

UNIVERSIDAD DE CHILE
FACULTAD DE MEDICINA
ESCUELA DE TECNOLOGÍA MÉDICA



TESIS PROFESIONAL

Para optar al título profesional de Tecnólogo Médico con mención en
Otorrinolaringología

**“COHERENCIA ESPECTRAL DE PISTAS MUSICALES EN EL
IMPLANTE COCLEAR”**

Alumnos: Eugenio Cuevas Orellana y Juan Guzmán Cortez

Tutores: Enzo Aguilar Vidal y Carolina Espinoza Onate

Firma

Fecha: 17 de julio del 2023

UNIVERSIDAD DE CHILE
FACULTAD DE MEDICINA
ESCUELA DE TECNOLOGÍA MÉDICA



TESIS PROFESIONAL

Para optar al título profesional de Tecnólogo Médico con mención en
Otorrinolaringología

**“COHERENCIA ESPECTRAL DE PISTAS MUSICALES EN EL
IMPLANTE COCLEAR”**

Alumnos: Eugenio Cuevas Orellana y Juan Guzmán Cortez

Tutores: Enzo Aguilar Vidal y Carolina Espinoza Onate

Firma

Fecha: 17 de julio del 2023

EVALUACIÓN FINAL



Santiago, 23 de octubre de 2023

ACTA CURSO TRABAJO DE INVESTIGACIÓN

El **Sr. Eugenio Andre Cuevas Orellana**, estudiante de Tecnología Médica con Mención en Otorrinolaringología, cumpliendo con los requisitos establecidos en el plan de estudio, realizó durante el noveno semestre de la carrera, la Tesis Profesional titulada: "**COHERENCIA ESPECTRAL DE PISTAS MUSICALES EN EL IMPLANTE COCLEAR**", dirigida por el **Prof. Enzo Aguilar Vidal** académico del Laboratorio de Audiología y Percepción Auditiva del Departamento de Tecnología Médica, Facultad de Medicina de la Universidad de Chile y la **Prof. Carolina Espinoza**, académica Departamento de sonido Universidad de Chile.

La Escuela de Tecnología Médica designó para su corrección una Comisión integrada por: **TM. Alvaro del Valle**, de la Clínica Santa María y el **Prof. Cristian Aedo**, académico del Laboratorio de Audiología y Percepción Auditiva del Departamento de Tecnología Médica, Facultad de Medicina de la Universidad de Chile.

La calificación obtenida se detalla a continuación:

Corrector 1 TM. Alvaro del Valle	7.00	25%
Corrector 2 TM. Prof. Cristian Aedo	6.96	25%
Tutor(es) Guía:		
Evaluación intermedia	7.00	25%
Nota final tutor	7.00	25%
Nota final tesis profesional	6.99	

En consecuencia el estudiante **Eugenio Andre Cuevas Orellana** aprueba satisfactoriamente la asignatura.


Prof. Catalina Rita .
Coordinador(a) curso
Trabajo de Investigación


HERNAN RODRIGO
TORRES RIVERA
15.932.579-2
07/11/2023 - 15:55
ESTE DOCUMENTO CONTIENE FIRMA ELECTRÓNICA AVANZADA
<https://corpopsi.uc.cl/validacion/verificador>
CV: 654a0836ea0aed196d0bfb2

Prof. Hernán Torres.
PEC curso
Trabajo de Investigación

ACTA CURSO TRABAJO DE INVESTIGACIÓN

El **Sr. Juan Pablo Guzmán Cortez**, estudiante de Tecnología Médica con Mención en Otorrinolaringología, cumpliendo con los requisitos establecidos en el plan de estudio, realizó durante el noveno semestre de la carrera, la Tesis Profesional titulada: “**COHERENCIA ESPECTRAL DE PISTAS MUSICALES EN EL IMPLANTE COCLEAR**”, dirigida por el **Prof. Enzo Aguilar Vidal** académico del Laboratorio de Audiología y Percepción Auditiva del Departamento de Tecnología Médica, Facultad de Medicina de la Universidad de Chile y la **Prof. Carolina Espinoza**, académica Departamento de sonido Universidad de Chile.

La Escuela de Tecnología Médica designó para su corrección una Comisión integrada por: **TM. Alvaro del Valle**, de la Clínica Santa María y el **Prof. Cristian Aedo**, académico del Laboratorio de Audiología y Percepción Auditiva del Departamento de Tecnología Médica, Facultad de Medicina de la Universidad de Chile.

La calificación obtenida se detalla a continuación:

Corrector 1 TM. Alvaro del Valle	7.00	25%
Corrector 2 TM. Prof. Cristian Aedo	6.96	25%
Tutor(es) Guía:		
Evaluación intermedia	7.00	25%
Nota final tutor	7.00	25%
Nota final tesis profesional	6.99	

En consecuencia el estudiante **Juan Pablo Guzmán Cortez** aprueba satisfactoriamente la asignatura.



Prof. Catalina Rita .
Coordinador(a) curso
Trabajo de Investigación



HERNAN RODRIGO
TORRES RIVERA
15.932.579-2
07/11/2023 - 19:57
ESTE DOCUMENTO CONTIENE FIRMA ELECTRONICA AVANZADA
<https://corpapel.uchile.cl/validacion/verificador>
Cv. 654a0892ea0aedf96dbbfad

Prof. Hernán Torres.
PEC curso
Trabajo de Investigación

DEDICATORIA

A nuestra familia que nos apoyaron y acompañaron en este proceso.

A nuestros amigos y amigas que siempre nos dieron palabras de aliento.

AGRADECIMIENTOS

Quisiéramos agradecer a nuestros tutores, los profesores Enzo Aguilar y Carolina Espinoza por su colaboración a lo largo de este proceso, quienes nos dedicaron tiempo, nos proporcionaron información y orientación en diversos aspectos.

ÍNDICE

RESUMEN	6
INTRODUCCIÓN	7
Música	8
La música a través del IC	9
Distorsiones en el dominio de la frecuencia	10
Distorsiones en el dominio temporal	11
Distorsiones en el dominio de la amplitud	12
Características fundamentales de la música e IC	14
Tono	14
Ritmo	15
Timbre	16
Melodía	16
HIPÓTESIS	19
Objetivo general:	19
Objetivos específicos:	19
METODOLOGÍA	20
1. Tipo de diseño	20
2. Selección y clasificación de las piezas musicales	20
3. Preparación de las pistas musicales	22
3.1 Tratamiento de pistas musicales originales	23
3.2 Simulación de las piezas musicales:	23
4. Análisis de coherencia espectral	26
5. Análisis descriptivo y estadístico	27
6. Representación esquemática del abordaje	28
RESULTADOS	30
1. Coherencia espectral de las pistas musicales	31
2. Coherencia espectral de las pistas musicales según su predominio	32
3. Análisis factorial de coherencia espectral de las pistas musicales por predominio rítmico, armónico y melódico.	33
4. Coherencia espectral de las pistas musicales según género musical	34
5. Análisis factorial de coherencia espectral de las pistas musicales por género musical	36
6. Coherencia espectral de las pistas musicales según la presencia y ausencia de voz	38
DISCUSIÓN	39
CONCLUSIÓN	43
BIBLIOGRAFÍA	45

RESUMEN

Introducción: Desafortunadamente, no todas las personas pueden disfrutar de una buena percepción auditiva de la música. Uno de estos colectivos son los usuarios de implante coclear (IC). Se ha visto en muchos estudios que su percepción musical es deficiente, principalmente por dificultades para representar los componentes fundamentales como melodía y armonía. En base a lo anterior, nuestra hipótesis es que música con mayor predominio rítmico que melódico o armónico puede ser mejor percibida por usuarios de IC.

Metodología: Se realizó un análisis de 334 pistas pertenecientes a 11 géneros musicales distintos, estudiando su predominio (ritmo, melodía y armonía), presencia de voz y género musical. Cada pista fue procesada por el simulador de IC "Vocoder". Las pistas procesadas fueron comparadas con las originales mediante la función de coherencia espectral utilizando códigos escritos en el entorno de programación MATLAB 2016a.

Resultados: Se observó que las pistas musicales con predominio rítmico se conservaron de mejor manera tras el procesamiento de IC, en comparación a aquellas con predominancia melódica y armónica. Se observó además que, la voz no fue un componente relevante al realizar el análisis de coherencia espectral entre las pistas musicales originales y simuladas.

Conclusiones: En este trabajo se realizaron dos aportes significativos a la investigación propuesta: 1ro, se escribieron y depuraron códigos de simulación y análisis de pistas, estableciendo además un flujo de trabajo. 2do, se construyó una base de datos que sirve de insumo para la realización futura de experimentos con pacientes implantados. En pasos futuros se pretende comprobar la fiabilidad del método matemático planteado, evidenciando si los valores de coherencia espectral coinciden con la percepción de los usuarios de IC.

Palabras claves: simulador IC, implante coclear, música, percepción, coherencia espectral.

INTRODUCCIÓN

Se estima que la pérdida auditiva afectaba cerca de 1500 millones de personas en 2021, de las cuales más de un 5% de la población mundial tuvo una pérdida auditiva discapacitante que requiere rehabilitación [1,2]. Esta privación auditiva puede afectar la calidad de vida y acceso a la comunicación oral, lo que puede impedir el desarrollo del lenguaje hablado en niños y aumentan el riesgo de demencia en edades más avanzadas, además de propiciar el aislamiento y soledad, así como también, trastornos psiquiátricos como la depresión y ansiedad [3,4].

Es por esto y gracias al avance de la tecnología, que se han creado diferentes alternativas y dispositivos para recuperar, aunque sea parcialmente, la percepción auditiva, entre los cuales destaca el implante coclear (IC). Este dispositivo consta de una parte externa generalmente alojada en el pabellón auditivo, compuesta por un micrófono, que detecta las variaciones de presión asociadas a las ondas sonoras transmitidas por el aire, y un procesador, que está diseñado para convertir las características seleccionadas de las señales acústicas en un patrón de estímulos nerviosos eléctricos que evocarán las sensaciones auditivas adecuadas en el usuario del IC [5,6]. El procesador envía las señales codificadas a un transmisor o bobina, ubicado en la superficie de la piel en la región temporo-parietal que se mantiene por la atracción magnética de dos imanes. El transmisor emite señales por radiofrecuencia modulada, código digital que especifica los parámetros de los estímulos eléctricos, que atraviesan la piel y son recibidas por una antena y un receptor-estimulador, colocados mediante cirugía por debajo de la piel en la región retroauricular [6,7]. La señal transmitida continúa hacia el conjunto de electrodos que se encuentran dentro de la escala timpánica de la cóclea y estimulan eléctricamente las neuronas del ganglio espiral del nervio auditivo a través de una matriz de 12 a 22 electrodos incrustados en la perilinfa conductiva. Esta señal, que especifica la amplitud y los parámetros temporales del estímulo que debe generarse y los

electrodos que deben conducir la corriente, generalmente corresponde a una serie de pulsos simétricos y bifásicos [6,8].

Según la FDA, para diciembre del 2019, aproximadamente 736.900 dispositivos registrados se habían implantado en todo el mundo y, desde el año 2000, está aprobado su uso en niños a partir de los 12 meses de edad, si cumplen con ciertos requisitos [9]. Los usuarios de IC, principalmente personas con hipoacusia sensorineural severa a profunda, utilizan la información contenida en los patrones de estimulación eléctrica para extraer pistas importantes para la percepción del sonido [10,11]. Lo anterior ha permitido que la mayoría de los usuarios de IC pueden lograr excelentes resultados con estímulos de habla simple en entornos silenciosos [10]. Estudios han demostrado que la mayoría de los usuarios de IC son capaces de comprender aproximadamente el 90% del mensaje hablado dentro de un período de entrenamiento y adaptación de dos años, lo que mejora sus capacidades sociales [8,12]. Sin embargo, su desempeño en la entrega de música y habla con ruido de fondo ha sido menos que ideal, lo cual ocurre cómo resultado de limitaciones sensoriales y del dispositivo [11,13]. La capacidad del IC para extraer información sobre los sonidos está limitada en los dominios de frecuencia, tiempo y amplitud [11].

El éxito de los IC probablemente esté relacionado con la facultad de nuestro cerebro para adaptarse a nuevas entradas y su capacidad para entender el habla en situaciones muy distorsionadas [11]. Por estas razones, y muchas más, generalmente los usuarios de IC perciben los elementos fundamentales de la música con menos precisión e informan que disfrutan menos de la música [14].

Música

La música es el arte que involucra combinar sonidos y silencios en una secuencia temporal, acorde a principios fundamentales como el ritmo, la melodía y la armonía. Ésta tiene como

propósito la expresión y comunicación de sentimientos o emociones, generar agrado en el oyente y, además, producir distintas sensaciones, emociones, sentimientos y estados de ánimo en las personas, lo cual depende del contexto y nivel de familiaridad con el género o cultura musical en cuestión [15,16].

La música es un elemento crucial de la vida cotidiana y juega un papel central en todas las culturas humanas. Por ejemplo, el canto está íntimamente ligado a las primeras formas de interacción entre una madre y su hijo [17]. Se ha informado que utilizar la música como una intervención estructurada en el tratamiento de enfermedades mentales como la ansiedad, depresión y esquizofrenia es beneficioso para aliviar los síntomas, mejora el estado de ánimo y las interacciones sociales, entre otros factores biológicos e inmunológicos [17]. La música desempeña un papel importante en la vida de las personas y forma parte de muchos eventos socioculturales y educativos. Está profundamente arraigada en nuestro patrimonio cultural, pero también se considera un lenguaje universal que une a las personas de diferentes culturas [8]. Por lo tanto, la música no es simplemente entretenimiento, la investigación científica ha demostrado que puede influir en los procesos fisiológicos que mejoran el bienestar físico y mental [17].

Entonces, la música es importante para la vida de muchas personas con audición normal o con problemas de audición. Muchos candidatos a IC expresan la esperanza de que disfrutarán de la música después de la implantación [18].

La música a través del IC

Los usuarios de IC a menudo perciben la música como demasiado compleja y no pueden disfrutar de ella [8]. Por lo tanto, los resultados del IC sobre el procesamiento de sonidos complejos como la música, siguen siendo deficientes [10]. Las posibles causas de esta limitación se detallan a continuación:

Distorsiones en el dominio de la frecuencia

Los usuarios de IC perciben el tono a través de mecanismos espaciales y temporales. El mecanismo del tono espacial está relacionado con la ubicación de la estimulación en la cóclea, donde los electrodos ubicados hacia el ápice producen sensaciones de tono más bajo que los electrodos ubicados hacia la base de la cóclea. Es decir, para ser capaces de entregar la información del contenido frecuencial con precisión, los electrodos deben estimular distintas partes del nervio auditivo [8,10]. Este mecanismo puede estar limitado por el bajo número de contactos de electrodos, dispersión de corriente por electrodos individuales y la discrepancia frecuencial entre un electrodo y la frecuencia tonotópica natural de la ubicación particular en la cóclea [8].

La estimulación eléctrica requiere que la corriente fluya de un electrodo a otro, lo que se logra estimulando cada electrodo con referencia a una tierra de campo lejano. Sin embargo, los electrodos quedan distantes del nervio separados por una cámara llena de líquido y una pared de hueso poroso con una resistividad mucho mayor, por lo que la mayor parte de corriente se propaga longitudinalmente a través de los fluidos con menor resistencia, produciendo una excitación más amplia y menos específica [11]. Esto se intenta soslayar con el arreglo perimodiolar, que busca la optimización de la interfaz electrodo-nervio proporcionando una mayor especificidad en la estimulación eléctrica por su ubicación cercana al nervio auditivo y, en teoría, entregando una mejor resolución frecuencial al disminuir la propagación de la estimulación eléctrica debido a que debe atravesar una menor cantidad de estructuras [19,20].

La membrana basilar actúa como un verdadero banco de filtros, es capaz de proporcionar una fuerte selectividad de frecuencias, equivalente a unos 40 filtros que no se superponen [11]. La incapacidad de los IC para seleccionar espacialmente sin que se superpongan las señales va

a provocar una diafonía entre los electrodos y la representación frecuencial distorsionada de los sonidos.

Si bien la cóclea contiene aproximadamente 2.5 vueltas, la gran mayoría de arranques de electrodos solamente se implantan a lo largo de la primera vuelta y media, para evitar daños mecánicos al tejido. En consecuencia, es probable que el IC no estimule las fibras apicales que codifican naturalmente para las bajas frecuencias. Sin embargo, esta información es transmitida por electrodos que estimulan regiones que responderían a frecuencias más altas, provocando un desajuste en la información del sonido [11].

Es por esto que la estimulación eléctrica del IC produce la excitación de regiones más amplias del nervio auditivo, en comparación a la estimulación que ocurre en un oído sano, y muchas veces son incapaces de diferenciar la estimulación eléctrica de electrodos cercanos.

Distorsiones en el dominio temporal

El tono del sonido está determinado por la periodicidad de las modulaciones temporales aplicadas en los pulsos de cada electrodo. En algunas técnicas de procesamiento del sonido, los patrones de estimulación eléctrica tienen modulaciones que se repiten a la frecuencia fundamental del hablante [6,11]. Numerosos investigadores han informado que la variación de la frecuencia de un tren de pulso constante presentado en una región coclear resulta en un cambio del tono percibido [6]. Este mecanismo de tono temporal está relacionado con el patrón temporal de estimulación, se obtiene un tono de mayor frecuencia con una tasa de estimulación aumentada, y la evolución temporal de los pulsos en los diferentes electrodos, es decir, los electrodos no están igualmente activos durante toda la duración del sonido [11].

La señal portadora que viaja a cada electrodo para transmitir la información del sonido suele ser un tren de pulsos bifásicos simétricos, cada pulso tiene una corta duración de 50 μ s, lo que

permite muestrear la información a una velocidad muy alta [11]. Aunque las envolventes extraídas por el procesador contienen modulaciones temporales de hasta 400 Hz o más, la práctica ha mostrado que la mayoría de los usuarios de IC no pueden distinguir señales con diferentes frecuencias de modulación cuando son superiores a 300 Hz y cuando la tasa de pulso de la señal portadora no es lo suficientemente alta (al menos cuatro veces la frecuencia de la señal) [6,11,21].

Sin embargo, todavía no se ha informado de ninguna ventaja consistente en las tareas de percepción del habla para velocidades superiores a aproximadamente 500 pulsos por segundo [11]. Esto es contradictorio, ya que las frecuencias altas deberían muestrear de mejor manera la estructura temporal fina del sonido.

Algunas de las posibles causas, en primer lugar, es que luego de que la fibra nerviosa ha disparado, entra en un estado refractario donde no se producen potenciales de acción, incluso si los pulsos consecutivos se presentan demasiado cerca, es posible que no provoquen descargas por si solos y solo se disparen en el momento del segundo pulso, ambas situaciones no contribuyen a la codificación de sonidos y distorsionan la representación temporal del sonido. Además, los pulsos con tasa fija se utilizan sólo como portadores de información de la envolvente, perdiendo el momento exacto de la generación del potencial de acción, lo que difiere de los patrones de actividad neural temporalmente precisos de la audición normal [11].

Distorsiones en el dominio de la amplitud

Se produce una disminución del rango dinámico en usuarios de IC, lo cual no siempre es un problema [11]. En la audición normal, las fibras del nervio auditivo tienen diferentes propiedades morfológicas y no todas se excitan al mismo nivel de sonido. Las fibras gruesas se reclutan a niveles bajos, mientras que las más delgadas tienen umbrales más altos y responden a niveles más altos [11,22]. Ambas fibras se proyectan a distintas estructuras

centrales, y es posible que la información de sonoridad dependa de cómo se excitan estas estructuras. Sin embargo, la estimulación eléctrica de un IC no controla el reclutamiento de estas fibras nerviosas y es posible que falten las más finas, que son las más frágiles [11].

La intensidad del sonido percibido va a depender de la cantidad de energía eléctrica entregada a los electrodos [11]. Estudios psicofísicos han demostrado una función de crecimiento exponencial de la intensidad debido al aumento de la corriente eléctrica [8]. Sin embargo, en términos de resolución, los oyentes con audición normal pueden discernir un rango dinámico de 120 dB con 60-100 pasos perceptibles, mientras que los usuarios de IC pueden discernir un rango dinámico de 6-30 dB con solo 20 pasos perceptibles [8,23].

De esta manera se aumenta la actividad neuronal al reclutar las mismas fibras de manera más efectiva y al reclutar fibras más distantes. Sin embargo, esto colapsa al activarse varios electrodos simultáneamente, lo cual ocurre en la vida real [11]. En experimentos anteriores con IC se ha evidenciado que la estimulación simultánea de varios electrodos produce una reducción en la selectividad espacial de la excitación neuronal y un control deficiente de la intensidad percibida [6]. Sin embargo, cuando se mide en un solo electrodo, la capacidad de los usuarios de IC para percibir cambios en el nivel de corriente o intensidad, en algunos casos, es bastante cercana a sujetos con audición normal [11].

El deficiente procesamiento de los distintos dominios, principalmente, en sonidos complejos como la música, va a provocar la percepción alterada de sus características fundamentales. La música se compone de varias características fundamentales, que incluyen el tono, timbre, melodía y ritmo, donde cada uno tiene rendimientos distintos en usuarios de IC y provocan la pobre percepción musical [10]. Sin embargo, más allá de estas características objetivas, existen diversos fenómenos que también son importantes en la experiencia de escuchar música. Entre ellos están la calidad subjetiva, el estado de ánimo y el contexto situacional [6].

Características fundamentales de la música e IC

Tono

El tono se refiere al correlato perceptivo de la frecuencia [24]. La percepción del tono hace posible que las personas tengan una buena comprensión de la armonía musical, ayuda a definir el significado de las palabras y ayuda a segregar y dar sentido a las fuentes de sonido de la competencia [25]. El tono se transmite mediante señales espectrales y temporales, aunque la percepción del tono depende más de señales espectrales. Para la mayoría de los usuarios de IC, tanto la información de tono espectral cómo la de tono temporal están interrumpidas [24]. El IC ofrece la oportunidad de manipular de forma independiente los factores temporales y espectrales que contribuyen a la percepción del tono. Sin embargo, en los usuarios de IC, el mecanismo del tono temporal sólo es funcional a tasas bajas de hasta aproximadamente 300 pps. Es decir, el procesamiento del IC no ha podido emplear prácticamente la sensibilidad del tono temporal utilizando frecuencias de pulso [13]. Además, la guía de electrodos no estimula regiones apicales de la cóclea [24]. También, los campos eléctricos relativamente amplios utilizados en la audición mediada por IC carecen de la precisión que poseen las células ciliadas internas para excitar una neurona auditiva [24].

Leal et al. (2003) evaluaron las distintas características fundamentales de la música, entre ellas la discriminación e identificación de tono, para este último se utilizaron 12 pares de piezas musicales que variaban en una octava, dos octavas y algunos semitonos. Como resultado obtuvieron buenos resultados en términos de discriminación de cambios entre 1 o 2 octavas, pero las puntuaciones fueron más bajas con cambios de semitonos [26]. Galvin et al. (2009) obtuvieron resultados similares, los sujetos con audición normal eran capaces de tener un buen desempeño con una entonación de solo 1 semitono, sin embargo, la mayoría de los

usuarios de IC no eran capaces de tener un buen desempeño incluso con una entonación de 5 semitonos [27].

Ritmo

El ritmo se define como patrones temporales auditivos que contienen elementos estructurales de pulso, métrica y tendencias de agrupación en función del tiempo, que van desde los 50 ms hasta los 5 s, lo que corresponde a frecuencias entre 0.2 y 20 Hz [6,8,10]. Se transmite en la mayoría de los estilos musicales mediante variaciones moderadamente rápidas del volumen [6]. En el procesamiento de la señal en el IC se preserva la información de la envolvente temporal, lo que permite una percepción precisa de la envolvente de amplitud y la extracción del ritmo [8,10]. El sistema neural de los usuarios de IC se ajusta más estrechamente a la información de tiempo que un sistema auditivo normal, por lo que las diferencias rítmicas se codifican como intervalos temporales, modulaciones de amplitud o ambas, lo que permite una discriminación de ritmo generalmente buena [13].

Gfeller et al. (1997) utilizó la prueba Adapted Primary Measures of Musical Audiation (PMMA) para evaluar la discriminación de patrones rítmicos y una tarea de 6 pulsos, obteniendo como resultado un desempeño similar en sujetos con audición normal y usuarios de IC en la prueba de patrón rítmico de PMMA [28]. Lassaletta et al. (2008) también utilizó la PMMA, obteniendo como resultado una mejor puntuación en las pruebas de discriminación de ritmo en comparación a las pruebas de discriminación de tono y timbre [29]. Sin embargo, se han encontrado hallazgos que resaltan los déficits de procesamiento en señales rítmicas más complejas y realistas [10].

Timbre

El timbre es el conjunto de cualidades auditivas que distingue a dos instrumentos diferentes que tocan la misma nota o tono, por lo que es un elemento esencial del valor estético de la música [6]. Por lo tanto, una tarea en la que los sujetos tienen que identificar instrumentos musicales corresponde a una tarea que se basa en la percepción del timbre, pero debido a su pobre percepción del timbre tienen dificultades para identificar instrumentos [25]. La percepción del timbre se basa en las características de banda ancha del sonido, incluyendo la distribución de energía espectral, la envolvente temporal y sus características de ataque y decaimiento [8]. Si bien la envolvente se conserva en el procesamiento de IC, la información espectral se reduce en comparación con normoyentes, por lo que el reconocimiento del timbre en usuarios de IC es mejor que el azar, pero no tan bueno como normoyentes [13].

Schulz y Kerber (1994) pidieron a 8 usuarios de IC que identificaran el instrumento que tocaba una melodía, obteniendo como resultado una puntuación de 35% en los usuarios de IC y un 90% en personas con audición normal [6]. Gfeller et al. (2002) realizaron un estudio similar en 51 usuarios de IC y le pidieron que identificaran grabaciones de ocho instrumentos diferentes tocando la misma secuencia breve de notas, obteniendo como resultado una puntuación de un 46.6% en usuarios de IC y un 90.9% en sujetos con audición normal [30].

Melodía

La melodía corresponde a una secuencia lineal de notas, que consiste en una combinación de tonos y ritmos. La percepción de la melodía depende principalmente de cómo sea la composición de ésta, lo cual varía según el género musical, pero es esta secuencia de tonorritmo, que hace que sea percibido de manera deficiente en los usuarios de IC, aunque las señales verbales y el ritmo pueden ayudar a la identificación de melodías [13]. Su estudio consiste en la identificación de melodías familiares que un oyente pueda nombrarlas al

escucharlas, aunque esta familiaridad puede estar influenciada por factores muy variables, tales como la formación musical, experiencia auditiva, cultura social y memoria de los títulos y melodías [6]. La percepción de la melodía se encuentra afectada por las limitaciones en la transmisión de la frecuencia fundamental y la distorsión en la transmisión de la armonía generadas por el IC. Por otro lado, el posible desajuste entre las frecuencias asociadas a los electrodos y las frecuencias tonotópicas pueden causar una transmisión inconsistente entre el tono espacial y el tono temporal, especialmente en frecuencias bajas [8].

Gfeller et al. (2002) presentaron 12 melodías conocidas a 49 usuarios de IC y 18 sujetos con audición normal, obteniendo como resultado una puntuación en el reconocimiento de la melodía de un 19% en usuarios de IC y un 83% en sujetos con audición normal [31].

Se entiende como armonía a la sucesión de acordes o al sonido simultáneo de dos o más notas, compuestas por tonos y timbres de distintos instrumentos. Este concepto fue estudiado junto a otras características musicales, a diferencia de los estudios descritos anteriormente, que se centran en el estudio individual de cada característica, la armonía no fue estudiada.

Es por estas dificultades que presenta el procesamiento de la señal en el IC para transmitir y representar fielmente las variaciones de tono, timbre, melodía y ritmo que los usuarios de IC presentan una percepción distorsionada de la música, lo que provoca una disminución en su calidad de vida, estado de ánimo e interacción social.

Por esta razón, con el fin de evaluar las características fundamentales de la música es el por qué se han creado diferentes pruebas para la evaluación de la percepción y apreciación musical. Looi y She (2010) desarrollaron el *University of Canterbury Music Listening Questionnaire (UCMLQ)* [32], Nimmons et al. (2008) desarrollaron la *University of Washington Clinical Assessment of Music Perception (CAMP)* [33], Peretz et al. (2003) desarrollaron y

perfeccionaron la *Montreal Battery of Evaluation of Amusia (BMEA)* [34], S. J. Brockmeier et al. (2002) desarrollaron el *Munich Music Questionnaire (MuMu)* [37].

En síntesis, los estudios enfocados en la percepción musical en usuarios de IC son consistentes en señalar que la calidad y apreciación de la música está deteriorada cuando se les compara a sujetos normoyentes. Adicionalmente, la calidad de la percepción musical está asociada a la calidad de vida, y a aspectos psicológicos de los usuarios de IC [38]. Cuando se analizan los aspectos musicales, que en mayor o menor medida se encuentran deteriorados por el procesamiento de la señal de audio realizada por el IC, se observa que el ritmo parece ser el aspecto mejor percibido, en comparación a otras características como el tono, armonía, timbre o melodía. Una pieza musical suele tener todos los componentes anteriormente mencionados entremezclados complejamente. Por lo tanto, de acuerdo con lo reportado por usuarios de IC, si dispusiéramos de piezas musicales con alto contenido proporcional en aspectos temporales como ritmo y tempo y/o voz, en detrimento de otros aspectos con alto contenido frecuencial como tono, armonía, timbre o melodía, se podría predecir, teóricamente, que estas pistas sean mejor percibidas por los usuarios de IC. En este trabajo se plantea testear esta idea. Recientemente, Ibáñez et al. (2022), evaluaron el nivel de deterioro de una pista musical, procesada por un IC, para distintos estilos musicales. Los resultados sugieren que no todas las pistas musicales sufren el mismo nivel de deterioro (expresado como una disminución en la coherencia entre la señal original y la procesada) y que la dirección y magnitud del deterioro, son difíciles de predecir. Por lo tanto, este trabajo se plantea como una continuación del anterior y, utilizando el mismo abordaje, se seleccionan diversas pistas musicales que contengan mayoritariamente un componente musical entre: armonía, melodía y ritmo, aumentando la cantidad de pistas estudiadas por género musical, utilizando un simulador de IC actualizado e innovando en las técnicas de análisis matemático utilizadas en el análisis de coherencia espectral [39].

HIPÓTESIS

El nivel de deterioro de una pista musical procesada por el implante coclear varía en función de las características musicales que contiene este, siendo el ritmo el componente que presenta un menor deterioro.

Objetivo general:

Analizar el grado de deterioro de la señal de audio musical, en función de su proporción de características musicales como tono, timbre, melodía y ritmo.

Objetivos específicos:

- Seleccionar pistas musicales con distintas proporciones de tono, timbre, melodía y ritmo.
- Procesar las señales de audio seleccionadas, correspondientes a canciones de distintos estilos musicales, utilizando un simulador de implante coclear.
- Comparar mediante un análisis de coherencia espectral, las pistas musicales procesadas y no procesadas por el simulador de implante coclear.
- Analizar las diferencias observadas en el nivel de deterioro de cada característica musical estudiada.

METODOLOGÍA

1. Tipo de diseño

Estudio de tipo descriptivo, observacional.

2. Selección y clasificación de las piezas musicales

Se seleccionaron canciones libres de derechos de autor obtenidas desde la página de la compañía Epidemic Sound [40]. Se seleccionaron los estilos musicales destacados, y se obtuvieron las piezas musicales (canciones completas) de la sección “populares” de la página Epidemic Sound, con el objetivo de seleccionar aquellos estilos y pistas musicales que formen parte del contexto musical de las personas durante su diario vivir.

De esta forma se obtuvo un total de 105 piezas musicales:

- Según el género musical se obtuvieron 7 piezas para Beats, 12 piezas para Classic, 9 piezas para Country, 7 piezas para Electronic, 10 piezas para Funk, 10 piezas para HipHop, 9 piezas para Jazz, 8 piezas para Pop, 6 piezas para RnB, 6 piezas para Rock y 11 piezas para Small Emotions.

De cada pieza seleccionada se recortaron y acotaron fragmentos con una duración de 15 segundos, generando distintas pistas musicales de 15 segundos, de las piezas ya seleccionadas, proceso realizado a través del software “Audacity ® Versión 3.3.1” debido a que permitió realizar dicho proceso de manera rápida y precisa para su posterior análisis, el cual será detallado más adelante [41].

Cada pista musical fue seleccionada al azar y clasificada según el componente musical predominante, entre: melodía, armonía y ritmo. La clasificación del componente musical fue realizada por los mismos integrantes de este trabajo, acompañada y contrastada con la

información que entrega la plataforma de Epidemic sound, permitiendo escuchar por separado los componentes melódicos, armónicos y rítmicos de las pistas en estudio, entendiéndose como:

- La **melodía** hace referencia a la sucesión de sonidos con diferentes frecuencias (tonos) y amplitudes (espectrograma en Anexo 1).
- El **ritmo** corresponde a un factor organizativo, que consta de acentos fuertes y débiles que estructuran la música (espectrograma en Anexo 2).
- La **armonía** corresponde a una sucesión de acordes o al sonido simultáneo de dos o más notas (tonos y timbres), que acompañan la melodía (espectrograma en Anexo 3).

Además, cada pista fue clasificada según la presencia o ausencia de señales verbales. Para las pistas musicales con presencia de voz son seleccionadas aquellas pistas que contienen un extracto mayor o igual a 5 segundos con presencia de oraciones verbales entendibles y/o con un significado, descartando todas aquellas pistas que utilizan la voz como un componente armónico o rítmico. Las pistas musicales con ausencia de voz son aquellas que no cumplen los requisitos antes mencionados.

Mediante los pasos señalados, se obtuvieron un total de 334 pistas musicales.

- Según el género musical se obtuvieron 30 pistas para Beats, 30 pistas para Classic, 30 pistas para Country, 30 pistas para Electronic, 30 pistas para Funk, 30 pistas para HipHop, 30 pistas para Jazz, 30 pistas para Pop, 31 pistas para RnB, 33 pistas para Rock y 30 pistas para Small Emotions.
- Según el componente musical predominante se obtuvieron 128 pistas en las que predomina la melodía, 101 pistas en las que predomina el ritmo y 105 pistas en las que predomina la armonía.

- Según la presencia o ausencia de señales verbales se obtuvieron 101 pistas con presencia de voz y 203 pistas con ausencia de voz.

Cada pista musical se trabajó utilizando el software “Audacity” y tanto el proceso previo como el análisis se realizó mediante el entorno de “MATLAB versión R2016a”.

3. Preparación de las pistas musicales

Todas las piezas obtenidas desde la plataforma Epidemic Sound contienen las siguientes características:

- Formato estéreo.
- Frecuencia de muestreo de 48.000 Hz.
- 32 bits flotantes.

Es de importancia conocer dichas características ya que esta información es utilizada en procesos posteriores.

Debido a la cuantiosa cantidad de pistas musicales trabajadas, se utilizaron distintas funciones de MATLAB como:

- “dir”: que enumera el contenido de las carpetas [42],
- “fullfile”: que crea rutas portátiles y seguras compatibles con otros sistemas operativos, permitiendo acceder a las carpetas de interés [43],
- “saveas” y “audiowrite”: guardan los distintos gráficos y archivos de audios utilizados [44,45], y
- “for”: repite procesos las veces que se estime necesario [46].

Con estas funciones y la creación de bucles en MATLAB, se creó un conjunto de códigos que permitieron realizar los distintos procedimientos en todas las pistas de forma simultánea, permitiendo aumentar de forma significativa la cantidad de pistas trabajadas.

3.1 Tratamiento de pistas musicales originales

Las pistas originales seleccionadas se trabajaron de forma separada en carpetas según su género musical.

Mediante MATLAB, estas pistas fueron convertidas desde su formato estéreo (original) a un formato mono para su posterior estudio y comparación, debido a que el código de análisis realizaba la comparación de un solo canal. Posteriormente, cada pista fue normalizada, para que de esta forma tanto la pista original como la pista simulada sean lo más similares posibles en cuanto a procesos realizados, ya que, en el guardado del archivo de las pistas musicales simuladas, era necesario dicho paso para el simulador de IC.

3.2 Simulación de las piezas musicales:

Las pistas originales seleccionadas fueron procesadas mediante un código de programación realizado en el entorno de MATLAB, al cual llamaremos "VoCoder". Este código fue construido y facilitado por el profesor Waldo Nogueira, director del "The Auditory Prosthetic Group of the German Hearing Center", parte de la Medical University Hannover. Este código se encarga de realizar el procesamiento de archivos de audio, funcionando como un simulador de implante coclear, de esta forma se obtienen las pistas musicales procesadas utilizadas para el análisis de resultados.

Como se mencionó en el punto 3.1, previo al guardado de las pistas musicales simuladas en su lugar correspondiente, las pistas musicales fueron convertidas a formato mono y normalizadas debido a que, de lo contrario, las pistas musicales simuladas sufrían clipping, es decir, recortes de extractos que excedían los límites de amplitud predefinidos y eran eliminados. Por lo tanto, la normalización previene la eliminación de información y permite ajustar la amplitud, manteniendo los elementos de la mezcla y evitando la distorsión generada

por el clipping [47]. La normalización realizada en las pistas originales fue la misma que para las pistas simuladas, para que de esta forma tanto las pistas originales como las pistas procesadas posean el mismo ajuste de amplitud, evitando así diferencias causadas por el tratamiento particular de las pistas.

Las pistas procesadas representan la percepción de la pista original por un usuario de IC, la cual corresponde a la señal que se analizará posteriormente.

Los parámetros que se utilizarán para realizar la simulación serán los siguiente:

- Límite inferior y superior del rango espectral procesado por el IC: 350-5500 Hz.
- Número de canales: 16 canales.

Los parámetros corresponden a los entregados por defecto en el código, ya que son aquellos que representan de manera óptima la escucha por parte de un usuario de IC.

Debido a las características y estrategias de procesamiento de señales de los IC, las pistas procesadas difieren en duración y en patrones temporales con su versión original no procesada. En este estudio, aquellas pistas musicales procesadas, además del procesamiento frecuencial, sufren una compresión temporal. Las pistas musicales originales poseen una duración de 15 segundos con 720.000 datos, pero su procesamiento entregaba como resultado pistas musicales con una duración de 14,5067 segundos y 696.320 datos. Si bien esta diferencia de datos no provoca una pérdida de información, representa una compresión de la pista original, lo que dificulta e impide realizar un análisis certero y confiable de la coherencia espectral dada esta diferencia temporal, en lugar de los cambios frecuenciales en las señales de audio procesadas por un IC.

Esta diferencia de tiempo y patrón temporal se revirtió mediante técnicas de procesamiento de señales. Según la función “factor de estiramiento” se obtuvo la diferencia entre la pista original

y la pista procesada, es decir, cuanto debe descomprimirse la pista procesada en función de su pista original para recuperar la duración y, de esa forma, el patrón temporal original. Debido a que todas las pistas poseen una duración de 15 segundos, la simulación de las pistas provoca en todas, la misma compresión temporal, por lo que se calculó el factor de estiramiento para todas las pistas musicales mediante la siguiente fórmula:

Factor de estiramiento

$$= \frac{\text{Duración total de la señal original} / \text{N}^\circ \text{ de datos de la señal original}}{\text{Duración total de la señal procesada} / \text{N}^\circ \text{ de datos de la señal procesada}}$$

Utilizando cualquiera de los datos señalados, tanto la duración como la cantidad de datos, se obtiene el mismo valor, que corresponde a $\approx 1,0340073$.

Una vez obtenido el valor del factor de estiramiento, en la misma línea se utilizaron distintas técnicas para revertir la compresión temporal, entre las cuales se encuentran: la interpolación, el resamplio, el cálculo de la distancia euclidiana, el dynamic time warping (DTW) o el cálculo de la correlación cruzada. De todas las técnicas utilizadas, la seleccionada fue la interpolación, debido a que los resultados obtenidos mediante este procedimiento no contenían ruido ni segmentos rellenados con silencio como ocurría utilizando alguna de las otras técnicas mencionadas.

La interpolación fue realizada mediante “MATLAB” utilizando la función “interp1” o interpolación de datos unidimensional (1D) [48], de esta manera se logró igualar la duración y la cantidad de datos entre las pistas originales y las pistas procesadas, logrando así superar la limitación que suponía la compresión temporal por el simulador del IC, de esta forma se puede realizar el análisis de coherencia espectral.

4. Análisis de coherencia espectral

Se analiza la coherencia espectral entre la pista original y la pista procesada por el simulador de IC. Para ello, se utilizará un script en el entorno MATLAB que permitirá calcular la magnitud cuadrática de tal coherencia utilizando la función “mscohere” (Magnitud Square Coherence). Esta función entrega un valor máximo de 1 cuando la correlación es perfecta y 0 cuando las señales son totalmente diferentes (ver Anexos 4 y 5).

Esta magnitud (C_{xy}) se encuentra dada por:

$$C_{xy} = (abs(P_{xy}).^2)./(P_{xx}.* P_{yy})$$

donde P_{xx} y P_{yy} corresponden a la Densidad de la Potencia Espectral (Power Spectral Density, PSD) estimada de la señal original y de la señal simulada, respectivamente, y P_{xy} es la Densidad de la Potencia Espectral Cruzada (Cross-PSD, CPSD) estimada de la señal original y de la señal simulada [39,49].

En nuestro análisis, las señales sonoras serán divididas en 72 fragmentos de 0.2083 segundos, que serán estudiados en el rango frecuencial entre 350 Hz a 5500 Hz, correspondiente a lo trabajado mediante el simulador de IC. Cada uno de los 72 fragmentos se analiza frecuencialmente con el algoritmo de Coherencia Espectral basado en la Transformada de Fourier Rápida FFT. La FFT requiere de la discretización de cada fragmento a partir del uso de “ventanas” de diferente tamaño y forma, las que pueden tener ciertos efectos deseables e indeseables al momento de analizar señales. Lo anterior hace necesario testearlas antes de tomar una decisión sobre el algoritmo a aplicar. Se realizaron distintas pruebas utilizando diferentes tipos de ventanas, tales como: Hamming, Hanning y Kaiser, modificando el tamaño de ventana entre 256, 512 y 1024 obteniéndose los resultados más concordantes entre valor de MSC y la graficación espectral entre la pista original y simulada

con ventanas Kaiser con una resolución o tamaño de 1024 (ver Anexos 4 y 5) [50]. Cada valor de coherencia espectral obtenido para un tiempo y frecuencia determinado será promediado en el tiempo total de la pista estudiada, obteniendo valores entre 0 y 1.

Este análisis permitirá conocer cuánto se corresponden a nivel espectral, la señal original y la procesada en un mismo punto en el tiempo. Bajo el supuesto de que la señal original representa el nivel óptimo de calidad de percepción para esa pista, cualquier pérdida de coherencia supondrá un deterioro en la calidad de la percepción. Por lo tanto, un mayor valor de coherencia sugeriría una mayor posibilidad de ser percibida adecuadamente por usuarios de IC.

Finalmente, a través del código obtenido, se buscarán correlaciones entre los niveles de coherencia obtenidos para las distintas pistas musicales, en función de la preponderancia de los parámetros de interés (melodía, armonía y ritmo). Lo anterior permitirá probar o refutar la hipótesis de que las canciones ricas en ritmo poseen un menor nivel de deterioro una vez procesadas por el simulador de IC, por lo tanto, un mayor índice de coherencia. El análisis será realizado estadísticamente mediante técnicas de comparación de grupos, paramétricas y no paramétricas según corresponda. Esto permitirá finalmente organizar una base de datos, con altos y bajos niveles de coherencia, que en un futuro podrían ser utilizados en investigaciones con usuarios de IC para evaluar el grado de predictibilidad del análisis de coherencia espectral del nivel de audibilidad de la pista musical en usuarios de IC.

5. Análisis descriptivo y estadístico

Se describirán de forma cualitativa y cuantitativa a las pistas que formaron parte del estudio mediante tablas de frecuencia de aparición, contemplando la cantidad por característica estudiada mediante su frecuencia absoluta y acumulada, y el porcentaje por característica estudiada mediante su frecuencia relativa y frecuencia relativa acumulada en porcentaje.

Para el análisis estadístico, los datos fueron ingresados al programa del "IBM SPSS STATISTICS versión 29.0.1.0". Los datos ingresados corresponden al predominio de la pista musical, género de la pista musical, presencia o ausencia de voz y valor de coherencia espectral entre la pista original y la pista procesada.

De esta forma, se obtuvieron los valores de media aritmética y de desviación estándar. Los datos fueron analizados con: la significancia del estadístico de Levene en la prueba de homogeneidad de varianzas, que permite determinar la igualdad de varianzas; la significancia del ANOVA, que permite determinar la igualdad de medias aritméticas y si esta es significativa; y la significancia de las pruebas HSD Tukey y Bonferroni en las comparaciones múltiples, que permite determinar la existencia de diferencias significativas entre los grupos [51].

6. Representación esquemática del abordaje

La Figura 1 corresponde a una representación esquemática del abordaje planteado en este estudio. En donde se seleccionaron diversas pistas musicales (o fragmentos de canciones). Estas pistas serán procesadas por un simulador de implante coclear ("VoCoder"), obteniendo pistas musicales simuladas, luego tanto la pista original como la pista simulada son trabajadas mediante MATLAB para convertirlas y trabajarlas en formato mono y finalmente se realiza una interpolación a las pistas simuladas para igualar la duración y patrón temporal entre las pistas simuladas y originales, para realizar el análisis que se detalla más adelante. Posteriormente mediante un análisis de coherencia se evaluará el nivel de deterioro que presenta cada pista musical simulada. Este indicador presumiblemente nos dará el grado de audibilidad de la pista musical. Este abordaje tiene dos razonables supuestos: i) la comparación entre las señales nos otorgará en grado de deterioro de la señal. A mayor similitud, menor deterioro y ii) un mayor deterioro conlleva a una peor calidad en la calidad de la pieza musical.

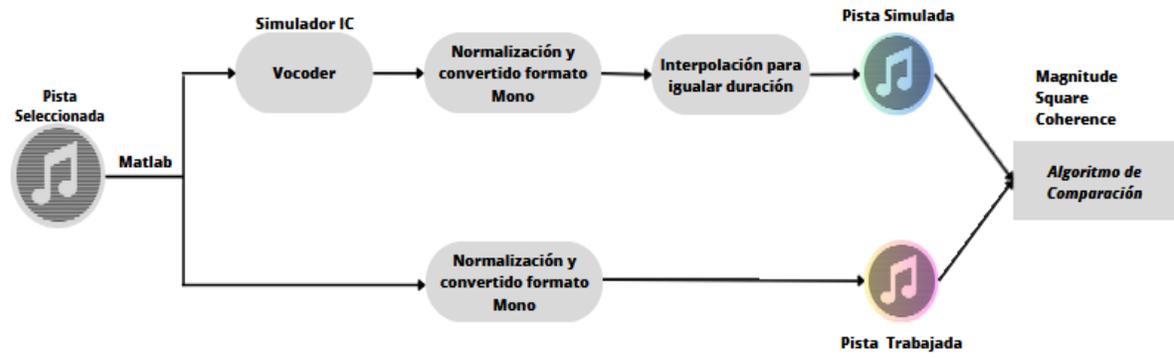


Figura 1. Esquema de abordaje general del estudio

RESULTADOS

A continuación, se presentan los resultados del análisis de las pistas musicales de 15 segundos según su predominio, su género musical y su presencia de voz.

Las 334 pistas seleccionadas, en su totalidad fueron procesadas y trabajadas, mostrando su detalle según su predominio en la Tabla 1, según su género musical en la Tabla 2 y según la presencia de voz en la Tabla 3.

	Frecuencia absoluta	Frecuencia absoluta acumulada	Frecuencia relativa en %	Frecuencia relativa acumulada en %
Ritmo	101	101	30,2	30,2
Armonía	105	206	31,4	61,7
Melodía	128	334	38,3	100,0
Total	334		100,0	

Tabla 1. Tabla de frecuencias de las pistas musicales según su predominio rítmico, armónico y melódico.

	Frecuencia absoluta	Frecuencia absoluta acumulada	Frecuencia relativa en %	Frecuencia relativa acumulada en %
Beat	30	30	9,0	9,0
Classic	30	60	9,0	18,0
Country	30	90	9,0	26,9
Electro	30	120	9,0	35,9
Funk	30	150	9,0	44,9
Hip Hop	30	180	9,0	53,9
Jazz	30	210	9,0	62,9
Pop	30	240	9,0	71,9
RnB	31	271	9,3	81,1
Rock	33	304	9,9	91,1
SmallEmotions	30	334	9,0	100,0
Total	334		100,0	

Tabla 2. Tabla de frecuencias de las pistas musicales según su género musical.

	Frecuencia absoluta	Frecuencia absoluta acumulada	Frecuencia relativa en %	Frecuencia relativa acumulada en %
Sin voz	233	233	69,8	69,8
Con voz	101	101	30,2	100,0
Total	334		100,0	

Tabla 3. Tabla de frecuencias de las pistas musicales según su presencia de voz.

1. Coherencia espectral de las pistas musicales

En la Figura 2 se observa la representación del nivel de coherencia espectral para las 334 pistas musicales, con valores de media y desviación estándar de 0.1574970 y 0.10561892, respectivamente. Se logra evidenciar una gran concentración de valores bajos y muy pocos casos con valores cercanos a 1.

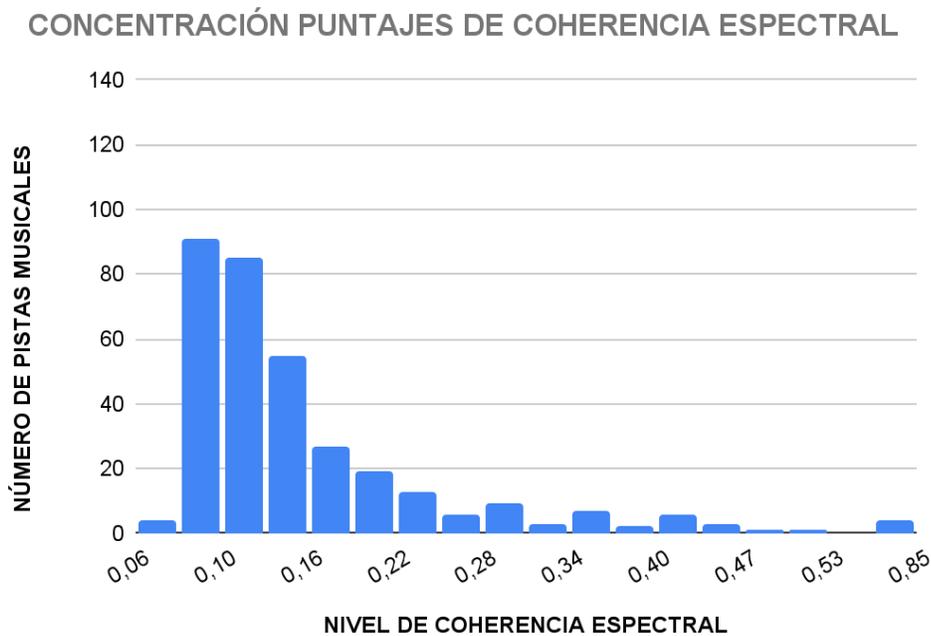


Figura 2. Concentración de los puntajes de coherencia espectral.

2. Coherencia espectral de las pistas musicales según su predominio

En la Figura 3 se observa la representación del nivel de coherencia espectral para las pistas musicales que poseen como predominio principal el ritmo, armonía y melodía, y en la Tabla 4 se observan los valores de media y desviación estándar de los valores de coherencia espectral para ritmo, armonía y melodía. Se logra evidenciar que las pistas musicales con predominio rítmico poseen un mayor valor promedio de coherencia espectral, a diferencia de las pistas musicales con predominio armónico y melódico.

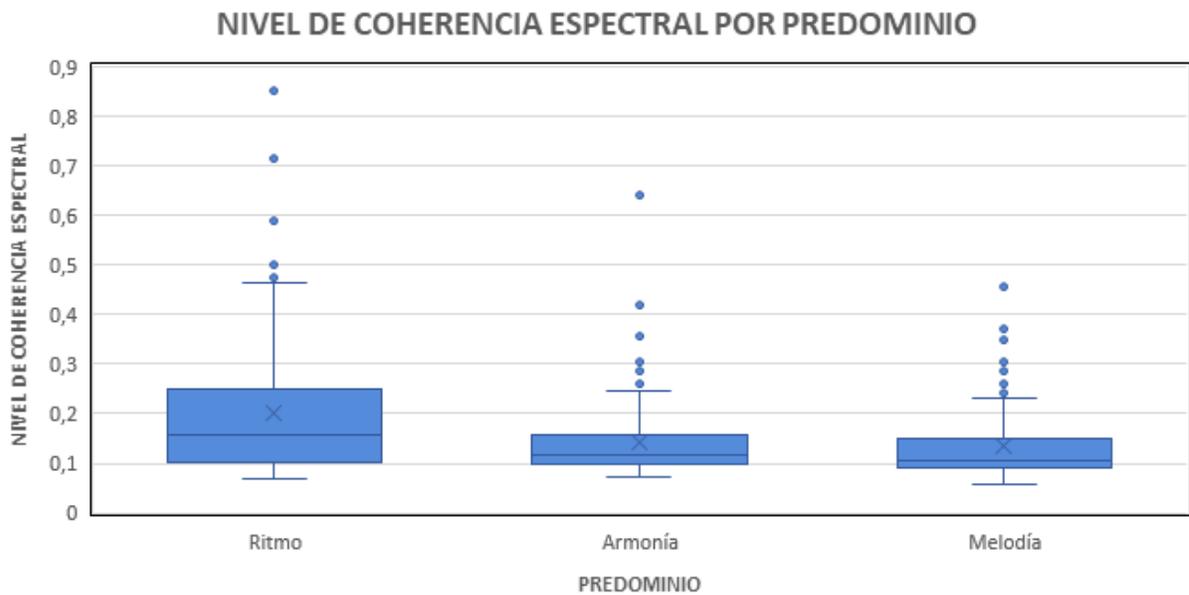


Figura 3. Nivel de coherencia espectral según el predominio de las pistas musicales (ritmo, armonía, melodía).

	Media	Desviación estándar
Ritmo	0,2019447	0,14078874
Armonía	0,1338590	0,07344360
Melodía	0,1435584	0,08526527

Tabla 4. Valores de media y desviación estándar para los valores de coherencia espectral de las pistas musicales según su predominio rítmico, armónico y melódico.

3. Análisis factorial de coherencia espectral de las pistas musicales por predominio rítmico, armónico y melódico.

Para conocer si el tipo de predominancia de la característica musical es relevante sobre el grado de la señal de audio procesada, se realizó un análisis factorial mediante un ANOVA para un factor en el programa "SPSS STATISTICS". Siendo el factor, la predominancia y la variable dependiente, el grado de coherencia espectral.

Como se puede observar en la Tabla 5 y en la Tabla 6, existe un efecto significativo en la predominancia de la pista musical y la magnitud del deterioro de la pista de audio procesada por el simulador de IC ($F(2,331) = 14,09$, $p < 0,05$, $n_2 = 0,078$).

Como se puede observar en la Tabla 7, las comparaciones múltiples evidencian diferencias significativas entre las pistas con predominio rítmico y predominio armónico ($t(331) = 4,14$, $p < 0,05$) y entre las pistas con predominio rítmico y predominio melódico ($t(331) = 5,23$, $p < 0,05$).

	Estadístico de Levene	gl1	gl2	Significancia
Se basa en la media	16,739	2	331	< 0.001
Se basa en la mediana	10,113	2	331	< 0.001
Se basa en la mediana y con gl ajustado	10,113	2	256,305	< 0.001
Se basa en la media recortada	13,619	2	331	< 0.001

Tabla 5. Prueba de homogeneidad de varianzas.

	Suma de cuadrados	gl	Media cuadrática	F	Significancia
Entre grupos	0,291	2	0.146	14,091	< 0.001
Dentro de grupos	3,423	331	0.010		
Total	3,715	333			

Tabla 6. ANOVA.

	Predominio	Predominio	Diferencia de medias	Error estándar	Significancia
HSD Tukey	Ritmo	Armonía	0,05838634	0,01417377	< 0.001
		Melodía	0,06808570	0,01353503	< 0.001
	Armonía	Ritmo	-0,05838634	0,01417377	< 0.001
		Melodía	0,00969935	0,01339015	0,749
	Melodía	Ritmo	-0,06808570	0,01353503	< 0.001
		Armonía	-0,00969935	0,01339015	0,749
Bonferroni	Ritmo	Armonía	0,05838634	0,01417377	< 0.001
		Melodía	0,06808570	0,01353503	< 0.001
	Armonía	Ritmo	-0,05838634	0,01417377	< 0.001
		Melodía	0,00969935	0,01339015	1,000
	Melodía	Ritmo	-0,06808570	0,01353503	< 0.001
		Armonía	-0,00969935	0,01339015	1,000

Tabla 7. Prueba de comparaciones múltiples. La diferencia de media es significativa con un valor de significancia igual o inferior a 0,05.

Entonces, la variable que menos se deteriora con diferencia es el ritmo, el cual es significativamente mejor que la armonía y melodía. Mientras que entre melodía y armonía no existen diferencias significativas en el deterioro.

4. Coherencia espectral de las pistas musicales según género musical

En la Figura 4 se observa la representación del nivel de coherencia espectral para las pistas musicales de los distintos géneros musicales estudiados, y en la Tabla 8 se observan los valores de media y desviación estándar de los valores de coherencia espectral para cada género musical. Se logra evidenciar que las pistas musicales pertenecientes al género “Beats” poseen un mayor valor promedio de coherencia espectral y las pistas musicales pertenecientes al género “Classic” poseen un menor valor promedio de coherencia espectral. El resto de los géneros musicales poseen valores intermedios.

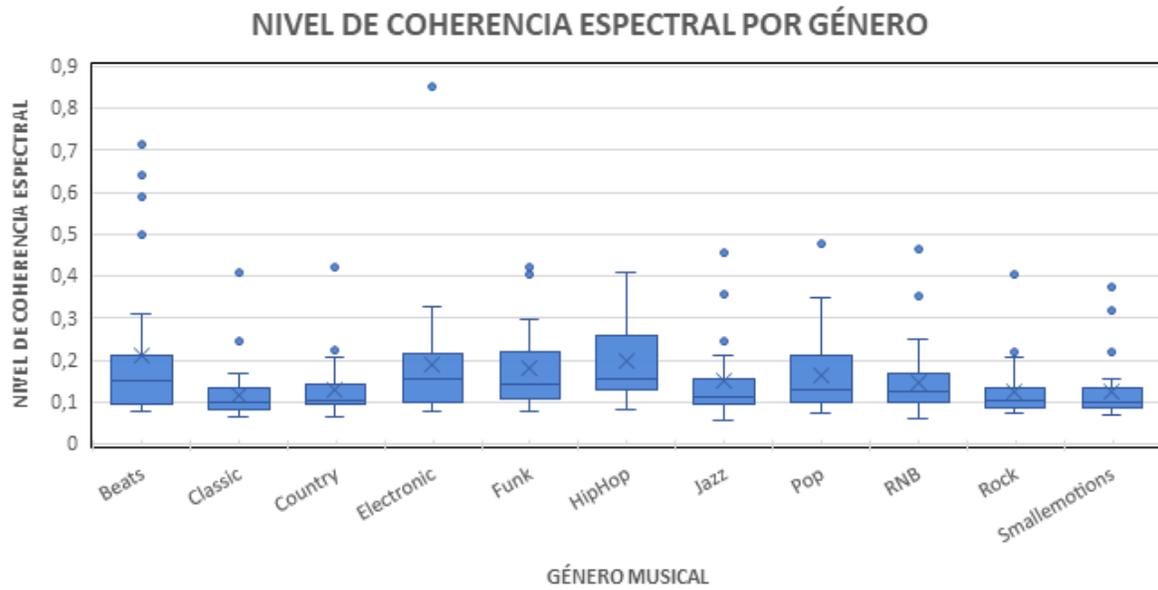


Figura 4. Nivel de coherencia espectral ordenados según el género de la pista musical.

	Media	Desviación estándar
Beats	0,2092568	0,17340590
Classic	0,1180553	0,06685461
Country	0,1296903	0,06998034
Electronic	0,1877334	0,15078359
Funk	0,1825163	0,10175247
HipHop	0,1962258	0,09210174
Jazz	0,1499609	0,10305822
Pop	0,1655823	0,09520369
RnB	0,1464661	0,08346685
Rock	0,1240901	0,06341289
SmallEmotions	0,1265983	0,07012397

Tabla 8. Valores de media y desviación estándar para los valores de coherencia espectral de las pistas musicales según su género musical.

5. Análisis factorial de coherencia espectral de las pistas musicales por género musical

Para conocer si el tipo de género musical de la pista musical es relevante sobre el grado de la señal de audio procesada, se realizó un análisis factorial mediante un ANOVA para un factor en el programa "SPSS STATISTICS". Siendo el factor, el género musical y la variable dependiente, el grado de coherencia espectral.

Como se puede observar en la Tabla 9 y en la Tabla 10, existe un efecto significativo en el género musical y la magnitud del deterioro de la pista de audio procesada por el simulador de IC ($F(10,323) = 2,98, p < 0,05, n^2 = 0,31$).

Como se puede observar en la Tabla 11, las comparaciones múltiples evidencian diferencias significativas entre las pistas del género beats y el género classic ($t(323) = 3,5, p < 0,05$) y entre las pistas del género beats y el género rock ($t(323) = 3,4, p < 0,05$). El resto de las combinaciones entre géneros no muestran diferencias significativas.

	Estadístico de Levene	gl1	gl2	Significancia
Se basa en la media	3,653	10	323	< 0.001
Se basa en la mediana	2,048	10	323	0.028
Se basa en la mediana y con gl ajustado	2,048	10	202,782	0.030
Se basa en la media recortada	2,975	10	323	0.001

Tabla 9. Prueba de homogeneidad de varianzas.

	Suma de cuadrados	gl	Media cuadrática	F	Significancia
Entre grupos	0,314	10	0,031	2,986	0.001
Dentro de grupos	3,423	323	0.011		
Total	3,715	333			

Tabla 10. ANOVA

	Predominio	Predominio	Diferencia de medias	Error estándar	Significancia
HSD Tukey	Beats	Classic	0,09120150	0,02849214	0,027
		Country	0,07956650	0,02849214	0,098
		Electro	0,02152340	0,02849214	0,999
		Funk	0,02674050	0,02849214	0,995
		HipHop	0,01303103	0,02849214	1,000
		Jazz	0,05929590	0,02849214	0,479
		Pop	0,04367453	0,02849214	0,860
		RnB	0,06279074	0,02827763	0,377
		Rock	0,08516674	0,02588305	0,043
		Small Emotions	0,08265857	0,02849214	0,071
Bonferroni	Beats	Classic	0,09120150	0,02849214	0,036
		Country	0,07956650	0,02849214	0,158
		Electro	0,02152340	0,02849214	1,000
		Funk	0,02674050	0,02849214	1,000
		HipHop	0,01303103	0,02849214	1,000
		Jazz	0,05929590	0,02849214	1,000
		Pop	0,04367453	0,02849214	1,000
		RnB	0,06279074	0,02827763	0,959
		Rock	0,08516674	0,02588305	0,061
		Small Emotions	0,08265857	0,02849214	0,108

Tabla 11. Prueba de comparaciones múltiples resumida. La diferencia de media es significativa con un valor de significancia igual o inferior a 0,05.

Entonces, la variable que menos se deteriora con diferencia es el género musical "Beats", el cual es significativamente mejor que el resto de los géneros musicales, y que la variable que más se deteriora con diferencia es el género musical "Classic", el cual es significativamente

peor que el resto de los géneros musicales. Mientras que, entre Country, Electronic, Funk, Hip Hop, Pop, RnB y Small Emotions, no existen diferencias significativas en el deterioro.

6. Coherencia espectral de las pistas musicales según la presencia y ausencia de voz

En la Figura 5 se observa la representación del nivel de coherencia espectral para las pistas musicales con presencia y ausencia de voz, y en la Tabla 12 se observan sus valores de media y desviación estándar de los valores de coherencia espectral. Se logra evidenciar que las pistas musicales con ausencia de voz poseen un mayor valor promedio de coherencia espectral, a diferencia de las pistas musicales con presencia de voz.

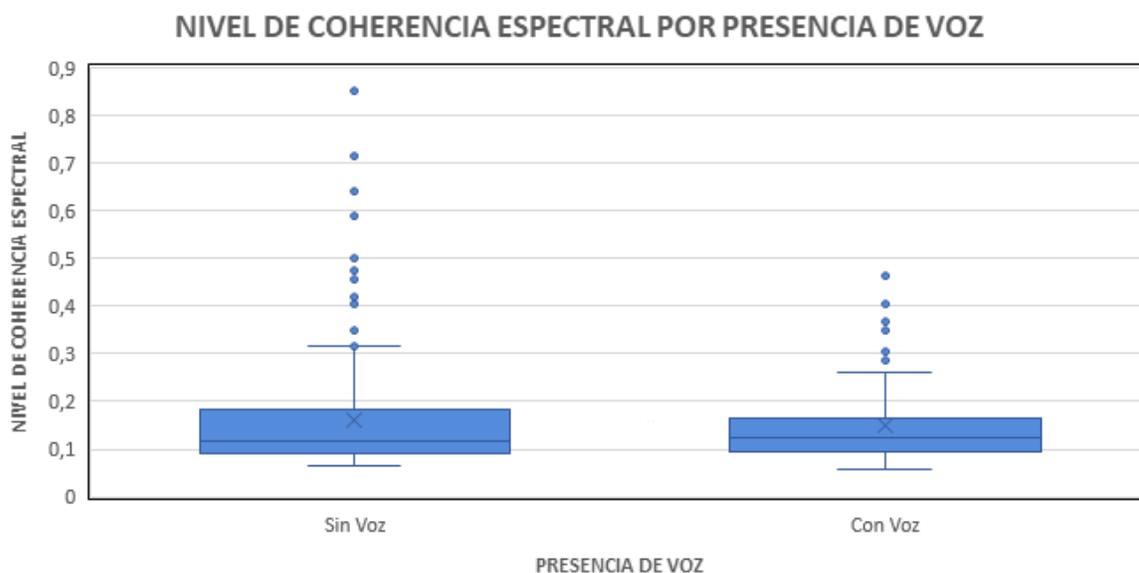


Figura 5. Nivel de coherencia espectral ordenados según la presencia o ausencia de voz.

	Media	Desviación estándar
Pistas musicales sin voz	0,1614466	0,11407756
Pistas musicales con voz	0,1483856	0,08267634

Tabla 12. Valores de media y desviación estándar para los valores de coherencia espectral de las pistas musicales según la presencia o ausencia de voz.

DISCUSIÓN

El objetivo del trabajo consistió en analizar el grado de deterioro de la señal de audio musical, en función de su proporción de características musicales como tono, timbre, melodía y ritmo. Este procedimiento fue realizado mediante un análisis de coherencia espectral entre las pistas musicales originales y las pistas musicales procesadas por el simulador de IC. La comparación de los valores de coherencia espectral permitió estimar los componentes y géneros musicales con un mayor grado de deterioro debido al procesamiento de la señal de audio realizada por el IC. Los principales hallazgos se discuten a continuación:

En el procesamiento del IC, el componente musical que menos se deteriora es el ritmo:

El procesamiento de la señal acústica por el IC provoca el deterioro del mensaje auditivo, cuya distorsión es mayor en señales complejas como la música, perdiendo una gran cantidad de información musical. Si bien esta distorsión se produce en todas las canciones, los componentes de dichas pistas musicales, tales como el ritmo, armonía y melodía, influyen en la percepción de la música, provocando que canciones que difieren en su composición rítmica, melódica y armónica, presenten distintos valores de coherencia espectral.

En nuestros resultados, evidenciamos que el componente musical que menos se deteriora es el ritmo, cuyos valores de coherencia espectral son significativamente mayor que los valores de coherencia espectral para la armonía y para la melodía. Además, logramos establecer estadísticamente que existe un efecto significativo de la predominancia o componentes musicales sobre el deterioro de la pista musical procesada por el simulador de IC.

La mejor percepción de los componentes rítmicos en las pistas musicales, así como la peor percepción de los componentes armónicos y melódicos, ha sido evidenciada en estudios de distintos autores, tales como: Lassaletta et al. (2008), obteniendo mayores puntuaciones en

pruebas de discriminación de ritmo en comparación a pruebas de discriminación de tono y timbre [29]; y Gfeller et al. (1997), donde al estudiar la discriminación de patrones rítmicos y una tarea de 6 pulsos, los usuarios de IC obtuvieron un desempeño similar a sujetos normoyentes [28].

Esta mayor percepción del ritmo se debe a la preservación de la envolvente temporal durante el procesamiento de la señal acústica por el IC, lo que permite una percepción y extracción del ritmo muy cercana a sujetos con audición normal. Sin embargo, la melodía y la armonía se encuentran estructuradas por distintos tonos y timbres, ordenados de forma sucesiva y simultánea, cuya percepción se ha visto significativamente distorsionada gracias al pobre procesamiento de la información espectral por parte del IC, proceso detallado previamente en las limitaciones del IC.

Si bien, la clasificación de predominio se llevó a cabo de la manera más objetiva y rigurosamente posible por los integrantes del trabajo, dada la complejidad de la música, ninguna se encuentra formada completamente por la clasificación asignada. De esta forma, existen componentes musicales que no pertenecen a la clasificación asignada que pueden repercutir en la valorización del puntaje obtenido del análisis de coherencia espectral.

En el procesamiento del IC, el género musical que menos se deteriora es el Beats, el cual es significativamente mejor que los géneros Classic y Rock:

Cada género musical se caracteriza por poseer diferentes características, lo que influye en la percepción musical y, en este trabajo, provoca que se obtengan distintos valores promedio de coherencia espectral dependiendo del género musical.

Las pistas musicales pertenecientes a los géneros Beats, Hip Hop y Electronic procesadas por el simulador de IC fueron aquellas que obtuvieron mayores promedios de coherencia espectral

(Tabla 8). Este fenómeno se podría explicar debido a que estos géneros musicales poseen una gran cantidad de componentes rítmicos y, como mencionamos anteriormente, el procesamiento de la señal acústica por parte del IC se encarga de preservar la envolvente temporal, cuya información es clave para la percepción de los componentes rítmicos.

Por otro lado, las pistas musicales pertenecientes al género Classic procesada por el simulador de IC, fue aquella que presentó peores promedios de coherencia espectral, seguido del género musical Rock y Smallemotions (Tabla 8). Este bajo nivel de coherencia de la música perteneciente al género musical Classic, se podría explicar debido a que, este género en especial, posee un gran componente melódico y armónico en su configuración basada en la sucesión de tonos de distintos instrumentos. Lo anterior se puede relacionar a que los usuarios de IC presentan dificultades para diferenciar cambios en la tonalidad de las pistas, visto en el estudio de Galvin et al. (2009), donde la mayoría de los usuarios de IC presentaron un mal desempeño incluso en la discriminación con una entonación de 5 semitonos [27] o en el estudio de Gfeller et al. (2002), donde el reconocimiento melódico en usuarios de IC era muy bajo en comparación a normoyentes [31]. Este bajo nivel de coherencia espectral se podría correlacionar también con la pobre percepción musical de la música clásica en usuarios de IC, ya que se ha demostrado que los usuarios de IC prefieren la música pop y country regularmente estructurada por sobre la música clásica [6,8,52].

En el procesamiento del IC, la presencia o ausencia de voz no provocó cambios significativos en el nivel de coherencia espectral:

La presencia de voz es un elemento muy utilizado en la música, éste varía entre pistas y género musical, existiendo algunos donde predomina la presencia y, en otros, la ausencia de voz, como en los géneros Hip Hop y Classic, respectivamente. Además, existen géneros donde la voz presenta distintos objetivos y no solo se utiliza para emitir mensajes verbales.

Si bien, en los resultados se observó que aquellas pistas musicales con ausencia de voz obtuvieron mayores valores de coherencia espectral (Tabla 12), la diferencia entre ambos grupos no fue significativa. Estos hallazgos no coinciden con el objetivo fundamental del IC, el cual consiste en la percepción de estímulos verbales fundamentales para la comunicación entre personas. En esta línea, la presencia de voz debería haber aumentado los valores de coherencia espectral, es decir, que la señal procesada por el simulador de IC sea similar a la señal original gracias a los componentes verbales de la pista musical. Sin embargo, esta tendencia puede deberse a otro tipo de factor y no necesariamente a la ausencia o presencia de voz.

Uno de estos factores puede deberse a que consideramos como una pista musical con presencia de voz, aquellas pistas que contenían un extracto mayor o igual a 5 segundos con presencia de oraciones verbales entendibles y/o con un significado, descartando todas aquellas pistas que utilizan la voz como un componente armónico o rítmico. Por lo tanto, algunas pistas musicales clasificadas con presencia de voz apenas contenían un tercio de su duración total con estímulos verbales, lo que pudo haber distorsionado los valores promedio de coherencia espectral para las pistas musicales con presencia de voz.

Por otro lado, debido a la cuantiosa cantidad de pistas trabajadas, fue difícil mantener el mismo margen de pistas con voz y predominio en ritmo, armonía y melodía, así como en la misma proporción que las pistas sin voz y el mismo predominio, lo cual puede haber afectado la valorización de la presencia de voz en las pistas musicales.

CONCLUSIÓN

El disfrute y escucha de la música es parte importante de nuestras vidas, tanto social como culturalmente, permitiendo que los seres humanos se desenvuelven a gusto en distintas situaciones y contextos. Lamentablemente, no todos pueden acceder al disfrute de la música como es el caso de los usuarios de IC, dado por el procesamiento de la señal acústica por parte del IC, la que modifica de forma significativa distintas características musicales importantes.

Como se ha visto, diversos estudios han abordado y estudiado la percepción de los distintos componentes musicales en usuarios de IC, buscando en cuales se obtiene una mejor percepción. Gracias a esto, existen distintos grupos de investigación que trabajan en distintas áreas en busca de aumentar el disfrute de la música en usuarios de IC, por ejemplo, a través de la remezcla de música con el objetivo de que sea más agradable para los usuarios de IC, acentuando elementos como el ritmo y la voz en deterioro de otros elementos como el tono y timbre [53-55].

En este trabajo se estudió la correlación espectral, es decir, la similitud frecuencial entre la señal acústica procesada por el simulador de IC y la señal acústica original, clasificadas según el componente musical predominante, el género musical y la presencia de voz.

En base a los resultados obtenidos se evidenció que:

- Aquellos factores propensos a la conservación corresponden a los elementos rítmicos, en lugar de los componentes armónicos y melódicos, compuestos principalmente por tono y timbre, características musicales deterioradas debido al procesamiento de la señal de los IC.

- Aquellos géneros musicales con un mayor valor promedio de coherencia espectral, tales como Beats, Funk, Electronic y Hip Hop, presentan una mayor cantidad y proporción de elementos musicales rítmicos, cuya información se conserva de mejor manera en el procesamiento del IC, permitiendo que puedan ser percibidos de manera similar a sujetos con audición normal. Por otro lado, aquellos géneros musicales con un menor valor promedio de coherencia espectral, tales como Classic, Country, Rock y Small emotions, presentan una mayor cantidad y proporción de elementos musicales como el tono y timbre.
- Aquellas pistas con ausencia de voz tuvieron un mayor valor promedio de coherencia espectral en comparación a las pistas con presencia de voz, pero este factor no provocó diferencias significativas en el análisis factorial de coherencia espectral.

Con este trabajo, se espera dejar una base de datos para la futura realización de experimentos perceptuales en pacientes, utilizando las pistas procesadas o el script de MATLAB utilizado en este estudio para el procesamiento de las pistas musicales. Los esfuerzos en este trabajo fueron dirigidos en recopilar y mejorar los métodos de análisis permitiendo una mejor valorización de las piezas simuladas.

El desafío actual corresponde a la comprobación mediante estudios perceptuales de la fiabilidad del método matemático planteado, dando la posibilidad de evidenciar si los valores de coherencia espectral coinciden con la percepción de los usuarios de IC y, posteriormente, aplicar este conocimiento con el objetivo de seleccionar o crear música específicamente dirigida a usuarios de IC, aumentando así su disfrute por la música y, por consiguiente, mejorar la calidad de vida de las personas.

BIBLIOGRAFÍA

- [1] World Health Organization. 2021. Sordera y pérdida de la audición. <https://www.who.int/es/news-room/fact-sheets/detail/deafness-and-hearing-loss> (accessed November 21, 2022).
- [2] Informe mundial sobre la audición. 2021. Informe Mundial Sobre La Audición:1–5. <https://doi.org/10.37774/9789275324677>.
- [3] Powell DS, Oh ES, Lin FR, Deal JA. 2021. Hearing Impairment and Cognition in an Aging World. *Journal of the Association for Research in Otolaryngology* ;22:387–403. <https://doi.org/10.1007/s10162-021-00799-y>.
- [4] Blazer DG, Tucci DL. 2019. Hearing loss and psychiatric disorders: a review. *Psychol Med* ;49:891–7. <https://doi.org/10.1017/S0033291718003409>.
- [5] Naples JG, Ruckenstein MJ. 2020. Cochlear Implant. *Otolaryngol Clin North Am* ;53:87–102. <https://doi.org/10.1016/j.otc.2019.09.004>.
- [6] McDermott HJ. 2004. Music Perception with Cochlear Implants: A Review. *Trends Amplif* ;8:49–82. <https://doi.org/10.1177/108471380400800203>.
- [7] Manrique M, Ramos Á, de Paula Vernetta C, Gil-Carcedo E, Lassaletta L, Sanchez-Cuadrado I, et al. 2019. Guía clínica sobre implantes cocleares. *Acta Otorrinolaringol Esp* ;70:47–54. <https://doi.org/10.1016/J.OTORRI.2017.10.007>.
- [8] Nogueira W, Nagathil A, Martin R. 2019. Making Music More Accessible for Cochlear Implant Listeners: Recent Developments. *IEEE Signal Process Mag* ;36:115–27. <https://doi.org/10.1109/MSP.2018.2874059>.

- [9] Implantes cocleares: Cómo funciona, quién los recibe | NIDCD n.d. <https://www.nidcd.nih.gov/es/espanol/implantes-cocleares> (accessed November 22, 2022).
- [10] Jiam NT, Limb CJ. 2019. Rhythm processing in cochlear implant-mediated music perception. *Ann N Y Acad Sci* ;1453:22–8. <https://doi.org/10.1111/nyas.14130>.
- [11] Macherey O, Carlyon RP. 2014. Cochlear implants. *Current Biology* ;24:R878–84. <https://doi.org/10.1016/j.cub.2014.06.053>.
- [12] Wilson BS, Dorman MF. 2008. Cochlear implants: A remarkable past and a brilliant future. *Hear Res* ;242:3–21. <https://doi.org/10.1016/j.heares.2008.06.005>.
- [13] Drennan WR, Rubinstein JT. 2008. Music perception in cochlear implant users and its relationship with psychophysical capabilities. *The Journal of Rehabilitation Research and Development* ;45:779–90. <https://doi.org/10.1682/JRRD.2007.08.0118>.
- [14] Dritsakis G, van Besouw RM, O' Meara A. 2017. Impact of music on the quality of life of cochlear implant users: a focus group study. *Cochlear Implants Int* ;18:207–15. <https://doi.org/10.1080/14670100.2017.1303892>.
- [15] Guylaine Vaillancourt. 2009. Música y musicoterapia: Su importancia en el desarrollo infantil. Madrid, España: Narcea.
- [16] Cabrera IM. 2013. Influence of music on emotions: a brief review. *Realitas: Revista de Ciencias Sociales, Humanas y Artes*.
- [17] Rebecchini L. 2021. Music, mental health, and immunity. *Brain Behav Immun Health*;18:100374. <https://doi.org/10.1016/j.bbih.2021.100374>.

- [18] Mirza S, Douglas S, Lindsey P, Hildreth T, Hawthorne M. 2003. Appreciation of music in adult patients with cochlear implants: a patient questionnaire. *Cochlear Implants Int* ;4:85–95. <https://doi.org/10.1179/cim.2003.4.2.85>.
- [19] Tarabichi O, Jensen M, Hansen MR. 2021. Advances in hearing preservation in cochlear implant surgery. *Curr Opin Otolaryngol Head Neck Surg* ;29:385–90. <https://doi.org/10.1097/MOO.0000000000000742>.
- [20] Jwair S, Prins A, Wegner I, Stokroos RJ, Versnel H, Thomeer HGXM. 2021. Scalar Translocation Comparison Between Lateral Wall and Perimodiolar Cochlear Implant Arrays - A Meta-Analysis. *Laryngoscope* ;131:1358–68. <https://doi.org/10.1002/lary.29224>.
- [21] Macherey O, Deeks JM, Carlyon RP. 2011. Extending the Limits of Place and Temporal Pitch Perception in Cochlear Implant Users. *Journal of the Association for Research in Otolaryngology* ;12:233–51. <https://doi.org/10.1007/s10162-010-0248-x>.
- [22] Bharadwaj HM, Verhulst S, Shaheen L, Liberman MC, Shinn-Cunningham BG. 2014. Cochlear neuropathy and the coding of supra-threshold sound. *Front Syst Neurosci* ;8. <https://doi.org/10.3389/fnsys.2014.00026>.
- [23] Limb CJ, Roy AT. 2014. Technological, biological, and acoustical constraints to music perception in cochlear implant users. *Hear Res* ;308:13–26. <https://doi.org/10.1016/j.heares.2013.04.009>.
- [24] Jiam NT, Caldwell MT, Limb CJ. 2017. What Does Music Sound Like for a Cochlear Implant User? *Otology & Neurotology*;38:e240–7. <https://doi.org/10.1097/MAO.0000000000001448>.

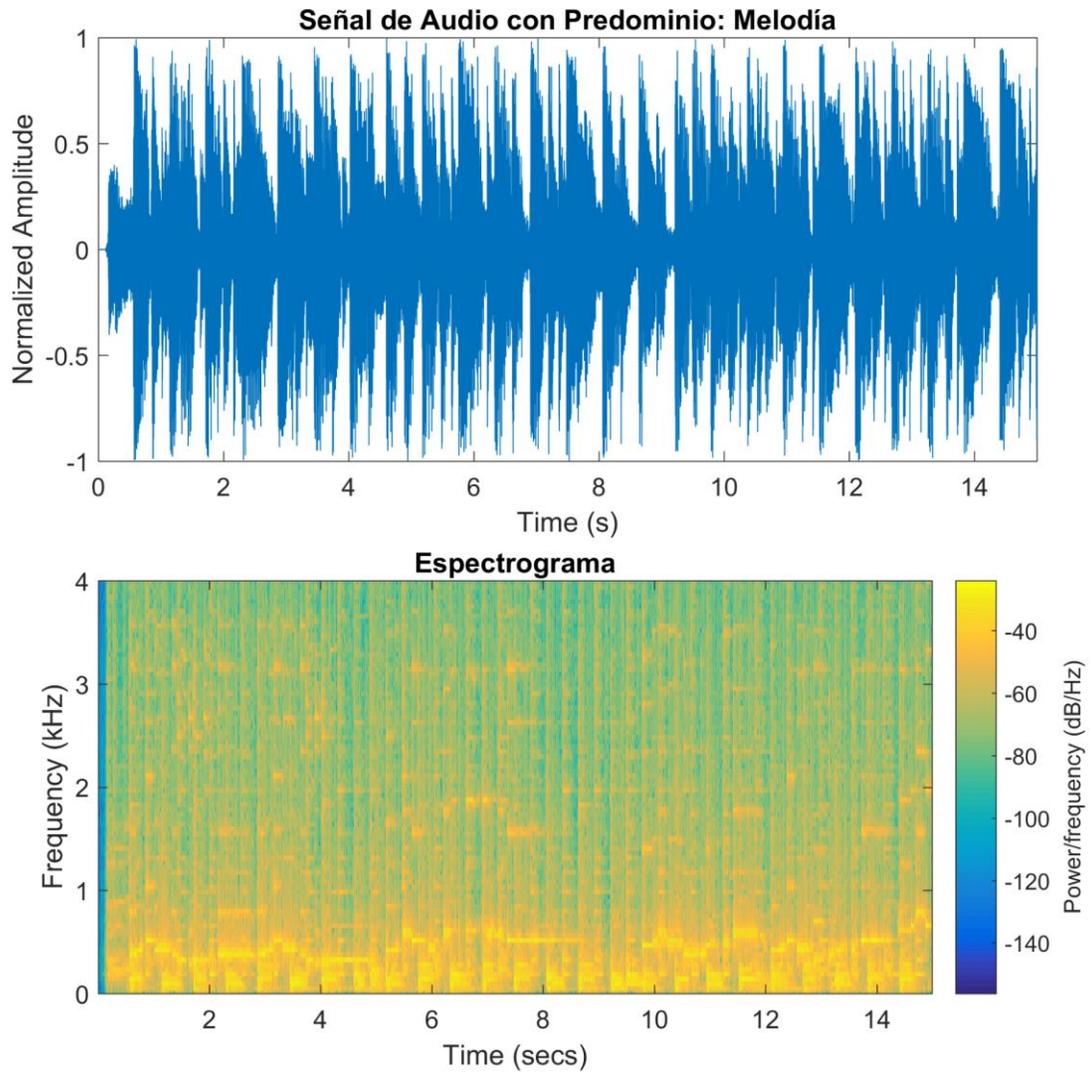
- [25] Sharp A, Delcenserie A, Champoux F. 2018 Auditory Event-Related Potentials Associated With Music Perception in Cochlear Implant Users. *Front Neurosci*;12. <https://doi.org/10.3389/fnins.2018.00538>.
- [26] Leal MC, Shin YJ, Laborde M, Calmels M, Verges S, Lugardon S, et al. 2003. Music Perception in Adult Cochlear Implant Recipients. *Acta Otolaryngol*;123:826–35. <https://doi.org/10.1080/00016480310000386>.
- [27] Galvin JJ, Fu Q-J, Shannon R V. 2009. Melodic contour identification and music perception by cochlear implant users. *Ann N Y Acad Sci* ;1169:518–33. <https://doi.org/10.1111/j.1749-6632.2009.04551.x>.
- [28] Gfeller K, Woodworth G, Robin DA, Witt S, Knutson JF. 1997. Perception of Rhythmic and Sequential Pitch Patterns by Normally Hearing Adults and Adult Cochlear Implant Users. *Ear Hear* ;18:252–60. <https://doi.org/10.1097/00003446-199706000-00008>.
- [29] Lassaletta L, Castro A, Bastarrica M, Pérez-Mora R, Herrán B, Sanz L, et al. 2008. Musical Perception and Enjoyment in Post-Lingual Patients With Cochlear Implants. *Acta Otorrinolaringologica (English Edition)* ;59:228–34. [https://doi.org/10.1016/S2173-5735\(08\)70228-X](https://doi.org/10.1016/S2173-5735(08)70228-X).
- [30] Gfeller K, Witt S, Mehr MA, Woodworth G, Knutson J. 2002. Effects of Frequency, Instrumental Family, and Cochlear Implant Type on Timbre Recognition and Appraisal. *Annals of Otolology, Rhinology & Laryngology* ;111:349–56. <https://doi.org/10.1177/000348940211100412>.

- [31] Gfeller K, Turner C, Mehr M, Woodworth G, Fearn R, Knutson JF, et al. 2002. Recognition of familiar melodies by adult cochlear implant recipients and normal-hearing adults. *Cochlear Implants Int* ;3:29–53. <https://doi.org/10.1179/cim.2002.3.1.29>.
- [32] Looi V, She J. 2010. Music perception of cochlear implant users: A questionnaire, and its implications for a music training program. *Int J Audiol* ;49:116–28. <https://doi.org/10.3109/14992020903405987>.
- [33] Nimmons GL, Kang RS, Drennan WR, Longnion J, Ruffin C, Worman T, et al. 2008. Clinical Assessment of Music Perception in Cochlear Implant Listeners. *Otology & Neurotology* ;29:149–55. <https://doi.org/10.1097/mao.0b013e31812f7244>.
- [34] Peretz I, Champod AS, Hyde K. Varieties of Musical Disorders. 2003. *Ann N Y Acad Sci*;999:58–75. <https://doi.org/10.1196/annals.1284.006>.
- [35] Wipe U B, Kuroiwa R M, Délano R PH. 2013. Trastornos de la percepción musical. *Revista de Otorrinolaringología y Cirugía de Cabeza y Cuello*;73:189–99. <https://doi.org/10.4067/S0718-48162013000200012>.
- [36] Cooper WB, Tobey E, Loizou PC. 2008. Music Perception by Cochlear Implant and Normal Hearing Listeners as Measured by the Montreal Battery for Evaluation of Amusia. *Ear Hear*;29:618–26. <https://doi.org/10.1097/AUD.0b013e318174e787>.
- [37] Brockmeier SJ, Fitzgerald D, Searle O, Fitzgerald H, Grasmeyer M, Hilbig S, et al. 2011. The MuSIC perception test: A novel battery for testing music perception of cochlear implant users. *Cochlear Implants Int*;12:10–20. <https://doi.org/10.1179/146701010X12677899497236>.

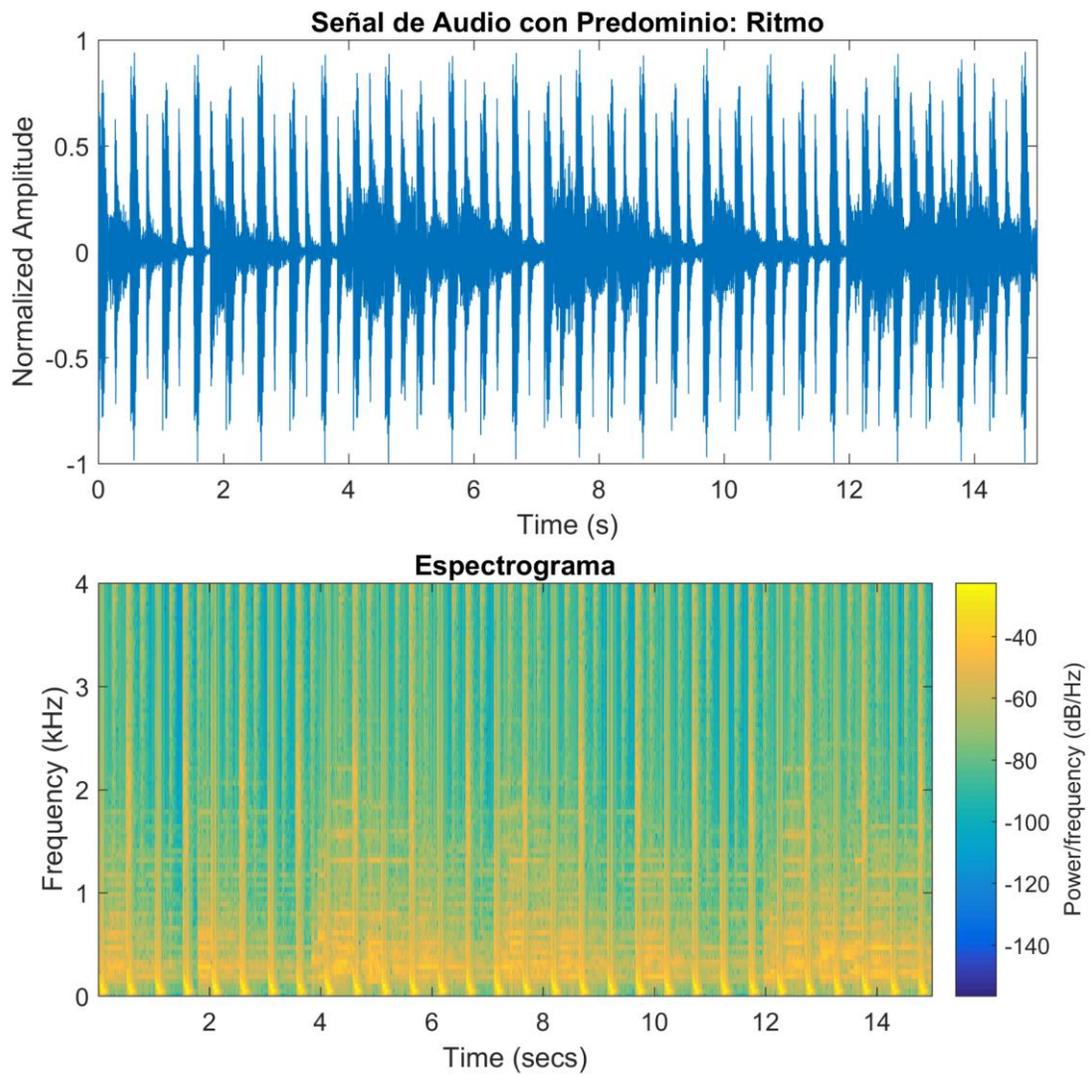
- [38] Dritsakis G, van Besouw RM, Kitterick P, Verschuur CA. 2017. A Music-Related Quality of Life Measure to Guide Music Rehabilitation for Adult Cochlear Implant Users. *Am J Audiol*;26:268–82. https://doi.org/10.1044/2017_AJA-16-0120.
- [39] Ibañez J, Aguilar E, Espinoza C. 2022. PROCESAMIENTO DE PISTAS MUSICALES EN EL IMPLANTE COCLEAR. Universidad de Chile.
- [40] Epidemic Sound. 2023. Epidemic Sound: Royalty-Free Music and Sound Effects . Epidemic Sound. <https://www.epidemicsound.com/> (accessed July 5, 2023).
- [41] Audacity Team. 2023. Audacity® 2023.
- [42] MATLAB. 2023. dir. <https://la.mathworks.com/help/matlab/ref/dir.html>.
- [43] MATLAB. 2023. fullfile. <https://la.mathworks.com/help/matlab/ref/fullfile.html>.
- [44] MATLAB. 2023 saveas. <https://la.mathworks.com/help/matlab/ref/saveas.html>.
- [45] MATLAB. 2023. audiowrite. <https://la.mathworks.com/help/matlab/ref/audiowrite.html>.
- [46] MATLAB. 2022. for. <https://la.mathworks.com/help/matlab/ref/for.html>.
- [47] Závíška P, Rajmic P, Mokrý O. 2022. Audio declipping performance enhancement via crossfading. *Signal Processing*;192:108365. <https://doi.org/10.1016/j.sigpro.2021.108365>.
- [48] MATLAB. 2022. interp1. <https://la.mathworks.com/help/matlab/ref/interp1.html>.
- [49] MATLAB. 2023. mscohere. <https://la.mathworks.com/help/signal/ref/mscohere.html>.

- [50] MATLAB. 2023. kaiser.<https://la.mathworks.com/help/signal/ref/kaiser.html>.
- [51] Bakieva M, González Such J, Jornet J. SPSS: ANOVA de un Factor. Universitat de Valencia.https://www.uv.es/innomide/spss/SPSS/SPSS_0702b.pdf.
- [52] Limb CJ, Roy AT. 2014. Technological, biological, and acoustical constraints to music perception in cochlear implant users. *Hear Res* ;308:13–26. <https://doi.org/10.1016/j.heares.2013.04.009>.
- [53] Pons J, Janer J, Rode T, Nogueira W. 2016. Remixing music using source separation algorithms to improve the musical experience of cochlear implant users. *J Acoust Soc Am* ;140:4338–49. <https://doi.org/10.1121/1.4971424>.
- [54] Tahmasebi S, Gajęcki T, Nogueira W. 2020. Design and Evaluation of a Real-Time Audio Source Separation Algorithm to Remix Music for Cochlear Implant Users. *Front Neurosci* ;14. <https://doi.org/10.3389/fnins.2020.00434>.
- [55] Gajęcki T, Nogueira W. 2018. Deep learning models to remix music for cochlear implant users. *J Acoust Soc Am* ;143:3602–15. <https://doi.org/10.1121/1.5042056>.

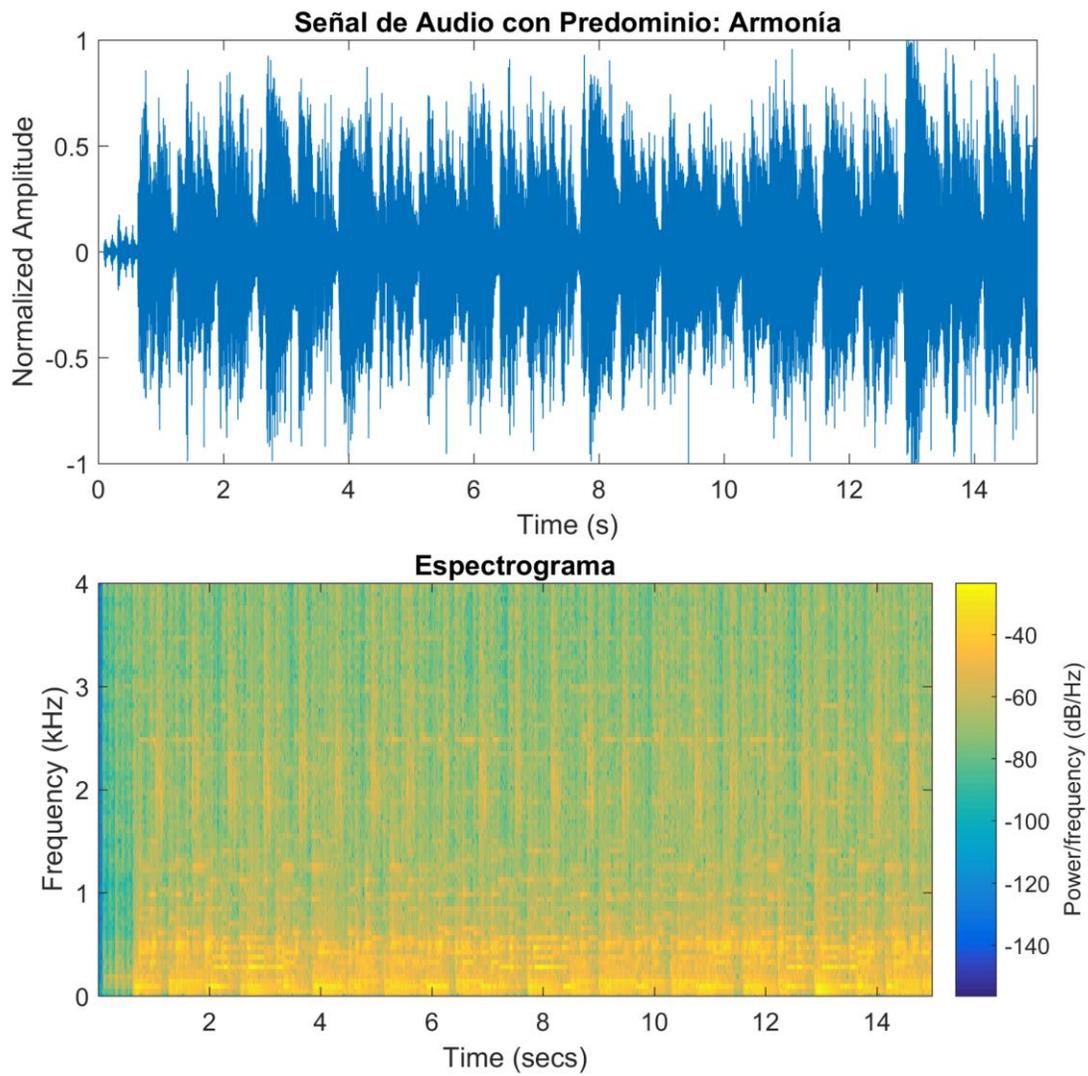
ANEXOS



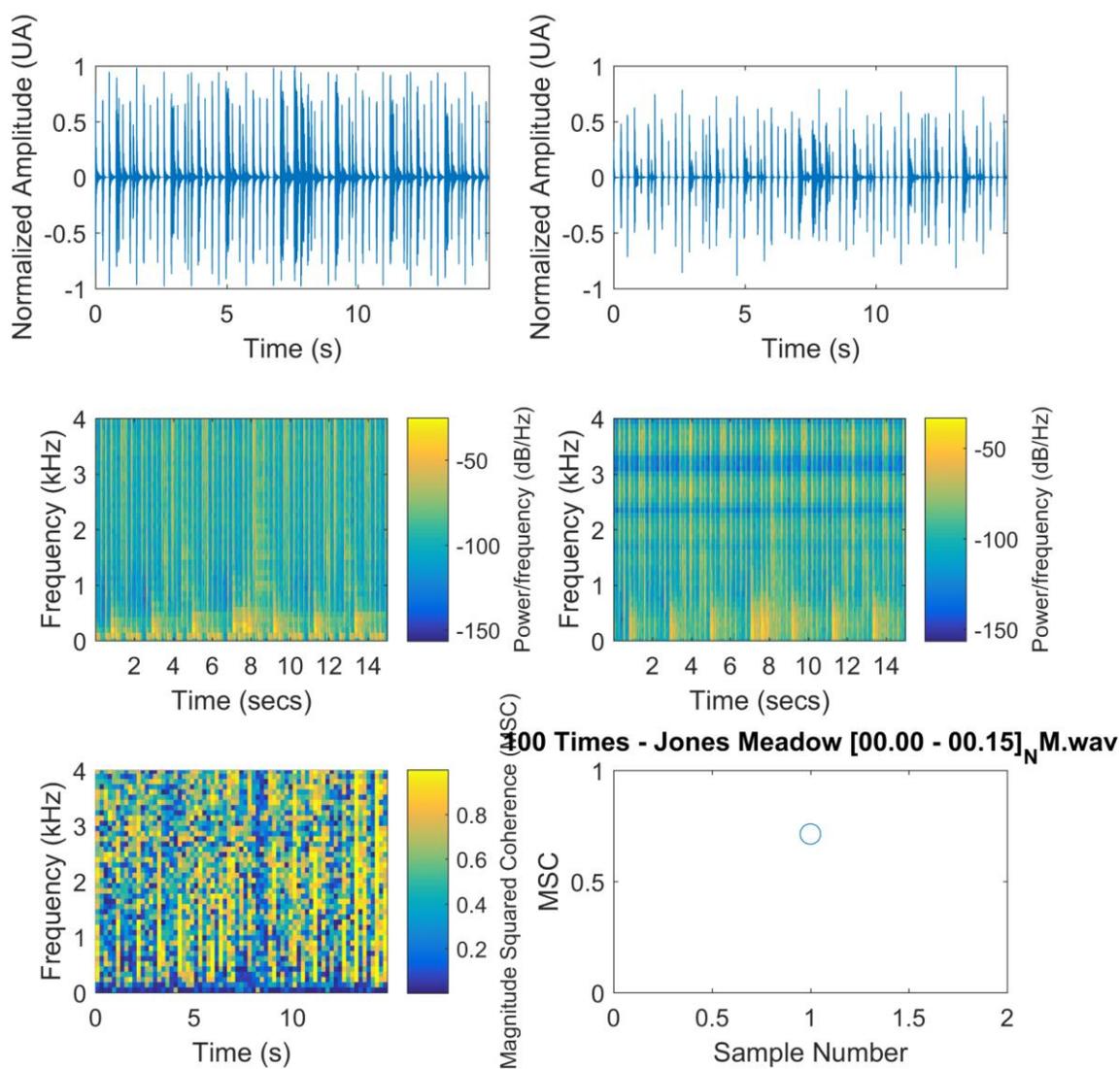
Anexo 1. Gráfico y espectrograma de una señal de audio con predominio melódico dentro de su patrón temporal y frecuencial.



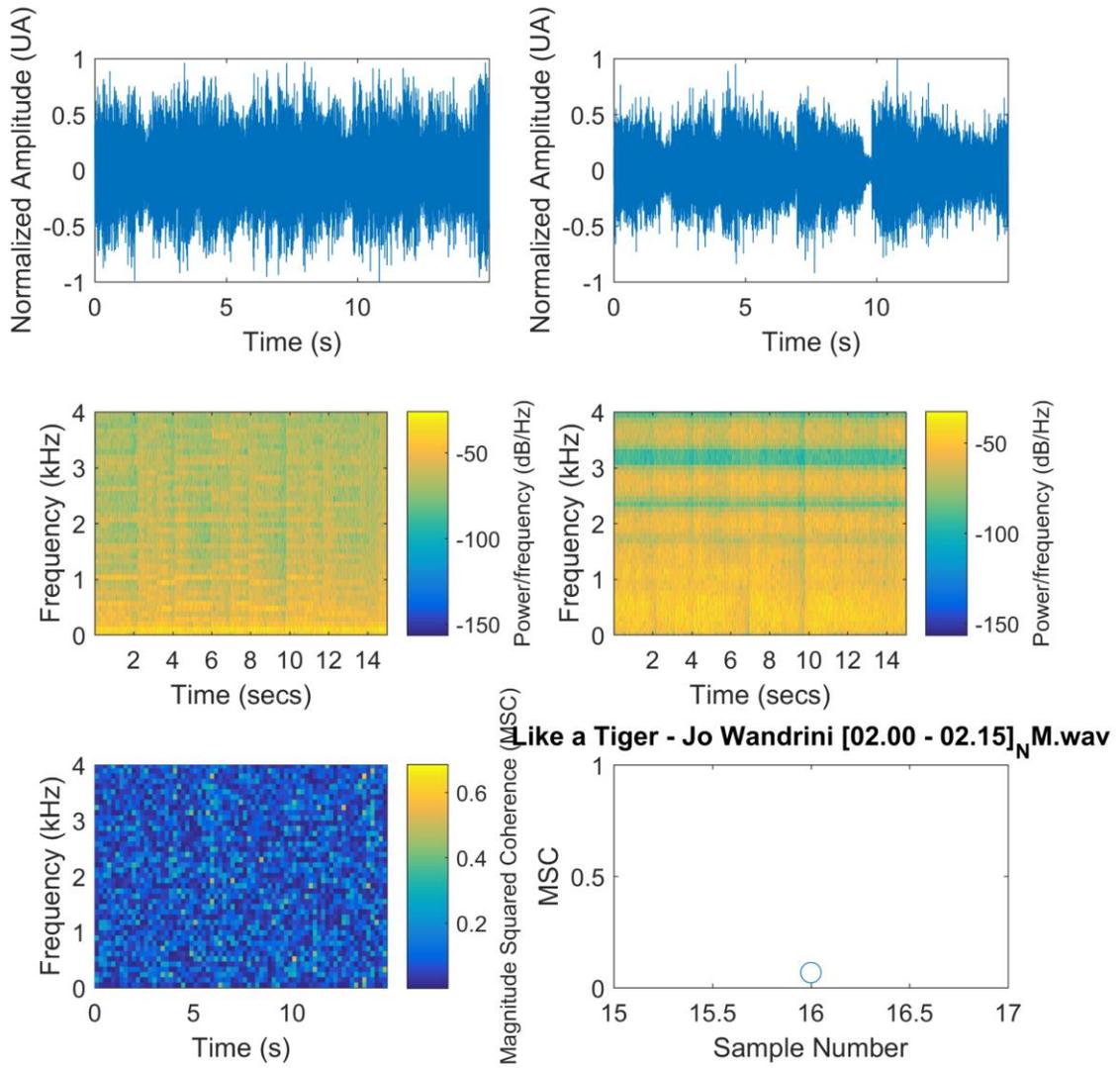
Anexo 2. Gráfico y espectrograma de una señal de audio con predominio rítmico dentro de su patrón temporal y frecuencial.



Anexo 3. Gráfico y espectrograma de una señal de audio con predominio armónico dentro de su patrón temporal y frecuencial.



Anexo 4. Gráficos utilizados en la comparación de la coherencia espectral entre la pista simulada con su versión original con alto valor de coherencia espectral (0.71331).



Anexo 5. Gráficos utilizados en la comparación de la coherencia espectral entre la pista simulada con su versión original con bajo valor de coherencia espectral (0.067448).