-cambiando-la-industria-musical>
- Rumsey, F. «Immersive audio Objects, mixing, and rendering». *Journal Audio Engineering Society* (2016): 7-8.
- Toledo, Luciano Augusto, Marcos Fernando Garber y Adriana Beatriz Madeira. «Consideraciones acerca del Design Thinking y Procesos». *Revista Gestão & Tecnologia*, vol. 17, n.º 3 (2017): 284-304.
- Zacharov, Nick, Toben Holm Pedersen y Chris Pike. «A common lexicon for spatial sound quality assessment-latest developments». *2016 Eighth International Conference on Quality of Multimedia Experience (QoMEX)*. Lisboa: 2016.





Vol. 3 N.º 9 / Guayaquil
mayo 2024
ISSN 2953-657x

Estudio de viabilidad del residuo del cultivo de cáñamo como material de acondicionamiento acústico

Timothy Antonio Carrera Bleiksloot

Facultad de Ingeniería y Ciencias Aplicadas
Universidad de las Américas

xxxxxxxxxxxx@xxxxxxxxc

Resumen

La industria del cannabis crece globalmente y llama la atención de investigadores en diversas áreas. En acústica, el cannabis se propone como una alternativa para reducir el ruido en interiores debido a sus propiedades en este campo. Las mediciones de cannabis en un tubo de impedancia son una técnica no invasiva y no destructiva que determina las propiedades eléctricas y acústicas de las muestras. La técnica es útil para determinar la calidad de la muestra, ya que las propiedades acústicas están relacionadas

con la densidad y homogeneidad de la muestra. Una muestra de alta calidad tendrá una impedancia acústica uniforme. La técnica es rápida y precisa, proporcionando resultados en minutos. Además, el cannabis es una alternativa económica y sostenible a los materiales acústicos tradicionales para reducir el ruido en interiores. En resumen, las mediciones de cannabis en un tubo de impedancia proporcionan información valiosa para productores e investigadores que estudian las propiedades acústicas de la muestra por lo que ayudará a determinar la viabilidad de los residuos como material de acondicionamiento acústico.

Palabras clave: absorbente acústico, cannabis sativa, cáñamo, economía ecuatoriana, industria

Abstract

The cannabis industry is growing globally and attracting the attention of researchers in various areas. In acoustics, cannabis is proposed as an alternative

to reduce indoor noise due to its acoustic properties. Cannabis measurements in an impedance tube are a non-invasive and non-destructive technique that determines the electrical and acoustic properties of cannabis samples. The technique is useful in determining the quality of the sample, as the acoustic properties are related to the density and homogeneity of the sample. A high-quality sample will have a uniform acoustic impedance. The technique is fast and accurate, providing results in minutes. In addition, cannabis is an economical and sustainable alternative to traditional acoustic materials for reducing indoor noise. In summary, cannabis measurements in an impedance tube provide valuable information for producers and researchers studying the acoustic properties of the sample and will help us determine the viability of the waste as an acoustic conditioning material.

Keywords: acoustic absorber, cannabis sativa, hemp, Ecuadorian economy, industry

1. Introducción

La sociedad desconoce las dimensiones reales del cannabis.¹ Se ha desarrollado un pensamiento erróneo generalizado sobre su uso, limitándolo a la recreación, cuando en realidad, dependiendo de la variedad, se lo puede emplear en distintos campos industriales. *Cannabis sativa* es una planta que puede ser utilizada casi en su totalidad: de ella se adquieren combustibles, medicamentos, alimentos, incluso fibras textiles.²

Esta especie vegetal pasó de ser recolectada a ser cultivada. Una de sus clasificaciones corresponde al cáñamo que se

ha sometido a industrialización. Al ser una nueva industria generará numerosas plazas de trabajo y esto favorecerá la dinamización económica en diversas áreas productivas del Ecuador, comenzando principalmente por la reactivación y renovación de la industria agrícola, pasando por la industria de los procesos y finalmente la comercialización de la materia prima o producto elaborado (importación y exportación de cáñamo).³ Aunque se ha utilizado desde hace mucho tiempo, recientemente ha surgido un gran interés en el uso de todas sus dimensiones, incluida la industria arquitectónica en la que su fibra actúa como material absorbente acústico natural y ecológico

1 Psyckus, «La gente no sabe. Cannabis», Desmotivaciones.es, 27 de agosto de 2012, <https://desmotivaciones.es/5848292/La-gente-no-sabe>

2 Guadalupe Ángeles-López, Fernando Brindis, Sol Cristians-Niizawa y Rosa Ventura-Martínez, «Cannabis sativa L., una planta singular», *Revista Mexicana de Ciencias Farmacéuticas*, vol. 45, n.º 4 (2014): 2.

3 Henry Gallegos-Dávila, «Aplicaciones de cáñamo como alternativa rentable a la reactivación económica de ecuador tras la pandemia de covid-19», *Perfiles*, vol. 1, n.º 25 (2021): 46.

que se aplica en el ámbito de aislamiento o acondicionamiento acústico.

El ruido es indudablemente parte de nuestras vidas, por lo que cada vez es mayor la preocupación por desarrollar mejores condiciones de aislamiento y acondicionamiento acústico en nuestra arquitectura. En los últimos años se han desarrollado teorías y modelos matemáticos que ayudan a predecir el comportamiento acústico de algunos materiales; y debido a la creciente tendencia en el uso de nuevas materias primas, el mercado se está abriendo cada vez más al uso de materiales alternativos.⁴

1.1. Antecedentes

Las fibras naturales como materiales son llamativas por la fuerza, la tenacidad, la resistencia al impacto y la flexibilidad. Ofrecen aislamiento térmico, acústico y lumínico para disminuir el consumo de combustibles y mitigar las emisiones.⁵

Se pretende utilizar esta agrofibra como sustituto de gran parte de polímeros mucho menos ecológicos⁶ aplicados a la acústica. En el exterior ya se ha implementado el uso de la fibra para la elaboración de paneles acústicos en compañías como SilentFiber, empresa emergente de Alta Austria de Viena, galardonada al premio German Business Awards 2022 como mejores proveedores de paneles acústicos de la región de habla alemana con la fabricación de paneles de cáñamo, un material ecológico que ab-

⁴ David Fernando Tapia-Gavilanes, «Caracterización de materiales absorbente y modelos matemáticos» (tesis de máster, Universidad Politécnica de Madrid, 2019), 11.

⁵ Jonnathan Rivas e Inés Torres, «Paneles de revestimiento acústico. Uso de materiales compuestos reforzados con fibras de totora adaptados al centro pedagógico de la ciudad de Cuenca» (tesis de grado, Universidad Católica de Cuenca, 2022), 62, <https://dspace.ucacue.edu.ec/handle/ucacue/13142>

⁶ Roberto Oltra-Vercher, «Materiales absorbentes acústicos basados en fibras de esparto» (tesis de máster, Universitat Politècnica de València, 2019), 7, <http://hdl.handle.net/10251/129234>

sorbe el sonido, y todo ello de forma natural, sin necesidad de usar fibras minerales artificiales.⁷ «Las fibras porosas de la corteza de la planta de cannabis la convierten en las fibras más resistentes que se conocen, es por esto que el cultivo agrícola del cáñamo industrial es uno de los más antiguos»⁸.

Según Fuentes-Pérez y Acurio-Arcos, «en Ecuador a partir del 21 de junio del 2020 se ha legalizado la siembra, cultivo y cosecha del cannabis (cáñamo industrial) de acuerdo con el Código Orgánico Integral Penal (COIP), además de su producción, industrialización, comercialización y exportación [...]»⁹. Sin embargo, no existen estudios ecuatorianos sobre la absorción de esta fibra como material para paneles acústicos. En función de esto, pensar en esta fibra como alternativa para ser utilizada es absolutamente coherente, lo que resultó en una investigación de campo innovadora para el país en la que se utilizaron fibras de cáñamo para determinar si es un material sustituible, rentable y más amigable con el medio ambiente como materia prima para la fabricación de paneles acústicos.

Adicionalmente, se realizó una visita a una plantación de cannabis y se obtuvo un dato de rentabilidad, considerando temas de inversión en el cultivo, en la extracción de fibras y en el proceso de elaboración de los paneles acústicos. Posteriormente se presentará con exactitud la forma en la que se obtuvo esta información.

⁷ Robert Schwemmer, «La paz sostenible – Concesión sostenible. SilentFiber recibe el premio German Business Awards 2022», SilentFiber, 2022, <https://www.silentfiber.net/es/nachhaltige-ruhe-nachhaltig-ausgezeichnet/?v=3fd6b696867d>

⁸ Luisa Martínez-Cubides, «Propuesta de paneles aislantes termoacústicos divisorios a base de cáñamo y resina de pino, como alternativa al uso de polímeros sintéticos en construcción» (tesis de grado, Universidad La Gran Colombia, 2021), 62, <http://hdl.handle.net/11396/6179>

⁹ Esteban Fuentes-Pérez y Liliana Acurio-Arcos, «El Cañamo (*Cannabis sativa L.*) para uso industrial y farmacéutico una visión desde la industria alimentaria», CienciAmérica: Revista de Divulgación Científica de la Universidad Tecnológica Indoamérica, vol. 9, n.º 4 (2020): 100, <https://dialnet.unirioja.es/servlet/articulo?codigo=7746466>

2. Materiales y métodos

2.1. Técnicas de levantamiento y recolección de datos

Para la presente investigación se empleó una metodología investigativa cuantitativa, ya que se obtuvo mediciones de las muestras de cáñamo y lana de roca para su comparación. Adicionalmente, se empleó el método de medida en tubo de Kundt para obtener la absorción acústica de estas.

Como punto de partida se realizó una visita a una plantación de cannabis ubicada en la ciudad de Quito, como se puede observar en la figura 1. Se recolectó información para determinar el promedio requerido para el desarrollo exitoso de una hectárea de cannabis y evaluar su rentabilidad. Además, como se indica en la figura 2, se obtuvieron tallos secos de la cosecha anterior para la extracción de fibras (demonstrado en la figura 3). Observaciones preliminares sugirieron que las fibras podrían funcionar como un material acústico adecuado debido a su estructura



Figura 2: residuos de cáñamo.

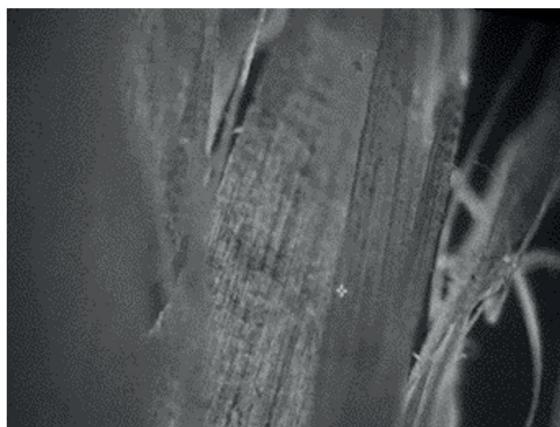


Figura 3: fibras de cáñamo vistas desde un microscopio.



Figura 1: cultivo de cáñamo, por CBD Solutions.

A continuación, como se muestra en la figura 4, se llevó a cabo la separación de las fibras de los tallos y ramas mediante un proceso de descomposición en agua. Después de varios días de descomposición, la separación de las fibras indicada en la figura 5 se volvió más sencilla. Posteriormente, las muestras se golpearon manualmente para romper las fibras de manera controlada y suavizarlas. Los elementos

más delgados que surgieron durante este

proceso se separaron manualmente, como se señala en la figura 6.



Figura 4: proceso de descomposición tallos de cáñamo.



Figura 5: separación de las fibras del tallo.



Figura : fibras de cáñamo.

La cantidad necesaria de fibra se conformó en una muestra compacta y se dejó secar, demostrado en la figura 7. Después de varios días, se cortó la muestra a un tamaño óptimo para el tubo de impedancia-Kundt utilizando un bisturí, como se evidencia en las figuras 8 y 9. Se tomaron las mediciones correspondientes con el método de función de transferencia siguiendo la norma ISO 10534-2:1998 con el fin de determinar el coeficiente de absorción acústica mediante el tubo de impedancia, dos posiciones de micrófono y un sistema digital de análisis de señal (ISO 10534-2:1998). Además, se agregó una muestra adicional de lana de roca con una densidad de 40 kilogramos por metro cúbico para fines de comparación, ya que es uno de los materiales más utilizados en este medio.



Figura 7: muestra de cáñamo.

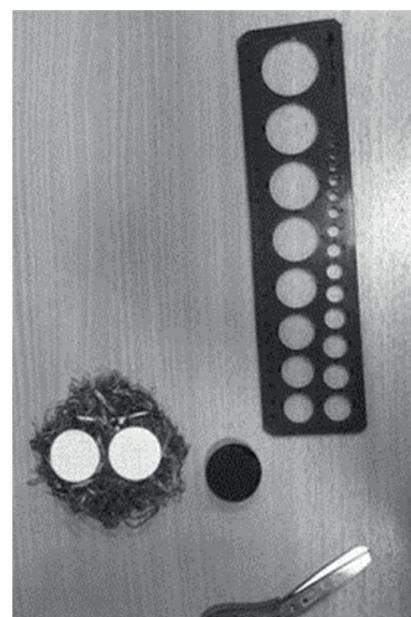


Figura 8: recorte de la muestra para el tubo de Kundt.

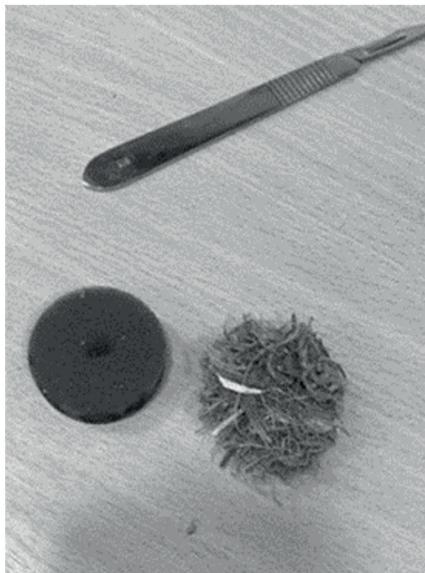


Figura 9: muestra de cáñamo recortada.

2.2. Instrumento de medición o recolección

Para calcular el nivel de rentabilidad, se consultó al propietario de la plantación y a profesionales del área. Se realizaron preguntas generales relacionadas con el número de personas que trabajan por hectárea, los costos de siembra y riego, y el tiempo de cada cosecha.

Para determinar el coeficiente de absorción del cannabis se empleó un tubo de impedancia en el laboratorio de acústica de la Universidad de Las Américas. Para ello, se llevaron a cabo varios pasos que tuvieron en cuenta la cantidad de componentes involucrados. La calibración adecuada del tubo fue crucial para obtener resultados lo más cercanos posible a la realidad.

2.3. Tubo de impedancia-Kundt (Figura 12)

Componentes

- Computador con el programa ACU-PRO

- 1 interfaz Data Translation dt9837a
- 2 cables BNC
- 1 cable BNC to RCA
- 1 cable USB 2.0
- 1 amplificador de Audio Pyle pta2
- 2 micrófonos de precisión Grass
- 1 tubo de impedancia-Kundt Spectro-nics ACUPRO
- Altavoz
- Porta muestras
- Pistón rígido
- Tubo
- Adaptadores de micrófonos

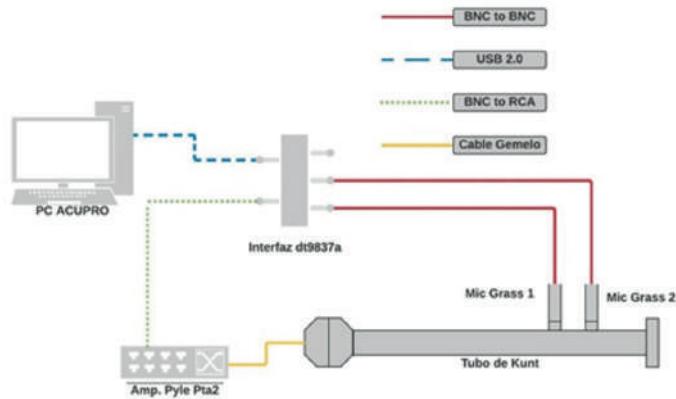


Figura 10: diagrama de conexión tubo de Kundt, por Velasco Nicolalde Juan Manuel.

2.4. Calibración

La calibración se lleva a cabo con la ayuda del software ACUPRO. A continuación, se detallan los pasos a seguir:

- Insertar el micrófono dentro de su respectiva cavidad en el tubo de impedancia.
- Seleccionar un material bastante absorbente (material de muestra/esponja de 25 mm de grosor) e insertarlo en la cavidad de muestras.
- Fijar la cavidad de muestras en su respectivo extremo del tubo de impedancia.
- Conectar la salida D/A del DT9835A al driver de compresión como se muestra en la figura 11:

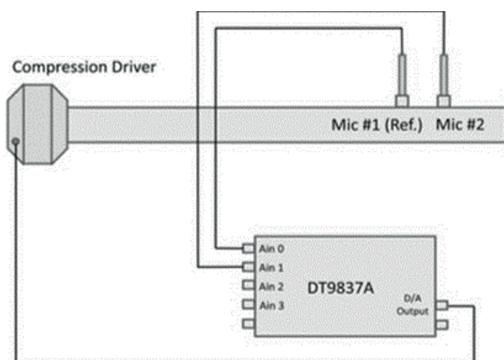


Figura 11: diagrama para la calibración del tubo de Kundt, por Velasco Nicolalde Juan Manuel.

- Conectar la salida de los micrófonos a las entradas Ain 0 para el micrófono 1 (el más cercano al driver) y Ain 1 para el micrófono 2 (Figura 13).
- Dentro del software seleccionar la opción «Calibration 1» de la ventana Ab. Meas. Options en la pestaña Calibration/TF Measurement.
- Una vez dentro de la opción mencionada, seleccionar al menos 100 promedios y presionar el botón «START» (una vez terminada la medición seleccionar «OK» para guardar los resultados).
- Intercambiar la posición (no las conexiones del DT9835A) de los micrófonos, de forma que el micrófono 1 esté ahora en la posición del micrófono 2 y viceversa (Figura 14).
- Seleccionar la opción «Calibration 2» de la ventana Ab. Meas. Options en la pestaña Calibration/TF Measurement, y dar nuevamente al botón «START».
- Guardar los resultados y regresar los micrófonos a su posición original.



Figura 12: tubo de Kundt.

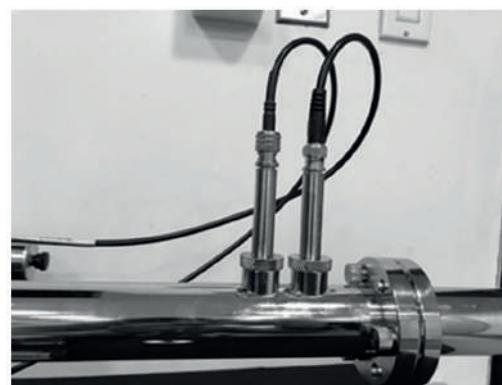


Figura 13: micrófonos colocados para medir absorción.

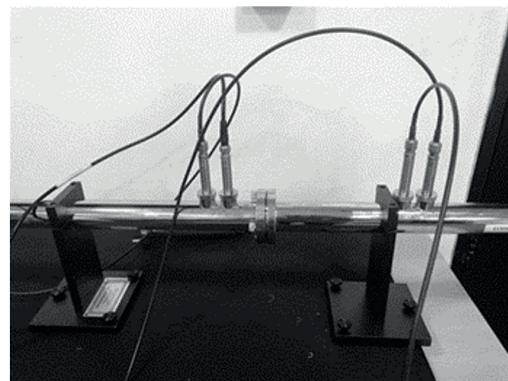


Figura 14: micrófonos insertados para medir transmisibilidad.

2.5. Método de organización de datos

Para la recolección de datos en el laboratorio, se tomaron dos muestras de cannabis y una muestra de lana de roca con una densidad de 40 kg por metro cúbico. La primera muestra de cannabis tenía un grosor de 0.7 mm, y se realizaron siete mediciones en diferentes instancias, con el tubo de impedancia siempre calibrado (Figura 15).



Figura 15: muestra de cáñamo en el tubo de Kundt.

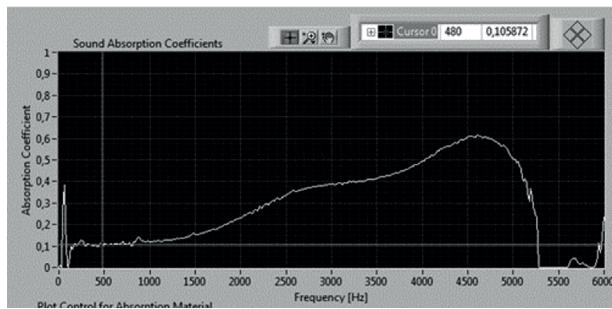


Figura 16: absorción acústica en ACUPRO, muestra 7 mm.

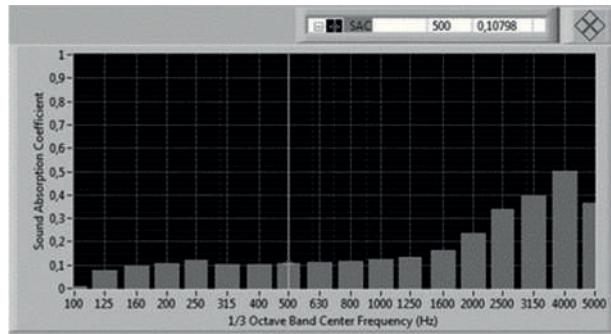


Figura 17: 1/3 de octava absorción acústica en ACUPRO, muestra 7 mm.

Los resultados mostraron que el coeficiente de absorción del cannabis se encuentra en valores medianos de frecuencia, con un pico en 4000 Hz y un coeficiente de absorción de 0.6.

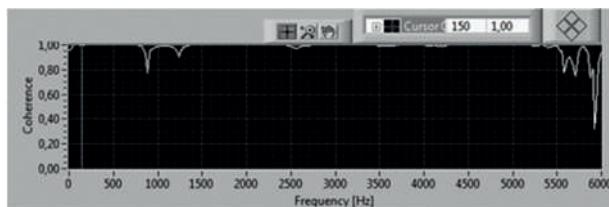


Figura 18: coherencia de la medición, muestra 7 mm.

La coherencia de los datos fue alta, aproximándose al 100 % para todas las frecuencias, excepto para 6000 y 1000 Hz, que mostraron niveles aceptables y validaron la medición. Los siete datos obtenidos de la primera muestra fueron promediados aritméticamente utilizando Excel.

Luego, se repitió el procedimiento con una segunda muestra de cannabis de 0.4 mm de grosor, y también se tomaron diferentes valores de tiempo para cada medición de manera aleatoria.

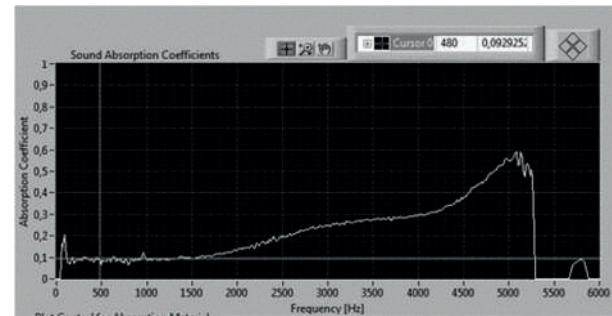


Figura 19: absorción acústica en ACUPRO, muestra 4 mm.

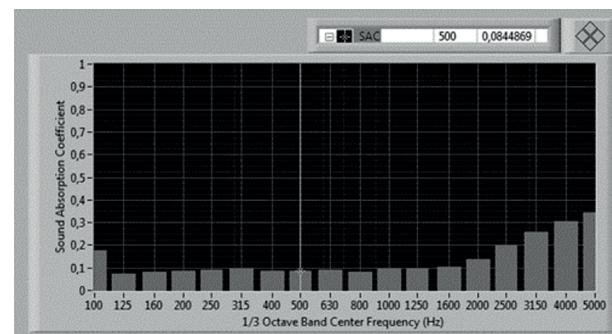


Figura 20: 1/3 de octava absorción acústica en ACUPRO, muestra 4 mm.

La muestra de 0.4 mm de grosor mostró un pico en 5000 Hz y una disminución en todas las frecuencias. Esto puede deberse a que el material más delgado absorbe principalmente frecuencias más agudas, lo que sugiere que una muestra más gruesa o densa podría absorber una mayor cantidad de frecuencias.

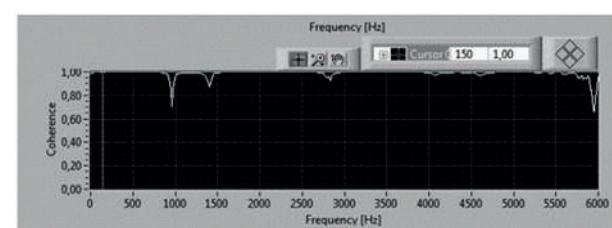


Figura 21: coherencia de la medición, muestra 4 mm.

La medición de la lana de roca de 40 kg por metro cúbico se llevó a cabo de manera similar (Figura 22) con el objetivo de compararla con el cannabis, pues es uno de los materiales más utilizados en el tratamiento acústico. Los resultados indicaron que la lana de roca tenía un coeficiente de absorción mucho mayor en comparación con el cannabis; mostraron

también un pico en 4000 Hz y una diferencia de 0.3 en Alpha.



Figura 22: muestra de lana de roca en el tubo de Kundt.

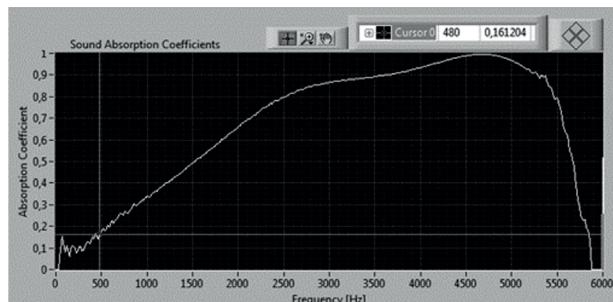


Figura 23: absorción acústica en ACUPRO, muestra lana de roca.

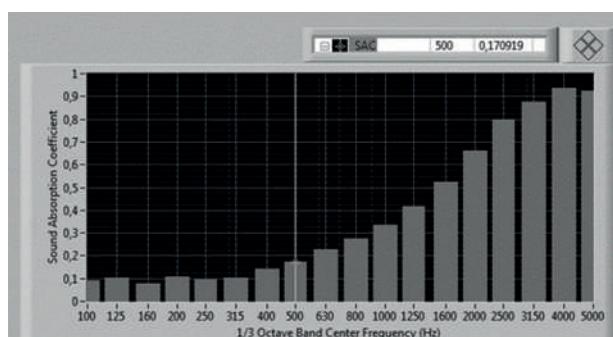


Figura 24: 1/3 de octava absorción acústica en ACUPRO, muestra lana de roca.

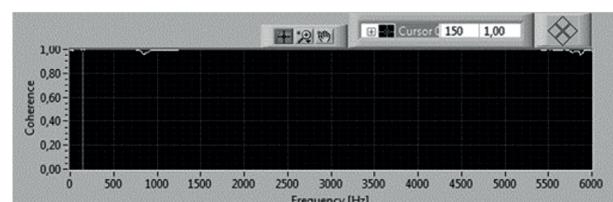


Figura 25: coherencia de la medición, muestra lana de roca.

Para complementar los datos, se utilizó el tubo de impedancia para medir la pérdida de transmisibilidad del sonido a través de ambos materiales, lo que podría ser relevante para futuras consideraciones sobre

su uso como material de aislamiento. Debido a la porosidad del material, se esperaba una pérdida de transmisibilidad mínima, lo que se comprobó mediante tres mediciones para cada muestra (un total de nueve mediciones). Los resultados se calcularon mediante un promedio logarítmico, utilizando decibelios como unidad de medida.

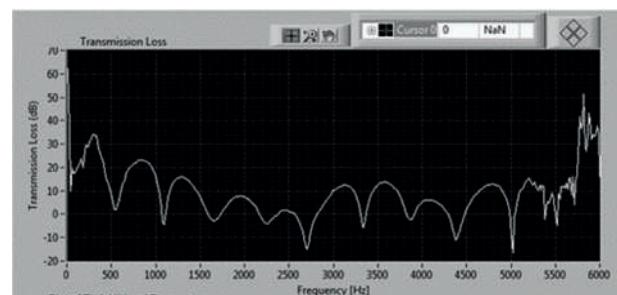


Figura 26: transmisibilidad del cáñamo.

Los datos obtenidos para el coeficiente de absorción de ambas muestras de cannabis y de lana de roca, así como su pérdida de transmisibilidad, serán explicados más adelante.

3. Resultados

En este estudio se realizaron mediciones y análisis para evaluar el potencial del cannabis como material absorbente acústico para reducir el ruido en interiores. Con ese fin, se empleó un tubo de impedancia no invasivo y no destructivo para determinar las propiedades acústicas y eléctricas de muestras de cannabis, así como de una muestra de lana de roca para comparación.

Los resultados de las mediciones del coeficiente de absorción del cannabis mostraron que las propiedades acústicas del material se ubicaron en valores medianos de frecuencia. Específicamente, se identificó un pico de absorción en torno a los 4000 Hz, con un coeficiente de absorción promedio registrado de 0.6. La coherencia de los datos obtenidos fue alta, con aproximadamente un 100 % para la mayoría de las fre-

cuencias, salvo para 6000 y 1000 Hz, donde los niveles de absorción fueron igualmente aceptables, validando así las mediciones realizadas.

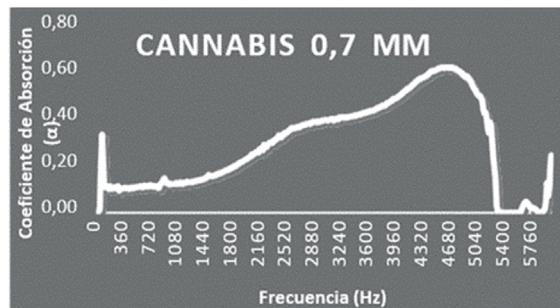


Figura 27: absorción acústica promediada, muestra 7 mm.

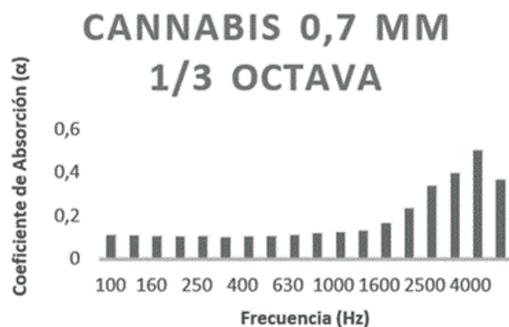


Figura 28: 1/3 de octava absorción acústica promediada, muestra 7 mm.

Por otro lado, se procedió a medir una muestra de cannabis de menor grosor, con un espesor de 0.4 mm. Los resultados de esta muestra mostraron un pico de absorción en torno a los 5000 Hz, con una disminución general en todas las frecuencias. Esta disminución podría atribuirse a que el material más delgado tiende a absorber predominantemente frecuencias agudas, sugiriendo que muestras de mayor grosor o densidad podrían absorber una mayor cantidad de frecuencias sonoras.

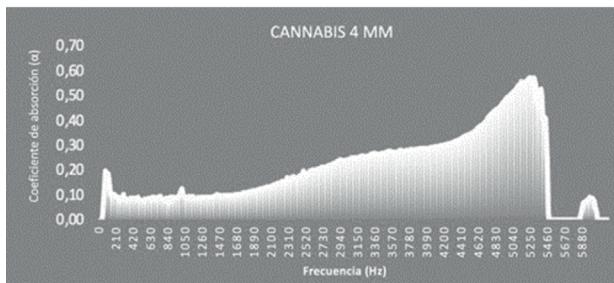


Figura 29: absorción acústica promediada, muestra 4 mm.

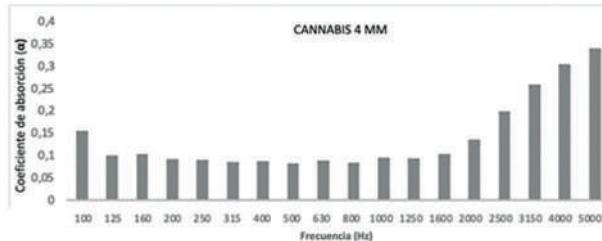


Figura 30: 1/3 de octava absorción acústica promediada, muestra 4 mm.

Para comparar con el cannabis, se realizó una medición de la lana de roca, un material ampliamente utilizado en el tratamiento acústico. Los resultados indicaron que la lana de roca exhibía un coeficiente de absorción mucho mayor al cannabis. Al igual que el cannabis, la lana de roca también mostró un pico de absorción en torno a los 4000 Hz, pero su coeficiente de absorción promedio fue de 0.9, lo que indica una mayor capacidad para absorber el sonido en este material.

Además de la medición del coeficiente de absorción, se llevó a cabo una medición de la pérdida de transmisibilidad del sonido a través de ambos materiales, como parte de las consideraciones para su uso como materiales de aislamiento. La pérdida de transmisibilidad del sonido resultó ser mínima para ambas muestras, lo cual es alentador, ya que sugiere que el cannabis, al igual que la lana de roca, podría ser una opción efectiva para reducir la transmisión de ruido entre espacios.

Tabla I
Costos de las muestras de cáñamo

1 costal /\$3	Longitud muestra	Superficie de la muestra	Total costal
35 muestras	5x5 cm	25 cm ²	8,75 m ²

En la tabla I se muestra el precio de un costal de tallos de cáñamo, la longitud de las muestras y su superficie, así como del total que se podría llegar a obtener de un costal. Así, se pudieron obtener los siguientes datos en cuanto al impacto económico, con-

siderando que Ecuador, por su geografía, es capaz de producir hasta cuatro cosechas anuales.

Tabla II
Cantidad de plantas y dinero entrante por las muestras

Plantas	Superficie	Una hectárea	Cantidad de costales	Dinero entrante
4	1 m ²	40 000	10 000	\$30 000

En la tabla II se muestra que en una superficie de 1 m² se puede obtener 4 plantas, es decir, 40 000 plantas por hectárea, lo que equivale a alrededor de 10 000 costales y una entrada de \$30 000 para la industria del cáñamo en el Ecuador.

Los resultados obtenidos revelaron múltiples beneficios medioambientales asociados con los cultivos de cáñamo. A continuación, se presentan las principales ventajas identificadas:

- Bajo requerimiento de agua

El cáñamo ha demostrado ser una planta resistente que requiere menos agua en comparación con cultivos tradicionales como el algodón. Esta característica reduce significativamente la presión sobre los recursos hídricos, lo que es especialmente relevante en regiones con escasez de agua y en un contexto de cambio climático.

- Reducción del uso de pesticidas

Las propiedades naturales del cáñamo lo hacen resistente a plagas y enfermedades, lo que reduce la necesidad de utilizar pesticidas y productos químicos agrícolas. Esta disminución en el uso de agentes químicos minimiza la contaminación del suelo y el agua, promoviendo la salud de los ecosistemas circundantes.

- Fortalecimiento de la fertilidad del suelo

El cáñamo posee un crecimiento vigoroso y un sistema de raíces extenso, lo que favorece la estructura del suelo y la retención de nutrientes. Además, su capacidad de fijar dióxido de carbono en el suelo contribuye a la mejora de la calidad del terreno y a la mitigación del cambio climático.

- Contribución como sumidero de carbono

El cáñamo actúa como un sumidero de carbono, capturando CO₂ durante su crecimiento y almacenándolo en su biomasa. Esto ayuda a reducir la concentración de gases de efecto invernadero en la atmósfera, contribuyendo a la lucha contra el calentamiento global.

- Menor huella de carbono en su producción

Comparado con otros materiales como el plástico o el acero, la producción de cáñamo requiere menos energía y genera menos emisiones de carbono. Además, los productos derivados del cáñamo son biodegradables, lo que contribuye a reducir los residuos y su impacto ambiental.

- Diversidad de usos y productos derivados

El cáñamo ofrece una amplia variedad de aplicaciones industriales, lo que promueve una economía circular y reduce la dependencia de recursos no renovables. Su uso en diferentes sectores contribuye a una mayor sostenibilidad y resiliencia económica.

- Conservación de la biodiversidad

Los cultivos de cáñamo pueden integrarse en sistemas agrícolas mixtos y agroecológicos, fomentando la biodiversidad agrícola

y protegiendo las especies autóctonas. Esto es esencial para el mantenimiento de ecosistemas saludables y la protección de los polinizadores.

- Capacidad de biorremediación

El cáñamo tiene la capacidad de absorber metales pesados y otras sustancias tóxicas del suelo, un proceso conocido como fito-rremediación. Esta propiedad lo convierte en una opción potencial para descontaminar suelos contaminados y revitalizar terrenos degradados. Por el contrario, la lana de roca inicia con la extracción y trituración de rocas volcánicas o basalto, seguido de su fusión a elevadas temperaturas. Las rocas fundidas son convertidas en fibras mediante un proceso de hilado, y luego se incorporan aglutinantes y aditivos químicos para mejorar sus propiedades acústicas y mecánicas. Posteriormente, la lana de roca es moldeada en forma de paneles o rollos, listos para ser utilizados en aplicaciones de aislamiento y acondicionamiento acústico en la industria de la construcción. El costo de la energía requerida para el proceso de fusión y la incorporación de aditivos químicos puede representar una parte significativa del costo de fabricación. Asimismo, la inversión en maquinaria y tecnología avanzada para lograr una alta calidad y eficiencia en el proceso de producción también incide en el costo final del producto.

3.1. Principal problema de la lana de roca

El principal desafío asociado con la lana de roca es su impacto ambiental durante el proceso de fabricación. Como se mencionó previamente, la producción de lana de roca consume grandes cantidades de energía y conlleva emisiones de gases de efecto invernadero, contribuyendo al cambio climá-

tico y a la contaminación atmosférica. Adicionalmente, el proceso de fusión de rocas volcánicas libera dióxido de azufre y otros contaminantes atmosféricos, lo que puede tener efectos adversos en la calidad del aire y la salud de las personas que residen cerca de las plantas de producción. Aunado a los impactos ambientales y de salud, la extracción de rocas volcánicas para la producción de lana de roca puede tener consecuencias negativas en el paisaje natural y los ecosistemas locales.

En síntesis, el costo de fabricación de la lana de roca puede ser significativo debido a la complejidad y la intensidad energética del proceso de producción. No obstante, el principal inconveniente asociado con la lana de roca radica en su impacto ambiental y el posible riesgo para la salud debido al uso de químicos durante la fabricación y al consumo de energía en el proceso. Estos problemas han generado una búsqueda continua de alternativas más sostenibles y respetuosas con el medio ambiente en el ámbito del acondicionamiento acústico, como el empleo del cáñamo como material absorbente acústico.

Los resultados de este estudio indican que el cáñamo ofrece ventajas económicas significativas en comparación con la lana de roca. En términos de costos de producción, el cáñamo es una planta de rápido crecimiento y alta productividad, lo que reduce los costos asociados a su cultivo y procesamiento. Además, el cáñamo requiere menos insumos como pesticidas y fertilizantes, lo que disminuye los costos de producción en comparación con la lana de roca.

La disponibilidad y sostenibilidad del cáñamo también juegan un papel importante en su ventaja económica. Su cultivo puede realizarse en cortos períodos de tiempo, lo que asegura una mayor disponibilidad del material en el mercado. Además, el cultivo de cáñamo tiene un menor impacto ambiental en comparación con la extracción de minerales para la fabricación de

lana de roca, lo que lo posiciona como una opción más sostenible y eco amigable.

El cáñamo también ofrece beneficios económicos en términos de transporte y logística. Dado que es más liviano que la lana de roca, los costos de transporte son menores, lo que puede reducir los gastos asociados a la distribución de los materiales. Asimismo, su naturaleza más maleable facilita su instalación, lo que puede traducirse en una disminución de los costos de mano de obra y tiempo de instalación. Además, el uso del cáñamo como material absorbente acústico tiene el potencial de generar nuevas oportunidades de empleo en la cadena de producción, desde el cultivo hasta la fabricación y la instalación de los paneles.

Al comparar la fibra de cáñamo con otras fibras naturales, como la totora, cabuya y yute, se evidencia que estas fibras tienen un comportamiento óptimo a partir de 1000 Hz. En ese nivel crece notablemente su coeficiente de absorción, pues es un comportamiento típico de las fibras ser más absorbentes en altas frecuencias por su porosidad con pérdida de energía por calor.¹⁰ Estos autores aportan con datos sobre los coeficientes de absorción de fibras naturales de 1.61 cm, 1.5 cm y 1.97 cm de espesor, respectivamente:

5.2.2 TOTORA

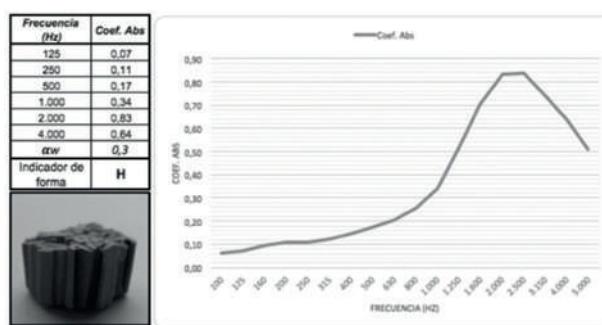


Figura 31: coeficiente de absorción de totora obtenido de Cazorla y Córdova.¹¹

10 Marcela Cazorla y Ricardo Córdova, «Caracterización de la absorción acústica de materiales para la construcción fabricados en Ecuador» (tesis de grado, Universidad de Las Américas, 2017), 33, <http://dspace.udla.edu.ec/handle/33000/8984>

11 Cazorla y Córdova, «Caracterización de la absorción...», 34.

5.2.4 FIBRA DE CABUYA

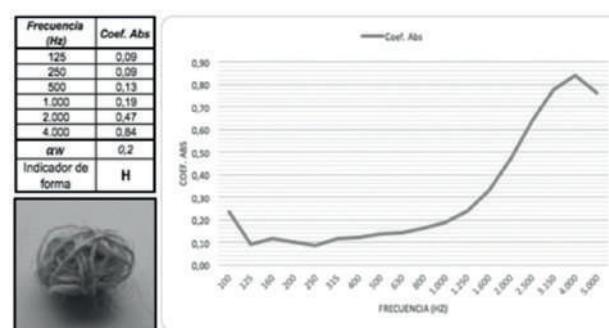


Figura 32: coeficiente de absorción de cabuya obtenido de Cazorla y Córdova.¹²

5.2.5 FIBRA DE YUTE

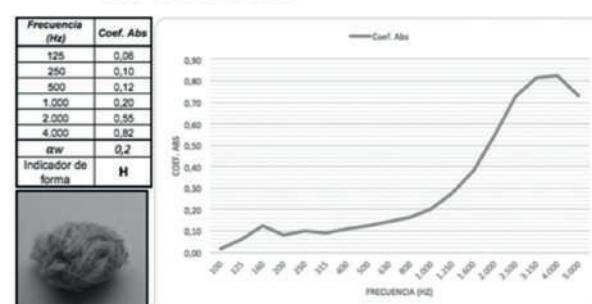


Figura 33: coeficiente de absorción de yute obtenido de Cazorla y Córdova.¹³

A partir de los datos obtenidos, se afirma que mientras mayor sea la frecuencia, mayor será el coeficiente de absorción en las fibras naturales. Por otra parte, al analizar el comportamiento de la fibra de totora (Figura 31), tomando en cuenta las mediciones realizadas y considerando que, a mayor densidad el coeficiente de absorción del cannabis aumenta, se puede deducir que al tener una muestra de igual densidad que la muestra de totora, el cannabis sería un mejor absorbente acústico.

4. Conclusiones

Los resultados de este estudio demuestran que el cannabis presenta propiedades acústicas interesantes y puede considerarse como

12 Cazorla y Córdova, «Caracterización de la absorción...», 36.

13 Cazorla y Córdova, «Caracterización de la absorción...», 34.

un material absorbente acústico viable. Aunque su coeficiente de absorción promedio fue inferior al de la lana de roca, el cannabis sigue siendo una alternativa económica y sostenible para reducir el ruido en interiores, especialmente en aplicaciones donde se requiere un tratamiento acústico más moderado. Su utilización como material absorbente acústico podría tener un impacto positivo en la industria de la construcción y en el ámbito de aislamiento o acondicionamiento acústico en la arquitectura, ya que sus propiedades le atribuyen un gran potencial.

Es importante destacar que este estudio representa una investigación inicial sobre las propiedades acústicas del cannabis en el contexto ecuatoriano, y se requieren estudios adicionales y más detallados para comprender completamente su viabilidad y aplicaciones prácticas en el campo de la acústica. Los resultados aquí presentados abren la puerta a nuevas investigaciones y al desarrollo de soluciones acústicas innovadoras y sostenibles basadas en el uso del cannabis como material absorbente.

Referencias

- Ángeles-López, Guadalupe, Fernando Brindis, Sol Cristians-Niizawa y Rosa Ventura-Martínez. «Cannabis sativa L., una planta singular». *Revista Mexicana de Ciencias Farmacéuticas*, vol. 45, n.º 4 (2014): 1-6. https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1870-01952014000400004
- Cazorla, Marcela y Ricardo Córdova. «Caracterización de la absorción acústica de materiales para la construcción fabricados en Ecuador». Tesis de grado, Universidad de Las Américas, 2017. <http://dspace.udla.edu.ec/handle/33000/8984>
- Dahal, Raj Kumar, Bishnu Acharya y Animesh Dutta. «Mechanical, Thermal, and Acoustic Properties of Hemp and Biocomposite Materials: A Review». *Journal of Composites Science*, vol. 6, n.º 12 (2022); 1-21. <https://doi.org/10.3390/jcs6120373>
- Fernea, Raluca, Daniela Tamas-Gavrea, Daniela Manea, Ioan Rosca, Claudiu Aciu y Constantin Munteanu. «Multicriteria Analysis of Several Acoustic Absorption Building Materials Based on Hemp». *Procedia Engineering* 181 (2017): 1005-1012. <https://doi.org/10.1016/j.proeng.2017.02.500>
- Fuentes-Pérez, Esteban y Liliana Acurio-Arcos. «El Cañamo (*Cannabis sativa L.*) para uso industrial y farmacéutico una visión desde la industria alimentaria». *CienciAmérica: Revista de Divulgación Científica de la Universidad Tecnológica Indoamérica*, vol. 9, n.º 4 (2020): 99-106. <https://dialnet.unirioja.es/servlet/articulo?codigo=7746466>
- Gallegos-Dávila, Henry. «Aplicaciones de cáñamo como alternativa rentable a la reactivación económica de ecuador tras la pandemia de covid-19». *Perfiles*, vol. 1, n.º 25 (2021): 45-53. <https://www.10.47187/perf.v1i25.112>
- Hash Museum. «Fibra de cáñamo industrial». Hash Museum, 2014. <https://hashmuseum.com/es/conocimiento-del-cannabis/canamo-industrial/fibra-de-canamo-industrial/>
- ISO 10534-1: 1998. Acoustics — Determination of sound absorption coefficient and impedance in impedance tubes — Part 2: Transfer-function method. <https://www.iso.org/standard/22851.html>
- Martínez-Cubides, Luisa. «Propuesta de paneles aislantes termoacústicos divisorios a base de cáñamo y resina de pino, como al-

- ternativa al uso de polímeros sintéticos en construcción». Tesis de grado, Universidad La Gran Colombia, 2021. <http://hdl.handle.net/11396/6179>
- Muriel-Páez, Miguel y Marco Pullas. «El cáñamo, una fibra textil sostenible». Ciencia Latina Revista Científica Multidisciplinaria, vol. 6, n.º 6 (2023): 12155-12182. https://doi.org/10.37811/cl_rcm.v6i6.4246
- Nath, Mausum Kumar. «Benefits of Cultivating Industrial Hemp (*Cannabis sativa* ssp. *sativa*)—A Versatile Plant for a Sustainable Future». *Chemistry Proceedings*, vol 10, n.º 1 (2022): 3-6. <https://doi.org/10.3390/IOCAG2022-12359>
- Oltra-Vercher, Roberto. «Materiales absorbentes acústicos basados en fibras de esparto». Tesis de máster, Universitat Politècnica de València, 2019. <http://hdl.handle.net/10251/129234>
- Psyckus. «La gente no sabe. Cannabis». Desmotivaciones.es. 27 de agosto de 2012. <https://desmotivaciones.es/5848292/La-gente-no-sabe>
- Rivas, Jonnathan e Inés Torres. «Paneles de revestimiento acústico. Uso de materiales compuestos reforzados con fibras de totora adaptados al centro pedagógico de la ciudad de Cuenca». Tesis de grado, Universidad Católica de Cuenca, 2022. <https://ds-pace.ucacue.edu.ec/handle/ucacue/13142>
- Schwemmer, Robert. «La paz sostenible - Concesión sostenible. SilentFiber recibe el premio German Business Awards 2022». SilentFiber, 2022. <https://www.silentfiber.net/es/nachhaltige-ruhe-nachhaltig-ausgezeichnet/?v=-3fd6b696867d>
- Tapia-Gavilanes, David Fernando. «Caracterización de materiales absorbente y modelos matemáticos». Tesis de máster, Universidad Politécnica de Madrid, 2019. <https://oa.upm.es/55263/>



FUERA DE TONO



Vol. 3 N.º 9 / Guayaquil
mayo 2024
ISSN 2953-657x

Epílogo

Renato Zamora Arízaga
Chair AES Ecuador
Co-Chair AES-LAC 2023

El tiempo ha transcurrido desde la primera iniciativa voluntaria que buscó reunir a profesionales, estudiantes y entusiastas del audio, en 2012, para crear oficialmente una sección profesional de AES en Ecuador. La labor, que parecía un sueño complejo, se lo gró gracias al compromiso de cada uno de los integrantes, quienes buscábamos un capítulo oficial de AES en el Ecuador que nos acerque a la investigación, el desarrollo y la innovación del audio y la industria. Conseguimos nuestro reconocimiento oficial como sección en 2018 e inscribimos la primera directiva, convirtiéndonos en la sección AES oficial más joven de la región.

En este período de actividad, hemos logrado materializar muchos de los sueños que tuvimos. Hoy contamos con nuestra persona jurídica gracias al respaldo del Ministerio de Cultura y Patrimonio del Ecuador. Esta entidad nos ha brindado la posibilidad de establecer la administración de convenios y alianzas estratégicas con empresas e instituciones que han ampliado el ámbito de acción de AES Ecuador en todo el país y la región. A través de estas acciones hemos llevado a cabo, de manera ininterrumpida, conferencias, capacitaciones y eventos participativos de índole académica y técnica que han contribuido significativamente a la consolidación y fortalecimiento de nuestra sección.

Ecuador se encuentra en una región privilegiada dentro del continente americano, con una extensa biodiversidad y una heterogénea construcción pluricultural que lo convierten en un lugar único en el mundo. Teníamos muchas expectativas de mostrar nuestro país a toda la región y esto nos impulsó a aceptar el desafío de ser la sede y asumir la responsabilidad de llevar a cabo la Conferencia Latinoamericana de Audio Engineering Society AES-LAC 2023, del 18 al 25 de septiembre de ese año, en Quito.

Su ejecución requirió una ardua planificación. El comité organizador llevó a cabo un esfuerzo constante durante más de un año y medio para asegurar su éxito. La sede oficial fue el campus Granados de la Universidad de Las Américas en Quito y se formó un equipo multidisciplinario que estableció acuerdos con empresas, instituciones académicas y la industria ecuatoriana, así como con expositores, ponentes y patrocinadores.

El evento se enfocó en la realización de conferencias, paneles de discusión, talleres y demostraciones prácticas. Los temas abordados incluyeron la incorporación de las tecnologías innovadoras en la producción de audio, la evolución de los sistemas y dispositivos de grabación y la ejecución de sonido directo.

Reunió a profesionales y expertos en audio de todo el mundo para abordar las últimas tendencias en la tecnología, la investigación y el desarrollo de audio y la industria. Los participantes tuvieron la oportunidad de adquirir conocimientos sobre nuevas técnicas de grabación de audio y técnicas avanzadas de mezcla en diversos talleres. Uno de los espacios más interesantes del evento fue la sala de audio inmersivo, coordinada por la empresa Genelec y su distribuidor local EASA, que permitió poner en práctica las nuevas técnicas y tecnologías de vanguardia de este medio en la industria, con expositores disertantes de varios países.

La feria de tecnología y presentaciones de productos permitió a los asistentes ver y probar las últimas tendencias en dispositivos y equipos, como sistemas de so-

nido de alta fidelidad y software para hacer música. La conferencia también tuvo lugar con una amplia variedad de ponentes destacados en el ámbito del audio, como el ingeniero Bob McCarty, de la empresa Meyer Sound, y el ingeniero ganador del Grammy Ken Lewis, quienes compartieron su trayectoria y conocimientos con los asistentes.

La conferencia fue un éxito rotundo y brindó a los asistentes la posibilidad de adquirir conocimientos sobre las últimas tendencias en la tecnología de audio, así como de conectarse y colaborar con otros expertos en la industria. La ciudad de Quito, brindó un lugar idóneo para el evento, enriqueciendo la vivencia con su vasto pasado y vibrante cultura.

Como Chair de AES Ecuador y Co-Chair de AES-LAC 2023, cargo que me cupo el honor de presidir, siento una gran alegría por esta publicación, que constituye la memoria académica y el testimonio científico de la conferencia. Es el resultado de la reflexión e investigación realizada por todos los autores, quienes han desempeñado un papel fundamental. Es un privilegio contar con este instrumento que contribuye al conocimiento de la investigación y a la ejecución de música profesional del Ecuador en el ámbito internacional. Extiendo mi profundo agradecimiento a Diego Benalcázar, Chair del comité científico de AES Ecuador y editor de esta publicación, por su excelente trabajo en su realización, y un agradecimiento especial a la Universidad de las Américas por el apoyo brindado para ser la sede de la conferencia y a la Universidad de las Artes del Ecuador por acoger esta obra para exponer el con-

FUERA
DE
TONO



Vol. 3 N.º 9 / Guayaquil
mayo 2024
ISSN 2953-657x

Memorias AES LAC 2023

Juan Jiménez
Co-Chair AES LAC 2023

En septiembre de 2022 se anunció en la AES LAC de Buenos Aires, Argentina, que la sede oficial de la AES LAC 2023 sería Quito. AES Ecuador, después de un proceso de selección, sería la encargada de la producción del evento más importante de audio y sonido de toda la región, que además celebraba los 75 años de fundación de la Audio Engineering Society. Desde ese momento, junto con Renato Zamora, tomamos el reto como Co-Chairs de la conferencia y nos pusimos a trabajar. La primera acción fue reunirnos junto a Jorge Azama, vicepresidente de AES para Latinoamérica, y comenzar a trazar un cronograma de lo que vendría para todo el año de trabajo y de preparación. En un inicio nuestro equipo se conformó por Renato, Juan Steven Chiriboga en la tesorería y yo.

No propusimos, entonces, el primer paso: invitar a auspiciantes para poder financiar la LAC y sumar expositores de primer nivel. Teníamos definida ya la ciudad, pero aún no un recinto. Con esto y con paquetes armados nos lanzamos a vender la propuesta. Con la ayuda de María Clara Kelly, comenzamos a hacer las cartas de invitación para luego enviarlas a nuestros contactos. La primera marca que se sumó en este viaje fue Genelec, quienes ya venían apoyando a AES Ecuador en sus conferencias nacionales. Además, en 2023 estaban cumpliendo 45 años y lo conmemorarían con un *tour* alrededor del mundo; esto convirtió a la conferencia en el punto más álgido de su recorrido por la región. Poco a poco más marcas se irían incorporando a este evento, como Meyer Sound, EASA, CEDIA, IDEACUSTIK, AyR, IK Multimedia, por mencionar algunas.

Viendo cómo se venía formando la LAC, era tiempo de agregar a un par de personas más al equipo. Diego Benalcázar se unió como Chair del comité científico, encargado de todo lo que tenía que ver con investigaciones presentadas para la conferencia. Él cumpliría con la labor de realizar las bases de entrega, la selección de profesionales para la revisión por pares de las investigaciones presentadas, así como la fase de llamada, revisión y presentación de trabajos seleccionados.

Otra de las partes importantes de la conferencia es la Competencia de Grabación Estudiantil. Allí, Omar Miranda tomó la responsabilidad como Chair. La organización, la selección, la invitación de jueces y todas las etapas hasta el día de las defensas de proyectos estarían a su cargo.

El trabajo de diseño de la línea gráfica del evento estuvo a cargo de Anabel Garri-dó y Pamela Burbano. Luego, integraron al equipo a Jemina Montenegro como asistente general, quien se puso manos a la obra de inmediato y comenzó a diseñar cada arte de anuncio y la página web oficial del evento. Jemina fue una parte fundamental en este proceso, pues tenía todo claro acerca de la producción general para dar cada paso.

Después de analizar varias opciones de posibles recintos —tomando en cuenta ubicación, seguridad, espacios para conferencias y feria, cercanía con hoteles y restaurantes—, la Universidad de Las Américas (UDLA), en su sede Granados, fue la opción escogida. Nos puso a disposición sus espacios y su personal técnico, médico y de seguridad para desarrollar el evento de la mejor manera. Así, la carrera de Ingeniería en Sonido y Acústica nos abrió sus puertas y firmamos un convenio para realizar la AES LAC en sus instalaciones.

Con el recinto listo, anunciamos ofi-

cialmente el evento. La AES LAC 2023 se llevaría a cabo del 18 al 21 de septiembre (el 22 de septiembre para eventos especiales) en la Universidad de Las Américas, sede Granados en Quito, Ecuador. A la par se publicó el cronograma del Call for Papers y de la Competencia de Grabación Estudiantil. También se comenzó con la venta de entradas, algo importante para el financiamiento de la conferencia. Todo previo al anuncio de los expositores del evento; con esto generaríamos expectativa. Se realizó junto a Jemina un cronograma de publicaciones en redes sociales: teníamos listos los movimientos para los próximos meses, lo que incluiría, por supuesto, el anuncio diario de los expositores.

Los expositores estaban confirmados y el cronograma listo. Usaríamos tres auditorios, entre ellos uno enfocado netamente al audio inmersivo, patrocinado por Genelec con su tour por los 45 años. Se trata de la Sala Inmersiva Genelec, donde se montó un sistema 7.1.4 para que los participantes pudieran aprovechar al máximo su estancia en ella. Se adhirieron expositores de distintos países: Ecuador, Perú, Colombia, Bolivia, Venezuela, Argentina, Brasil, Chile, Panamá, México, Estados Unidos, Canadá, España y Alemania. Además de los auditorios, también realizamos talleres en La Increíble Sociedad y Graba Estudio, además de visitas guiadas por el Teatro Nacional Sucre, Universidad San Francisco de Quito (USFQ) y la sala de experiencias de EASA, nuestro gran aliado.

Es importante mencionar también que se realizó la feria de audio a la par. Empresas como EASA, IDEACUSTIK, Mejor Sonido, CEDIA, AyR In Ear Monitors y Black Mountain estuvieron presentes mostrando sus productos y propuestas durante los cuatro días de charlas dentro de la UDLA.

Llegó mayo y era momento de terminar de armar todo el equipo. Juan Manuel Velasco se unió como productor técnico de la conferencia y tenía como responsabilidades elegir a su equipo y encargarse de los auditorios. Por su parte, María Clara Kelly pasó a ser la encargada de la feria. Además, se sumaron estudiantes de las secciones estudiantiles AES del país como voluntarios, de esa forma participaron estudiantes de la UDLA, la USFQ, la Universidad Hemisferios (UHemisferios) y la Universidad de las Artes (UArtes).

Estábamos a pocos días de comenzar la conferencia. Teníamos confirmados los vuelos y hospedaje de los expositores e invitados especiales y estábamos listos para comenzar este gran reto. El viernes 15 de septiembre de 2023 iniciamos el montaje con todo el equipamiento necesario y, a su vez, comenzamos a recibir a los visitantes extranjeros. Todo el equipo trabajó arduamente para iniciar de manera puntual el lunes 18 de septiembre.

Llegó el gran día, comenzamos con la AES Latin American Conference 2023 en Quito. Nos estrenábamos con la presentación de *posters* en el auditorio B y la Competencia de Grabación Estudiantil en la Sala Inmersiva Genelec. Durante este día se llevarían a cabo las defensas de todas las categorías de la competencia.

A las 11:15 llegó el momento de la inauguración oficial. Sentados en la mesa directiva estaban Bruce C. Olson (presidente de AES a nivel mundial), Jorge Azama (vicepresidente de AES para Latinoamérica), Renato Zamora (Co-Chair AES LAC 2023, presidente AES Ecuador), Juan Jiménez (Co-Chair AES LAC 2023, Vice-Chair del Comité de Educación AES para Latinoamérica), Diego Buenaño (decano de la Facultad de Ingenierías y Ciencias Aplicadas de la

UDLA), Christiam Garzón (director de Ingeniería en Sonido y Acústica de la UDLA). Todos tuvimos unos minutos para dar unas palabras y dar la bienvenida a los asistentes. En este auditorio comenzaría Bruce C. Olson (Estados Unidos) con su exposición «Audio Engineering Society. 75 Years» y después el taller «Stadium Scale Stereo in the Round: Metallica M72 World Tour» junto a Bob McCarthy (Estados Unidos). A su vez, en el auditorio B tendrían sus presentaciones Ricardo Cortés (Colombia), Hernán Freire (Ecuador) y Jero Cilveti (Ecuador). Para cerrar el día se realizaron dos talleres en los estudios junto a Joel Hamilton (Estados Unidos) e Ian Corbett (Estados Unidos). Algo importante para la integración con todos los asistentes fueron las reuniones por la noche. Cada sección estudiantil era la encargada de elegir el lugar de encuentro y de ser la anfitriona; este primer día estuvo a cargo de la sección AES UArtes.

El martes 19 comenzaba en el auditorio B con la presentación de *posters*, seguido de presentaciones junto a Joffre Amaguaña (Ecuador), Janio Allauca (Ecuador), Hazel Burns (Canadá) y Juan Carlos Sotelo (Méjico). Simultáneamente, en el auditorio A, tuvimos a David Acosta (Ecuador), Ian Corbett (Estados Unidos) y Rafa Sardina (España), cuya presentación se realizó de manera virtual, y cerraríamos con el Listening Party de IK Multimedia. La Sala Inmersiva Genelec recibía presentaciones de Miguel Domínguez (España), Christopher Manhey (Chile), Francisco Bissone (Argentina), Daniel Pasquel (Ecuador) y Eric Horstmann (Alemania). Por la noche, Graba Estudio recibía el taller de Ken Lewis y la sección AES USFQ fue la anfitriona de la reunión de integración.

Durante este día también se tuvo la primera reunión con los directivos de AES,

en ella participaron Bruce C. Olson (EE. UU), Agnieszka Roginska (EE. UU), Jorge Azama (PER), Renato Zamora (ECU), Juan Jiménez (ECU), Juan Carlos Sotelo (MEX), Marcela Zorro (COL), Diego Benalcázar (ECU) y Carlos Fonseca (MEX). Esta reunión tuvo como objetivo el conversar sobre la organización de este tipo de eventos en la región, las necesidades que nuestros países presentan y el apoyo que se puede brindar desde AES HQ a Latinoamérica.

Comenzaba el tercer día de conferencia: el auditorio A recibía a Lucía Cadena (Ecuador) y junto con la comunidad SONIDAS se daría paso al conversatorio «Iniciativa de mujeres profesionales de la música y el audio, y la brecha de género en el sector». Allí participaron Lucía Villaruel (Ecuador), Caridad Espinosa (Ecuador), Bernarda Ubida (Ecuador), Jemina Montenegro (Ecuador), Emmie Stadler (Ecuador) y Flor Bonilla (Ecuador). Luego siguieron charlas junto a Luis Cabané (Chile) y Ezequiel Morfi (Argentina) junto a José Orea (México). En el auditorio B tuvimos presentaciones de Jorge Azama (Perú), Luis Soria (Bolivia), Ramiro Rodríguez (Ecuador), Hillary Van Scy (Estados Unidos) y Marcela Zorro (Colombia). En la Sala Inmersiva Genelec estaba Christopher Manhey (Chile) junto a Eric Horstmann (Alemania), Ken Lewis (Estados Unidos), Carlos Freitas (Brasil), Fernando Richard (Argentina) y Danny Dee (Brasil). En La Increíble Sociedad terminábamos el día con un taller dictado por Joel Hamilton (Estados Unidos). La sección estudiantil AES UHemisferios era nuestra anfitriona para la reunión de integración.

Este día también se llevó a cabo la reunión con los directivos regionales de AES, en la que participaron Bruce C. Olson (Estados Unidos), Agnieszka Roginska (Estados Unidos), Jorge Azama (Perú), Renato Za-

mora (Ecuador), Juan Jiménez (Ecuador), Juan Carlos Sotelo (Méjico), Carlos Fonseca (Méjico), José Orea (Méjico), Marcela Zorro (Colombia), Luca Gardani (Colombia), Iván Markovic (Argentina), Francisco Bissonne (Argentina) y Luis Soria (Bolivia). El objetivo era definir la sede de la AES LAC 2024, con la gran noticia de que pasaría a ser una Convención AES, la tercera en el mundo, contando a la Convención Mundial y a la Convención Europea.

El jueves 21 comenzó el último día de conferencia en la UDLA. El auditorio A recibió a Jesús Rubio (Venezuela), Ricardo Escallón (Colombia), Juan Sierra (Colombia), Joel Hamilton (Estados Unidos) y la ceremonia de cierre. En el auditorio B se encontraban Andrés Bracero (Ecuador), Fausto Espinoza (Ecuador), Pablo Santacruz (Ecuador), Santiago Salazar (Ecuador) y Marcelo Navía (Bolivia). En la Sala Inmersiva Genelec hubo un conversatorio de sistemas de audio inmersivo con Daniel Pasquel (Ecuador), Fernando Richard (Argentina), Carlos Freitas (Brasil), Christopher Manhey (Chile) moderado por Miguel Domínguez (España), seguido de una charla junto a Iván Markovic (Argentina), el conversatorio Genelec junto a Jero Cilveti (Ecuador) y Ezequiel Morfi (Argentina), y un taller dictado por Ken Lewis (Estados Unidos).

La ceremonia de cierre tenía preparadas algunas sorpresas. Iniciaríamos con las palabras de Renato Zamora, Co-Chair de la AES LAC 2023, y a continuación la premiación de la Competencia de Grabación Estudiantil con su Chair, Omar Miranda. Kev Santos Band puso el toque musical de la noche, interpretando canciones del género marimba fusionadas con algunos tipos de ritmos. Juan Jiménez, Co-Chair de la conferencia, daría sus palabras agradeciendo al equipo de trabajo. Genelec nos tenía preparada una

sorpresa: el sorteo de una pareja de monitores que lo llevaría a cabo Miguel Domínguez, representante de la marca. Para finalizar el evento teníamos la intervención de Jorge Azama, vicepresidente de AES para Latinoamérica, quien, además de dar por clausurada esta edición, anunciaría que la primera AES Latin American Convention 2024 se llevaría a cabo en La Paz, Bolivia. Después de esta ceremonia, nos dirigimos a la última reunión de integración organizada por la sección estudiantil AES UDLA.

La conferencia había terminado, pero el viernes 22 teníamos preparados eventos fuera de las instalaciones de la UDLA. Estos incluían una visita guiada por el Teatro Nacional Sucre a cargo de EASA y d&b Audio-technik, una visita a las instalaciones de la Universidad San Francisco de Quito a cargo del CoM USFQ y un taller dictado por Víctor Herrera (Ecuador) en la sala de experiencias EASA.

La AES Latin American Conference 2023 fue un reto que lo supimos afrontar con los más altos niveles de calidad. Estoy muy feliz con el resultado obtenido y los excelentes comentarios por parte de los asistentes. Durante una semana, el Ecuador pudo participar de la conferencia de audio y sonido más importante de la región, que contó con más de 300 asistentes. El trabajo se cumplió de la mejor manera, lo que, por supuesto, no hubiera sido posible sin el equipo que estuvo detrás por un año y cuyos integrantes merecen ser mencionados: Renato Zamora, Juan Jiménez, Jemina Montenegro, Juan Steven Chiriboga, Diego Benalcázar, Omar Miranda, María Clara Kelly, Santiago Salazar, Jorge Azama, Mayerli Chachalo, Jarek Morales, Verónica Pérez, César Ponce, Steven Copa, Martín Carvaljal, Pamela Salazar, Darío Simbaña, Eduardo Hipo, Gabriela Ushiña, Kevin Grunauer,

Deiby Veloz, Mateo Burbano, Nevir Montero, Felipe Obando, Fernando Pacheco, Consuelo Peña, Mariasol Pineda, Maximiliano Morán, Esteban Cahueñas, Sidney Calvopíñaa, Camila Velasteguí, Nemrod Reyes, José Álvarez-Torres, Devin Rea.

Y también quedo muy agradecido con los auspiciantes que se sumaron a esta gran aventura: Genelec, Universidad de Las Américas, Ingeniería en Sonido y Acústica Universidad de Las Américas, EASA, CEDIA, IDEAKUSTIC, AyR In Ear Monitors, IK Multimedia, Black Mountain, Sonarworks, Eventide, Clave Digital, Borealis Audio, Sound Dogs, SAYCE, Ableton, Meyer Sound, La Increíble Sociedad, Titánio Studios, Soria Records, IAS, Omni Soundlab, Alama Music Business, Santiago Salazar Sound, AES Colombia, Classic Master, Graba Estudio, SONIDAS, CoM USFQ, Jero Silveti Audio Freak, Universidad Hemisferios, Studio G Brooklyn, Renato Zamora, Juan Jiménez, Corporación Azende, Escuela de Artes Sonoras de la Universidad de Las Artes, Acon Digital, Martinic, Modartt, Native Instruments, iZotope, Quagliardi Pro Audio.

Y las secciones estudiantiles AES de Ecuador, que fueron protagonistas dentro de este evento: AES UDLA, AES USFQ, AES UArtes y AES UHemisferios.

Gracias a todos y todas por hacer posible este gran evento. Como Co-Chair de la AES LAC 2023 me siento inmensamente agradecido y feliz por el resultado y el aprendizaje obtenido.

¡Nos vemos en La Paz!

Memoria gráfica del evento AES LAC 2023

La Conferencia Latinoamericana de la Audio Engineering Society (AES LAC) 2023, organizada por el equipo de AES Ecuador, se llevó a cabo entre los días 18 y 21 de septiembre de 2023, en la sede Granados de la Universidad de las Américas (UDLA), Quito, Ecuador.

Fotografías de Fernando Pacheco.











Biografías de los autores

Michael Javier Acevedo Cutta

Nació el 12 de diciembre de 2001 en Yopal, Casanare, Colombia, y vivió hasta los 13 años de edad en Murcia, España. A los 10 años comenzó sus estudios de guitarra clásica. En 2014, regresó a Yopal, donde culminaría sus estudios de bachillerato. En sus dos últimos años de escuela, descubrió su interés por la electrónica, lo que lo llevó a ingresar al programa de Ingeniería de Sonido en 2020. En 2021, se unió al semillero de Salud, Discapacidad e Ingeniería gracias al trabajo desarrollado en la implementación de una consola mezcladora para acercar a los estudiantes de primer año al mundo del audio. Sus campos de interés son la producción, grabación y mezcla musical, donde ha colaborado con los docentes líderes de curso en diversas clases. Ha sido parte de la representación estudiantil de la Facultad de Ingeniería de la Universidad de San Buenaventura; vocalista principal de la banda de pop de la misma institución y ha participado como ponente en investigaciones a nivel nacional e internacional.

Isabella Álvarez Franco

Forma parte del grupo de investigación Arte y Nuevos Medios, semillero ACORDE. Estudiante de Artes Visuales del Instituto Tecnológico Metropolitano de Medellín (ITM).

Timothy Antonio Carrera Bleiksloot

Nacido el 16 de diciembre de 2001 en Quito, Ecuador. Es estudiante de ingeniería en Sonido y Acústica en la Universidad de las Américas (UDLA). Actualmente es un miembro destacado del programa de excelencia académica UDLA Honors. Su enfoque principal se centra en el diseño e implementación de tratamiento acústico, así como en el campo del sonido en vivo, donde combina su pasión por la música con su habilidad técnica para crear experiencias auditivas de alta calidad. Además, ocupa el cargo de vicepresidente de la sección estudiantil Audio Engineering Society (AES) UDLA.

Manuelita Castañeda

Graduada como técnica en Audio de la Escuela Fernando Sor y es estudiante de Ingeniería de Sonido en la Universidad San Buenaventura, Bogotá. Además de ser vocera y miembro de la seccional de AES San Buenaventura, ha consolidado una sólida trayectoria en el campo del sonido. Desde sus inicios como *roadie* en eventos en vivo hasta su experiencia en diseño sonoro, grabación, mezcla, *mastering* y

como asistente de audio en Noticias Caracol, ha cultivado su pasión por los audiovisuales, la música, la historia y la cultura. Este enfoque multifacético la motiva a explorar constantemente nuevas tecnologías y proyectos para ofrecer experiencias únicas en el ámbito del ocio en Colombia.

Julia Cubides

Estudiante de Ingeniería de Sonido en la Universidad de San Buenaventura, miembro y parte de la junta directiva de la seccional AES San Buenaventura. A muy temprana edad fue intérprete de instrumentos típicos en grupos musicales representativos de la región andina colombiana. Esa experiencia la llevó a involucrarse en el mundo del sonido en un nivel académico, donde ha descubierto su amor por la acústica, el diseño y la investigación, sin dejar de lado su parte musical, participando en las diferentes etapas de proyectos musicales y audiovisuales. La impulsa el deseo de adquirir conocimientos, comprender y participar activamente en el progreso y desarrollo de la industria del sonido en sus diversos campos.

Lina Alejandra González Ramírez

Forma parte del grupo de investigación Arte y Nuevos Medios, semillero ACORDE. Estudiante de Tecnología en Informática Musical del Instituto Tecnológico Metropolitano (ITM).

Daniel Jara-Hurtado

Oriundo de Yopal, Casanare, al oriente de Colombia. Realizó sus estudios de pregrado en la Universidad de San Buenaventura (Bogotá), en el programa de Ingeniería de Sonido. Continuó su formación como especialista en Docencia Universitaria, en la Universidad el Bosque, en la cual profundizó en el desarrollo curricular de cursos de ingeniería. Se formó en el programa de Maestría en Ingeniería Biomédica, adscrito a la Facultad de Medicina de la Universidad Nacional de Colombia (Bogotá), donde desarrolló su trabajo en las áreas de procesamiento (filtrado) y extracción de información (caracterización del movimiento) de imágenes médicas (ecocardiografías). Además, tiene una maestría en Educación, en la Universidad el Bosque. Cuenta con experiencia docente desde 2018, en programas de educación técnica y profesional y liderando semilleros y proyectos de investigación en las áreas de ciencias biomédicas, aplicaciones de métodos de inteligencia artificial para la clasificación de señales bioacústicas, desarrollo de dispositivos electrónicos y educación.

Daniel Leonardo Layton Rodríguez

Nació el 18 de octubre de 1999 en Tunja, Boyacá, Colombia. Desde muy pequeño surgió el interés por la música, y en 2016, al graduarse del Colegio de Boyacá, aprendió a interpretar la guitarra de manera autodidacta. Estudió dos semestres de Música en la Universidad Juan de Castellanos de Tunja en 2019. Debido a su deseo

de profundizar en el sonido desde el punto de vista físico y también en la electrónica aplicada al audio, en el año 2020 inició sus estudios de Ingeniería de Sonido en la Universidad de San Buenaventura en Bogotá. En 2021 se unió al semillero de Salud, Discapacidad e Ingeniería (SDI), que tiene como uno de sus enfoques el área de desarrollo de aplicaciones. Durante su proceso de formación han surgido intereses por el desarrollo de *software*, la producción audiovisual y el refuerzo sonoro.

Diego Alejandro Molina Quintero

Tutor del grupo de investigación Arte y Nuevos Medios, semillero ACORDE. Profesional en Artes de la Grabación y la Producción Musical del Instituto Tecnológico Metropolitano de Medellín (ITM). Especialista en Programación Aplicada de la Institución Universitaria Digital de Antioquia. Docente de la Facultad de Artes y Humanidades del ITM.

Jamir Mauricio Moreno Espinal

Tutor del grupo de investigación Artes y Humanidades, semillero TECNOARTE. Magíster en Gestión Cultural de la Universidad de Antioquia. Especialista en Artes, con énfasis en Composición de la Universidad de Antioquia. Docente de la Facultad de Artes y Humanidades del Instituto Tecnológico Metropolitano (ITM).

Josué Felipe Ordoñez Castro

Tiene 23 años. Disfruta de la música no solo desde el enfoque de la producción sonora, sino también como intérprete empírico de piano, guitarra y voz. Creció y se formó al sur de ciudad en el Liceo Manantial de Vida Eterna, graduado en el 2018 con uno de los mejores puntajes de la prueba de estado (ICFES) de la institución. Actualmente cursa el noveno semestre de Ingeniería de Sonido y quinto semestre de Ingeniería de Sistemas en la Universidad de San Buenaventura, sede Bogotá. Ha sido monitor en las asignaturas de Electroacústica y Dispositivos Semiconductores durante 2023. Pertenece al semillero de Salud, Discapacidad e Ingeniería (SDI) desde 2022.

Laura Sofía Pineda Gómez

Estudiante de Ingeniería de Sonido en la Universidad de San Buenaventura, Bogotá. Tiene destreza en la producción de medios sonoros y mezcla. Ha sido partícipe de eventos como el Festival de Fuego 2023 y ha trabajado en el estudio 5ta Studio. Actualmente es presidenta y miembro de la seccional AES San Buenaventura. Ha participado en proyectos de grabación, producción musical y contenido audiovisual. Fusiona su creatividad con habilidades técnicas, comprometida con la equidad de género. Aspira a ser una destacada ingeniera de sonido, dejando su huella en la industria.

Juan Diego Prieto Pedraza

Nacido en Bogotá, Colombia. Su fascinación por el mundo de la música y la matemática lo llevó a elegir la carrera de Ingeniería de Sonido, que estudia actualmente. Ha realizado cursos en producción musical y ha participado activamente en proyectos de investigación y desarrollo relacionados con la ingeniería de sonido, como el semillero de investigación en la Universidad de San Buenaventura por casi dos años.

Eric Julián Rodríguez Babativa

Nació el 11 agosto 2002, en Bogotá, Colombia. Es estudiante de Ingeniería de Sonido en la Universidad de San Buenaventura. Ha realizado estudios complementarios, por movilidad académica, en la Universidad Nacional de Tres de Febrero (Untref) en Buenos Aires, Argentina, cursando la asignatura de Electroacústica II durante el periodo 2023-2. Ha destacado como investigador en los semilleros de Tecnología en Audio (STA) entre 2021 y 2022 y en el semillero Salud, Discapacidad e Ingeniería (SDI) desde 2022. Además, se ha desempeñado como monitor en el curso de Oscilaciones y Ondas. También es vicepresidente de la seccional estudiantil AES USBBOG desde 2021.

Guía rápida para los autores

Envíos

a. Artículos de investigación (sección peer review)

Los artículos arbitrados por pares ciegos se publicarán como artículos de investigación científica original. El tratamiento del tema y su contenido deben ser de interés para el área de la producción musical y las artes sonoras. Deberán mostrar rigor académico en el abordaje metodológico y precisión analítica del enfoque temático. Los trabajos serán inéditos y novedosos, enfatizando las relaciones multi, inter y transdisciplinares en, desde y para las artes de lo sonoro.

Los artículos originales deberán contener título, un resumen estructurado de 250 palabras en español e inglés, un máximo de seis palabras claves o descriptores, cuerpo del artículo, referencias bibliográficas, y fuentes de financiamiento y/o permisos de uso de datos, con extensión total entre 5000 y 6500 palabras.

Es indispensable el anonimato de los manuscritos para garantizar la imparcialidad de los pares evaluadores. El tiempo promedio de revisión será de 20 días y se comunicará a los autores en un periodo máximo de 2 meses a partir de la recepción del artículo.

Se evaluará la calidad del artículo, los aportes al conocimiento, la actualidad de la bibliografía, la calidad y el manejo de las fuentes, la claridad en la argumentación, la coherencia en la redacción e importancia del tema. Una vez sometidos al arbitraje, los artículos podrán ser aprobados, aprobados con cambios o no aprobados. Para una segunda revisión, los autores contarán con 7 días laborales para una nueva entrega del manuscrito. Luego de recibir el artículo modificado, se le informará al autor acerca de su aprobación.

En los casos de controversia, se convocará al Comité Editorial a una reunión extraordinaria para su evaluación y viabilidad de la petición en un plazo no superior a los 8 días hábiles.

Comprobación para envíos:

- El artículo no ha sido publicado previamente ni propuesto a otro medio editorial.
- El artículo está en formato editable de procesador de texto DOC (.docx).
- El artículo cumple con los requisitos bibliográficos y de estilo señalados.
- La lista de referencias contiene un 40 % de los últimos cinco años.
- El texto no contiene referencia del autor.
- Evita el empleo de lenguaje de género excluyente.

b. Partituras de composición contemporánea

Las partituras deberán enviarse en formato PDF, en blanco/negro o color, con un máximo de 14 cuartillas en A4, incluyendo en la página inicial el título, nombre del autor y/o del transcriptor y año de composición. Podrá proponerse adicionalmente una página para

instrucciones de interpretación y/o conceptualización performativa en tipografía Helvética Neue, 11 puntos e interlineado sencillo. Los envíos de partituras no podrán contener numeración de página ni encabezado de ningún tipo. El tiempo promedio de revisión será de 40 días y se comunicará a los autores en un periodo inferior a 2 meses a partir de la recepción de las propuestas.

Se evaluará la capacidad de ejecución, los aportes científicos al lenguaje y producción musical, a la interpretación y a la composición contemporánea, la claridad del formato, la calidad y el manejo de edición y coherencia en la selección de los signos extralingüísticos sonológicos. El Consejo Asesor de la Revista Sonocordia elegirá para cada número las seleccionadas de entre aquellas presentadas por la dirección editorial.

c. Sección Miscelánea

Como miscelánea entendemos traducciones, críticas de eventos y/o festivales, reseñas de trabajos discográficos, entrevistas. Los envíos no deberán superar las 4000 palabras y deberán incluir nombre, filiación institucional, mail de contacto y tipo de vinculación con la investigación en artes sonoras y producción musical. Se tomarán en cuenta las normas de edición de los artículos de investigación científica.

Ética de la publicación científica original

Los artículos deberán ser originales, inéditos y no haber sido editados en ningún medio de publicación.

Los textos serán revisados mediante el sistema de plagio URKUND.

Para el arbitraje científico, los evaluadores señalarán sus apreciaciones en un formulario de evaluación y deberán informar al Comité Editorial los casos de conflicto de interés que pudieran ocaionarse para realizar el cambio de evaluador.

Autoría

Los autores conservan los derechos de autor y ceden el derecho de la primera publicación a la revista, registrándose en Creative Commons y permitiendo a terceros utilizar lo publicado mencionando la autoría, número de edición y fecha de publicación.

Se deberá indicar al final del manuscrito si la investigación fue financiada por parte de alguna institución, así como indicar los permisos respectivos de uso si se utilizan bases de datos que no sean de dominio público o no fueron creadas por el autor.

Normas de referencia

Las citas en el texto deben seguir el estilo [Chicago-Deusto](#). Algunos ejemplos:

Cita corta

Como afirma Maurel-Indart, «[...] la obra es la verdadera prolongación de la persona del autor, no sus avatares reproducidos en el soporte material».¹

Cita larga

Como afirma Maurel-Indart:

¹ Hélène Maurel-Indart, *Sobre el plagio* (Buenos Aires: Fondo de Cultura Económica, 2014), 205.

En el siglo XX, el principio de perpetuidad fue claramente aceptado en la legislación; pero en lugar de estar relacionado con el derecho de explotación, como lo deseaban los escritores militantes, se lo relacionó con el derecho moral. En efecto, en su inmaterialidad, la obra es la verdadera prolongación de la persona del autor, no sus avatares reproducidos en el soporte material.²

Referencia en nota (libro)

Fernando Ortega. *Amadeus. Una lectura teológica del filme de Milos Forman* (Buenos Aires: Agape, 2014): 68-69.

Referencia abreviada en nota

Ortega, *Amadeus*, 68-69.

Referencia en Bibliografía (la bibliografía va al final)

Ortega, Fernando. *Una lectura teológica del filme de Milos Forman*. Buenos Aires: Agape, 2014.

Referencia en nota (capítulo en libro)

Karina Borja. “Los paisajes vivos del equinoccio: la yumbada de San Isidro de El Inca”, en Esteban Ponce Ortiz (ed.), *Grado cero. La condición equinocial y la producción de cultura en el Ecuador y en otras longitudes ecuatoriales* (Guayaquil: UArtes Ediciones, 2016): 86-87.

Referencia en nota (artículo de revista)

Javier Marín-López y Virginia Sánchez-López. “El pensamiento musical de Pablo de Olavide y Jáuregui (1725-1803): perfiles inéditos de un peruano ilustrado”, en *Revista Musical Chilena* Vol. 74 N.º 234 (2020): julio – diciembre: 17.

Partitura

Adriana Verdié. *In “C” Nuation* (Virginia: Cayambis Music Press, 2018): 3.

Dirección de envíos y/o consultas

revista.sonocordia@uartes.edu.ec

² Hélène Maurel-Indart, *Sobre el plagio*, 205.

Una publicación de la Universidad de las Artes del Ecuador,
bajo el sello editorial UArtes Ediciones.

Julio de 2024, Guayaquil, Ecuador

Familias tipográficas: Uni Sans, Merriweather

En esta edición especial de *Sonocordia* se presenta una colección de investigaciones divulgadas en la reciente Conferencia Latinoamericana de la Audio Engineering Society en Quito, Ecuador, que se llevó a cabo en septiembre de 2023. Este encuentro reunió a expertos en la industria y la academia en el área de sonido, música y producción musical, creando un interesante espacio de discusión y divulgación de conocimiento. Los trabajos seleccionados cubren una amplia gama de temas dentro de las prácticas creativas en la producción musical, nuevas tecnologías y la acústica.

Estos trabajos subrayan la relevancia de la acústica y su aplicación práctica en la región. Esta compilación celebra la capacidad de Latinoamérica para contribuir significativamente a la discusión global sobre música y sonido, resaltando la creatividad y el espíritu innovador de nuestra comunidad científica y artística.

Diego Benalcázar

Editor invitado

