

OCR: Diseño de Paisajes Sonoros a través del Texto

OCR: Soundscape Design through Text

BRUNO PERELLI SOTO

Universidad de Chile (Chile)

<https://orcid.org/0000-0002-2461-9639>

bperelli@uchile.cl

RODRIGO DUEÑAS SANTANDER

Universidad de Chile / Universidad de las Américas (Chile)

<https://orcid.org/0000-0001-7035-0215>

rodrigoduena@uchilefau.cl

Recibido: 10 de marzo de 2025

Aceptado: 17 de junio de 2025

Publicación online: 26 de diciembre de 2025

19

RESUMEN:

Este trabajo explora la relación entre texto y sonido a través de la sonificación de letras y palabras utilizando tecnologías como el Reconocimiento Óptico de Caracteres (OCR) y la notación MIDI. A partir del diseño de instrumentos experimentales, se logró transformar los textos en composiciones sonoras, generando un paisaje sonoro que desafía las convenciones lingüísticas tradicionales. La obra fue presentada en dos instancias distintas, demostrando su versatilidad para aplicarse a textos diversos. Inspirada en las demandas de Avraamov sobre la creación de instrumentos con significado social, la obra genera una experiencia sensorial en tiempo real donde escritura y sonido coexisten. Asimismo, se exploran las nociones de performance y performatividad, permitiendo analizar tanto el proceso técnico como el carácter irrepetible de cada presentación. Esta investigación-creación contribuye a nuevas formas de experimentar el lenguaje escrito, resignificando su contenido y ampliando las posibilidades del arte sonoro mediante la tecnología.

PALABRAS CLAVE: Sonificación, Texto, LiveCoding, OCR, Performance.

ABSTRACT:

This work explores the relationship between text and sound through the sonification of letters and words using technologies such as Optical Character Recognition (OCR) and MIDI notation. By designing experimental instruments, texts were transformed into sound compositions, creating a soundscape that challenges traditional linguistic conventions. The work was presented in two different instances, demonstrating its versatility for applying the process to various texts. Inspired by Avraamov's demands for creating instruments with social significance, the piece generates a real-time sensory experience where writing and sound coexist. Additionally, the notions of performance and performativity are explored, allowing the analysis of both the technical process

and the unique nature of each presentation. This research-creation project contributes to new ways of experiencing written language, resignifying its content and expanding the possibilities of sound art through technology.

KEYWORDS: Sonification, Text, LiveCoding, OCR, Performance.

><><><><

1. Introducción

¿Cómo suenan las palabras? La pregunta, por simple que pueda parecer, inevitablemente nos lleva a contestar que, a priori, a través del habla —mediante la conversión de grafemas a fonemas— le otorgamos sonido a las palabras. No obstante, cuestionarse ¿qué sonidos encierran las palabras escritas? propone un acercamiento distinto, puesto que el ordenamiento de caracteres, la construcción de sílabas y, finalmente, de palabras, encierra una notación posible de ser sonificada de una manera distinta a la construcción histórica que como humanidad hemos vivido. Esta es la premisa de la obra OCR: el sonido de las palabras¹, trabajo que desde la experimentación tecnológica y un proceso de investigación-creación plantea una obra-performance en que a través de la recodificación del lenguaje escrito plantea una composición sonora para diversos textos los que ya no son leídos de manera tradicional, sino que son reinterpretados como una partitura de signos y códigos capaces de generar otros sonidos.

20

Esta posibilidad, que desplaza la voz hablada, enfatiza, impone y releva el valor del texto en función del orador, e introduce la noción de *difference* (Derrida, 1967), pues la escritura y el sonido no existen en una relación jerárquica ni secuencial, sino que ambos coexisten, y la escritura suena al mismo tiempo que abre el tiempo (Celedón Bórquez, 2023, p. 352). Es decir, el sonido no es un producto de la escritura, sino que surge simultáneamente, difiriendo sin descanso en una relación indisoluble.

Para John Cage, ‘*La música es sonidos, los sonidos que nos rodean, estemos dentro o fuera de un auditorio*’ (Schafer, 2013, p. 21). Los casos de John Cage², Jean-Luc Godard³ o Arseny M. Avraamov⁴ toman al sonido como un acontecimiento físico, entendiéndolo como principio estético de creación. Esta perspectiva invita a considerar que la escritura no espera ser hablada para sonar; ella misma porta un sonido latente que emerge en el acto de leer y escribir, abriendo nuevas formas de explorar su dimensión sonora. Ellen Lupton (2015) señala que ‘*la tipografía es la encarnación concreta del lenguaje escrito*’. Basándonos en ello podemos quedarnos con la construcción tradicional en que cada letra representa un sonido establecido y codificado. Sin embargo, podemos pensar que otro tipo de grafemas pueden representar otros sonidos, encarnando otros lenguajes y a la inversa los signos y códigos que ya conocemos podrían representar otro tipo de sonidos, desarrollando un nuevo lenguaje en la medida que lo decodificamos y desplazamos la voz hablada por otras técnicas y tecnologías.

Tanto Celedón Bórquez (2023) como Casati y Dokic (1994) definen al sonido como un acontecimiento, algo que difiere y que es independiente del objeto, convirtiéndose en un testimonio de algo que está sucediendo con origen en el actor: ‘...no deja de informar, de decirnos algo sobre el objeto’ actuando como una señal que trasciende su origen físico (Celedón Bórquez, 2023, p. 57). Lo que esta

¹ https://www.instagram.com/reel/C2iFdpMukDp/?utm_source=ig_web_copy_link

² 4'33" (1952) (Cage, 2005)

³ Histoire(s) du cinéma (1988-1998) (Zielinski, 2019, p. 105).

⁴ Symphony of Sirens (1922-1923) (Zielinski, 2019, pp. 261-265).

investigación propone se vincula precisamente a esto: un cambio en el origen físico de la lectura y el habla. No es el cuerpo el que lee y ejecuta el habla, sino que son dispositivos de captura de imagen, sonoros y visuales los que mediante controladores interpretan los códigos generando un nuevo acontecimiento sonoro. Para Schafer, estamos empezando a habitar en un mundo “*con un entorno acústico radicalmente diferente de cualquiera en el que se haya vivido hasta la fecha*” (2013, p.19).

Técnicas como la sonificación, hacen uso del sonido no hablado para transmitir y otorgarle sentido a la información (González-Castañeda et al., 2017). En esta área podemos encontrar las referencias de Temple (2020), quien recurre a la sonificación como método de conversión de datos genéticos en patrones de sonido, y Tay et. al (2021), quienes desarrollan la denominada *Protein Music*, un método que hace una analogía entre las secuencias de proteínas y las notas musicales, donde se mapean diversas propiedades de los aminoácidos a parámetros musicales como tono, duración, octavas y acordes. En este caso el propósito de la sonificación es combinar ciencia y estética, por lo que está diseñada no solo para ser funcional en términos de datos, sino también para mejorar la musicalidad, lo que la distingue de los ejemplos antes citados.

En definitiva, la sonificación en el ámbito científico, transforma datos complejos en patrones sonoros, permitiendo su análisis desde una perspectiva diferente y facilitando la identificación de patrones o anomalías en procesos, por ejemplo, como datos genómicos o biológicos (Temple, 2020; 2024). Este enfoque metodológicamente riguroso en su origen científico establece un puente interesante con la experimentación sonora en las artes. En este campo, la sonificación inspira nuevas formas de creación y exploración estética, permitiendo que el sonido funcione no solo como una representación de datos, sino también como una herramienta de expresión creativa.

A partir de estas definiciones y casos, establecemos que la sonificación transforma eventos y procesos (biológicos) en un acontecimiento sonoro, asimilándose a la manera en que el sonido en el arte es un fenómeno autónomo que trasciende su origen.

21

La música, por otro lado, se articula como una estructura organizada y deliberada de estos acontecimientos sonoros en “... *un tiempo estructurado [...]la colisión entre el mundo de los números y la pasión...*” (Sloterdijk, 2020, p. 7), convirtiendo el testimonio del sonido en una forma estética intencionada. Así, la música es el resultado de la relación entre los sonidos en un contexto organizado de manera formal, muchas veces mediante una escritura que conocemos como partitura y que resulta ser una serie de signos para ser leídos e interpretados por diversos instrumentos, mientras que el sonido en sí mismo sigue siendo un acontecimiento singular y autónomo, con su propio significado independiente del orden que se le quiera asignar.

Para Sloterdijk (2020, p.10), la musicalidad surge cuando el oído adulto se toma un descanso del ruido cotidiano, seleccionando activamente ciertos sonidos para escucharlos de forma distinta, lo que podemos entender como una organización del ruido, o bien una sistematización entre silencios y ruidos, tal como las letras o los códigos que sistematizan escritura en color contra un blanco, el concepto de forma y contraforma tipográfica llevada al sonido.

Este enfoque multidisciplinario refuerza la idea de que los métodos científicos pueden ser adaptados para nutrir la experimentación sonora en el arte, donde el objetivo no es tanto el análisis preciso, sino la creación de nuevos paisajes sonoros (Schafer, 2013) y experiencias sensoriales.

1.1. Escritura y Sonido

Habiendo visto la correlación que las diversas formas de sonificación aportan como método para la representación sonora de datos, algunas veces de carácter científico y otras de carácter artístico, es relevante volver a nuestra pregunta inicial ¿qué sonidos encierran las palabras escritas?

Ancestralmente la escritura y la lectura en voz alta estaban ligadas, la lectura silenciosa era privilegio de individuos letrados capaces de realizar el acto de lectura sin emitir sonido, lo que era visto casi como un acto de magia. Por esto, Celedón Bórquez (2023), nos dice que la escritura y el sonido coexisten: el sonido no es producto de la escritura. Interpretando más profundamente, nos sugiere que la escritura y el sonido se configuran en una especie de equilibrio dinámico. La escritura no solo está a la espera de ser leída en voz alta, tal como en los orígenes de la escritura con fonemas, sino que tiene una dimensión sonora intrínseca que está presente desde el momento en que se realiza. No anticipa el sonido, sino que ambos (escritura y sonido) emergen juntos, abriendo simultáneamente la experiencia del tiempo. El tiempo, en este caso, es tanto el acto de escribir como el de leer o escuchar, todos formando parte de un mismo proceso temporal.

La sonificación lo deja muy claro: los mismos datos (escritos) son escuchados como sonidos. La respuesta procedural a la pregunta comienza a tomar protagonismo en este trabajo. Ya lo anticipaba Lyotard cuando dice en torno al bit, que esta unidad tecnológica, se convierte en la “*nueva moneda de cambio abstracta... para una economía de producción de texto, imagen y sonido, que también incluirá aquellas formas de arte que se expresan y realizan con y a través de los medios*” (Lyotard, citado en Zielinski, 2019, p. 72).

Esta reflexión nos invita a pensar la escritura y el sonido como dos fenómenos que no se subyugan uno al otro, sino que coexisten en una relación fluida y continua, sin poder separarse completamente o medirse mutuamente. La escritura como una representación del sonido puede tomar diversas formas, como en este caso de estudio las letras que representan fonemas que elaboran palabras y estas palabras frases y así discursos, sin embargo, otros signos pueden representar otros sonidos, los que pueden ser interpretados mediante otra codificación. De igual manera pueden existir lenguajes no codificados, interpretables, en desarrollo, o más allá de la comprensión humana propios de otras especies.

22

En este contexto, la *différance*, término filosófico acuñado por Derrida y utilizado por Celedón Bórquez (2023, p. 352), refiere a la diferencia y al aplazamiento simultáneo en el proceso de significado. Esta *différance* entre la escritura y el sonido no puede medirse cuantitativamente; es una diferencia que siempre está presente, pero nunca completamente resoluble. La escritura y el sonido no son intercambiables ni uno deriva del otro, sino que coexisten en una relación de diferencia constante.

La *différance* (Derrida, 1967) se hace evidente en torno al bit descrito por Lyotard, puesto que los bits encierran más que una forma de almacenamiento. Los bits, dependiendo de su contexto de uso precisamente expresan ese diferencial y aplazamiento en los significados: 7 bits, por ejemplo, establece valores binarios entre 0 y 127 y puede precisamente almacenar tanto mapas de caracteres ASCII (American Standard Code for Information Interchange) como notaciones MIDI (Interfaz Digital de Instrumentos Musicales) generando en torno al mismo valor numérico tanto una expresión escrita como una expresión sonora.

El siguiente trabajo da cuenta de esta obra experimental que fusiona el uso del Reconocimiento Óptico de Caracteres (OCR) con la traducción a sonido. Utilizando textos escaneados como fuente, la obra transforma las letras en sonidos, y las palabras en armonías, dándole otra vida a textos que

son reinterpretados sonoramente. Cada palabra detectada se convierte en una armonía única, tejiendo un tapiz auditivo que explora la intersección entre la comunicación visual escrita y la expresión musical. La combinación de OCR y medios analógicos como el fotocopiado de textos (Hyde, 2020) permite sonificar tipografías y palabras y discursos, y transformarlos en originales composiciones musicales.

1.2. Captura y Codificación de Textos

Entendemos que la tipografía es física y que esta ha evolucionado a partir de la escritura manual (Kane, 2012), por lo que iniciar el proceso desde esta captura análoga traduce y adapta los procesos de composición manual de la construcción tipográfica, haciendo referencia a la evolución de la palabra escrita hasta los tipos de metal y su composición manual hasta llegar a la digitalización del proceso. Comprender su condición de objeto físico es validar su característica física de las letras, al igual que el sonido hablado.

Teniendo claridad en cuanto al método de correspondencia entre textos y sonido, el elemento clave reside en la captura de dichos textos. En el pasado tecnologías como los scanner cargaban softwares denominados OCR (Reconocimiento Óptico de Caracteres en español) los cuales capturaban un texto impreso mediante el scanner, cargaba un diccionario según el idioma y establecía una correspondencia entre el texto escaneado y la palabra reconocida.

Sin embargo, aspectos como las demandas expuestas por Avraamov (Zielinski, 2019), la coexistencia entre texto y sonido de Celedón Bórquez (2023), la “*la colisión entre el mundo de los números y la pasión*” de Sloterdijk (2020) y los conceptos de *Performance* y *Performatividad* definidos por Bal (2016) configuran, a nuestro juicio, la inmediatez en el proceso de captura de textos y su correspondiente sonificación.

23

Aquí la primera demanda expresada por Avraamov cobra fuerza: La Construcción de instrumentos experimentales y con significado Social. Los *artefactos construidos* hacen referencia a las tecnologías o instalaciones desarrolladas para producir y controlar el sonido, desde sintetizadores hasta dispositivos interactivos. De esta manera, el arte sonoro invita a una reflexión tanto racional como sensorial, donde el sonido es explorado como un fenómeno autónomo y no únicamente como un acompañante de lo visual (Zielinski, 2019).

1.3. Performance y Performatividad

Dos aspectos importantes por distinguir son la *performance* y la *performatividad*. Para Bal (2016) esta distinción favorece el análisis cultural de la obra, una distinción entre lo público y lo privado siendo la *performance* la producción hábil y reflexiva de la obra por parte de los intérpretes. En cambio, la *performatividad* se refiere al acto en sí mismo, que tiene lugar en un presente único. Estos dos conceptos no son excluyentes entre sí, por el contrario, se complementan de tal manera que la producción técnica y hábil de la *performance* a través del uso de artefactos y tecnologías, y por otro, la *performatividad* de la acción en el presente, se activa una experiencia irrepetible que confronta la memoria y el sonido en tiempo real.

1.4. El Paisaje Sonoro a partir del Texto

El concepto de paisaje sonoro propuesto por Schafer (2013) nos permite comprender la dimensión ecológica del entorno acústico. Para el autor, el paisaje sonoro es todo aquello que suena en un entorno determinado y configura una experiencia auditiva significativa. Este acercamiento del sonido para Schafer suma al acontecimiento (Celedón Bórquez, 2023; Casati y Dokic, 1994) una construcción social, cultural e histórica que refleja cómo nos relacionamos con nuestro entorno a través del oído.

Desde esta perspectiva, la obra no busca únicamente transformar letras en notas musicales, sino generar una experiencia acústica que resignifique el texto como materia sonora. Tal como plantea Schafer, vivimos en un entorno sonoro que hemos ido transformando sin mayor conciencia de sus consecuencias, y por tanto debemos aprender a escucharlo críticamente, con atención al diseño sonoro que producimos y habitamos (Schafer, 2013, p. 19)

El uso del Reconocimiento Óptico de Carácteres (OCR) como mecanismo de lectura del texto impreso y su posterior traducción en notas MIDI propone una reorganización del paisaje sonoro a partir de la escritura: El texto deja de ser un código para la lectura visual y se convierte en una partitura audible. Este desplazamiento propone un nuevo modo de habitar el lenguaje escrito, como una superficie que no solo se ve, sino que también se escucha.

A partir de esto, el proyecto constituye un paisaje en cuanto al diseño sonoro: así como es posible diseñar visualmente un espacio, también es posible hacerlo desde lo sonoro, proponiendo un entorno auditivo que dialogue con quien lo habita (Schafer, 2013, pp. 35–37). La obra aquí descrita contribuye a esa reflexión mediante un procedimiento artístico que convierte la lectura en escucha, transformando el texto en paisaje sonoro y posicionando al espectador como oyente activo.

Tal como mencionamos anteriormente, el *bit* descrito por Lyotard encierra como unidad tecnológica una versatilidad en cuanto la salida e interpretación del dato almacenado (texto, imagen y sonido). La binariedad de 1 bit, por ejemplo, encierra diversas interpretaciones más allá de un 0-1: Si el valor del *bit* es 0 no almacena nada, pero también, significa que estamos frente a un color Negro, una condición de Falso o un No como respuesta. Mientras que un valor 1 del bit almacena un color Blanco, una condición Verdadera o un Sí como respuesta.

En los procesos de sonificación antes referenciados, la correspondencia sonora en estos experimentos no es aleatoria, sino que está cuidadosamente mapeada a partir de propiedades fisicoquímicas y biológicas de las proteínas, lo que permite crear una experiencia musical que mantiene un vínculo con las estructuras y funciones biológicas de las moléculas. La música resultante puede variar en términos de melodía, armonía y dinámica dependiendo de las características de la proteína.

2. Materiales y Métodos

La metodología parte de la premisa de crear un paisaje sonoro artificial —en el sentido de Schafer (2013, p. 35–37)— mediante la codificación ASCII-MIDI del texto. Cada carácter es capturado por una cámara, codificado (ASCII) y convertido en una nota (MIDI), no solo como señal, sino como parte de un ecosistema sonoro diseñado. Igual que en el diseño acústico, los participantes actúan simultáneamente como audiencia e intérpretes de ese entorno construido.

A partir de los aspectos antes abordados previamente en torno a la sonificación: La traducción de expresiones en sonido y musicalización de datos, además de la revisión de métodos y referentes en el ámbito artístico y científico, nos proponemos responder a la pregunta que motiva esta obra: ¿Cómo suenan las palabras?

En primer lugar, tenemos una fuente de información de la cual emergen simultáneamente Texto y Sonido. En segundo lugar, identificamos en el *bit* un entorno de unidad que favorece la traducción de textos a sonido dada su abstracción en textos, imagen y sonido. En tercer lugar, Zielinski (2019) enfatiza que el arte sonoro, la creación y manipulación del sonido, no es simplemente un medio para la comunicación o el entretenimiento, sino un material que puede generar tanto placer intelectual como sensorial. El sonido puede ser percibido, manipulado y experimentado, utilizando artefactos o

herramientas diseñadas para generar o modificar los paisajes acústicos. En este punto Avraamov (1923), citado por Zielinski (2019), expresa demandas en torno a una *nueva ciencia musical*. Estas demandas consideran:

1. Construcción de instrumentos no solo un significado de laboratorio, sino también un colosal significado social.
2. Acústica topográfica: Análisis de las condiciones para el poderoso tintineo de los aparatos musicales.
3. Problemas de diseño en el proceso de creación musical: las leyes de la composición bajo la condición de representaciones artísticas al aire libre, el cambio del paisaje sonoro de la vida urbana.

Finalmente, en cuanto a la integración y producción de obra, y en concordancia con el problema de diseño propuesto por Avraamov, Bal (2016, pp. 104-108) nos habla de la *Performance* y *Performatividad* de la obra. El cambio del paisaje sonoro de Avraamov (1923), se desarrolla a partir de la interacción de *Performance*, entendida como la producción hábil y reflexiva de una obra por parte de intérpretes, y *Performatividad*, entendida como el acto en sí mismo, que tiene lugar en un presente único en el que la memoria hace sus trucos.

Para llevar a cabo estos requerimientos, se proponen cuatro etapas para el desarrollo de la obra. En primer lugar, se aborda la correspondencia entre texto y música mediante la construcción de instrumentos. La exploración de distintas formas de vincular el texto escrito con su musicalización tiene como objetivo establecer cómo pueden ser sonificados los caracteres, sílabas o palabras, permitiendo así su interpretación a través del sonido.

En segundo lugar, la captura, procesamiento y codificación de textos para su musicalización detalla el método de transformación de información escrita en música (acústica topográfica). Un elemento importante por destacar en esta etapa es que uno de los desafíos impuestos para el desarrollo de la obra es la inmediatez en el proceso de captura e interpretación musical del texto escrito.

En tercer lugar, una vez establecidas las correspondencias musicales del texto escrito, su captura y codificación en notas musicales, procedemos a estructurar la sonificación de la obra (Problemas de Diseño). En esta etapa se contemplan la elección de textos, la salida del sonido mediante sintetizadores y las primeras iteraciones.

Finalmente, la integración y producción da cuenta de la realización de la *performance* y *performatividad* en dos instancias de circulación de obra detallando los aspectos de composición de textos, su captura, codificación y sonificación final.

2.1. Correspondencia entre Textos y Música

Frente al desafío de establecer correspondencias entre caracteres y sonidos, el *bit* de Lyotard aflora como una solución considerando que actualmente, en entornos digitales, tanto los Mapas de Caracteres como la Notación MIDI se expresan ambas en 7 bits (valores binarios de 0-127).

De esta manera el proceso de sonificación de la obra concibe captura de textos, la separación en caracteres ASCII, su almacenamiento y posterior traducción a Notación MIDI quedando, por ejemplo, de la siguiente manera:

Carácter	Valor ASCII	Valor MIDI	Nota
R	82	82	A5
r	114	114	F#7

Tabla 1. Tabla de correspondencia entre caracteres ASCII y notación MIDI.

Esta tabla de correspondencia entre los caracteres del alfabeto y su equivalencia en valores ASCII y MIDI permite traducir texto a sonido utilizando una codificación común de 7 bits. Por ejemplo, el carácter “R” (mayúscula) tiene un valor ASCII de 82, lo que se traduce directamente en una nota MIDI A5, mientras que “r” (minúscula) corresponde al valor 114, equivalente a F#7 en notación musical.

A partir de esta correspondencia la palabra escrita comienza a estructurar su sonido en cuanto a los caracteres que la componen. La distinción entre mayúsculas y minúsculas aporta otro aspecto importante dado que una palabra completamente en minúsculas sonará distinto de la misma palabra escrita completamente en mayúsculas e incluso con mayúscula inicial. De esta manera “Sonido y Palabra” como frase queda estructurado de la siguiente forma:

Carácter	Valor ASCII	Valor MIDI	Nota
S	83	83	B5
o	111	111	D#8
n	110	110	D8
i	105	105	A7
d	100	100	E7
o	111	111	D#8
y	121	121	C#9
P	80	80	G#5
a	97	97	C#7
l	108	108	C8
a	97	97	C#7
b	98	98	D7
r	114	114	F#8
a	97	97	C#7

Tabla 2. Correspondencia ASCII-MIDI aplicado a la frase “Sonido y Palabra”.

Otro aspecto importante de esto es que los Mapas de Carácteres no almacenan solo letras: Glifos como la puntuación, acentuaciones e incluso espacios tienen su correspondiente en bits lo que por consiguiente le otorga al sonido un ritmo, énfasis, duración e intención a la par con su origen textual.

2.2. Construcción de Instrumentos Experimentales

En torno a la captura de los textos se opta por un inicio análogo y físico a partir de impresos modificados y re-editados mediante el proceso de fotocopiado. La construcción de instrumentos experimentales se traduce en el cruce entre hardware y software. Para ello se diseñaron dos instrumentos mediante software: El primero, es el instrumento de captura el cual mediante Computer Vision (OpenCV), programación de un *script* en Python con la librería *Python-tesseract* que permite implementar OCR mediante Python y la librería CSV que crea una base de datos CSV destinada a almacenar los textos escaneados.

Utilizamos el verbo *escanear* pese a no contar con un scanner físico como antaño. El uso de operadores y algoritmos de visión por computador nos permitió reemplazar el scanner por una Webcam generando así la captura de textos.

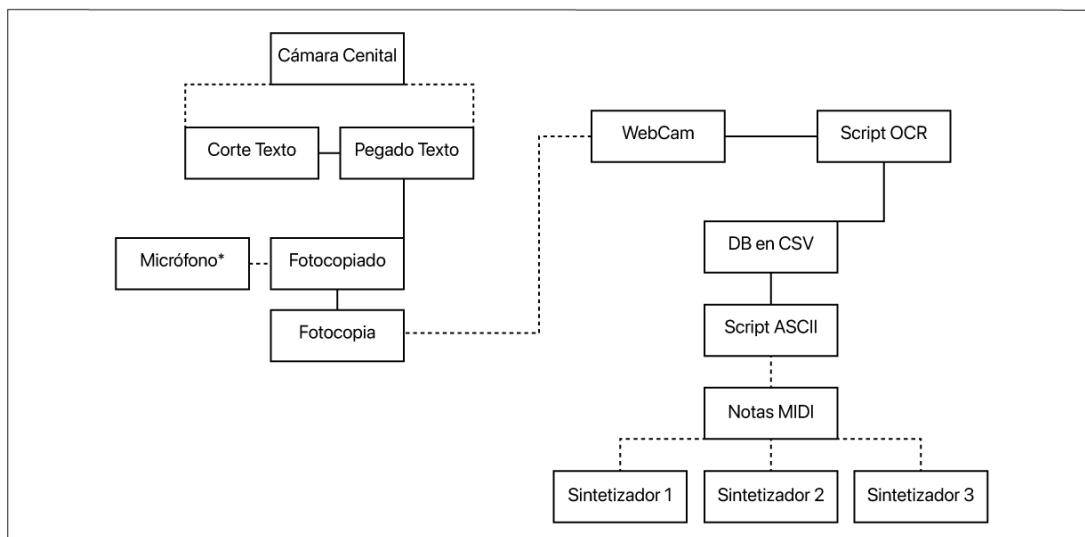


Figura 1. Esquema de montaje de obra.

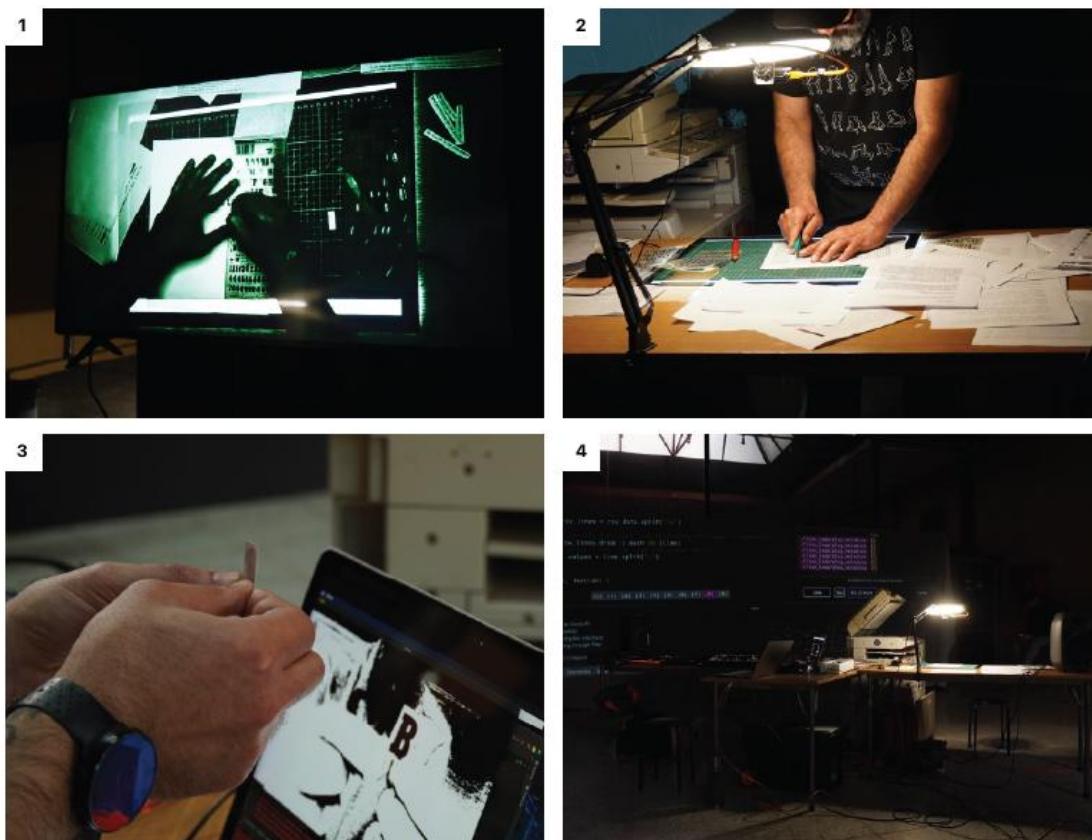


Figura 2. OCR: El Sonido de las Palabras. Los cuatro momentos más importantes de la obra: (1) Composición de Textos (Cámara Cenital); (2) Fotocopiado de Frases a sonicificar; (3) Captura de Textos mediante Webcam y Script utilizando OpenCV y python-tesseract además de su proyección; (4) Vista general de la obra durante proceso de sonificación con SonicPi. Fotografía Stephanie Orellana (Dirección de Investigación, Facultad de Arquitectura y Urbanismo. Universidad de Chile. Octubre de 2023).

El procedimiento de captura queda finalmente organizado de la siguiente manera: En primer lugar, la cámara web captura y detecta el impreso contenedor de textos. En segundo lugar, OpenCV convierte el modo de color proveniente de la cámara a escala de grises, generando una imagen similar a la que obtendríamos de una fotocopiadora (fig. 2). Posteriormente, *python-tesseract* comienza a reconocer palabras y construir las listas (*strings*) de textos según el diccionario definido (en el caso de la obra, español). Estas listas de caracteres detectados son almacenadas en la base de datos CSV la cual no solo almacena el texto, sino que lo separa glifo por glifo y posteriormente le asigna su valor numérico en ASCII. De esta manera una lista de caracteres procedentes de un texto queda estructurados como una lista/partitura de notas interpretables mediante MIDI (fig. 2). Cabe señalar que la lista una vez que deja de detectar textos, espera un periodo de tiempo de quince segundos para borrar su contenido a la espera de nuevos textos escaneados que entren en el *loop*.



28

Figura 3. Esquema de montaje de obra. Fotografía Stephanie Orellana (Dirección de Investigación, Facultad de Arquitectura y Urbanismo. Universidad de Chile. Octubre de 2023).

En este punto la base de datos se constituye en la partitura entre instrumentos. El segundo instrumento, diseñado es un script para *Sonic Pi* tiene como trabajo el acceder a la partitura de caracteres MIDI transformados a partir de las listas de caracteres ASCII. El código no solo interpreta las notas, sino que establece aspectos como la velocidad o tempo expresada en BPM (Beats per Minute), la velocidad de lectura de notas y el canal MIDI por el cual saldrá la señal. Cabe destacar que la elección de *Sonic Pi* no es azarosa, la capacidad de modificación del código durante la reproducción de sonidos, el *live coding*, permite establecer grados de intencionalidad y cambios en el desarrollo de la partitura.

2.3. Sonificación de Textos

Habiendo cumplido con la primera de las demandas de Avraamov (1923), es decir, la construcción de instrumentos que responden a la coexistencia texto-sonido, la inmediatez en la captura y condiciones tanto experimentales como sociales de estos, pasamos a la segunda demanda: Plantear una visión avanzada y experimental sobre cómo el sonido interactúa con el entorno físico y arquitectónico.

Bajo esta premisa, la sonificación de los textos, va más allá de la mera producción de sonido, también es importante considerar la capacidad del sonido de llenar y modificar un espacio transformando la

percepción de quienes son espectadores de la performance. De esta forma se propone generar un *acontecimiento acústico* que transforme el entorno circundante a la obra en una especie de instrumento musical masivo, donde el espacio y la acústica se combinan para crear una experiencia única.

Es importante entonces, dar cuenta que la circulación de la obra consta, hasta la fecha, de dos presentaciones. La primera fue en el Foro de las Artes 2023 de la Universidad de Chile, cuya convocatoria titulada "Memorar al porvenir" hacía eco de los cincuenta años del Golpe Cívico Militar (1973-1990), vinculando el arte con la memoria y la proyección futura de este acontecimiento histórico. La segunda tuvo lugar en el festival Soundtiago: cuerpo, sonido y espacio, organizado por la Universidad de las Américas en enero de 2024. Este evento, bajo la convocatoria "Amnesia", reflexionaba sobre el Chile a cincuenta años del golpe, postpandemia y postestallido social. El festival emergió como una propuesta que desafía el olvido, confrontando la memoria histórica de los lugares y la resiliencia del recuerdo, convertido en sonido en una ciudad en constante transformación.

Para el Foro de las Artes 2023 la selección de textos a sonificar fueron extractos de tres discursos del presidente Salvador Allende Gossens: El primero, realizado en la Universidad de Guadalajara, (México, 1972); el segundo, frente a la Asamblea General de Naciones Unidas (1972) y finalmente, el Discurso al país emitido el mismo 11 de septiembre de 1973 desde el palacio de la Moneda. Mientras, para el festival Soundtiago 2024, trabajamos con extractos de artículos sobre la propuesta de Constitución Política de la República de Chile de 2022.

2.4. Puesta en escena

La puesta en escena de la obra consiste en una instalación de estos artefactos de sonido, sumado a una fotocopiadora que permite seleccionar y extraer palabras o frases claves dentro de los textos utilizados. Este sistema de sonorización de lenguaje impreso se compone de una primera parte completamente análoga (fotocopiado, corte y montaje) que es registrado y expuesto mediante una cámara cenital sobre el área de trabajo, y la proyección de esta imagen a la audiencia, que se vuelve testigo del proceso. El uso de la fotocopiadora pone en evidencia el origen impreso de la palabra y su reproducción ilimitada mediante un proceso que activa la memoria del registro original (palabra escrita) y representa el concepto de la copia no autorizada, o bien el "robo" de un texto, enmarcado en el concepto de *Bricoleur*, desarrollado por Claude Lévi-Strauss en los años sesenta, ya que opera no con materias primas, sino con fragmentos de otras obras, reciclando elementos para darles nueva vida y ponerlos en un sentido distinto. En el *bricoleur* "*alguien que sin buscar obtener resultados y rendimientos, trabaja con los materiales a disposición [...] desarrolla una ciencia que se manifiesta a través de un pensamiento no disciplinado, en estado salvaje*" (Mata, 2018, p. 121).

29

De esta manera el discurso escrito es apropiado por los ejecutores, para la posterior recodificación sonora al que será sometido. Esta apropiación funciona como pequeños poemas visuales, tal como los ejercicios del poeta visual y sonoro Bob Cobbing en Inglaterra en los años sesenta y del artista Guillermo Deissler en Chile durante los años setenta, quienes a partir del uso de la fotocopiadora como herramienta, desarrollan poemas visuales basados en la repetición, ampliación y deformación del texto como son los casos de *Typstract 2* (1972), de Cobbing y *Xeno/Fremd* (1992) de Deissler, configurándose ambos casos como referentes directos para esta acción en la performance.

Una segunda parte de esta escenografía performática se compone de los controladores y procesadores de sonido e imagen, que registran la palabra o frase impresa mediante la cámara que escanea con tecnología OCR, la que es leída por los procesadores MIDI dispuestos para esto. El acto en sí es un proceso que asemeja un trabajo de oficina, un proceso mecánico, apurado, a veces desesperado por alimentar la lectura de los procesadores, al decir del músico y artista Martin Gubbins "*una máquina de poesía visual y sonora*", la que, a partir de la lectura de palabras impresas y su reinterpretación, emite

sonidos no codificados por el público los que perciben un paisaje sonoro aleatorio de armonías y melodías.

Si bien durante la performance la tendencia del público fue la de buscar naturalmente una relación entre lo que escuchaba y las letras que lo originaban, el interés del proyecto radica más en la experiencia sonora que en una decodificación literal del texto. Eso se hizo evidente dada la deformación del sonido que impidió realizar esta acción, generando una arquitectura de sonido a veces caótica que se transforma en lenguaje abstracto y que finalmente fue entendido como “música”.

A nivel escenográfico, la performance funciona con dos pantallas que proyectan, una la toma cenital del área de trabajo de corte y pegado de textos, y otra en que se puede ver el escaneo de las palabras mediante la cámara con OCR y el script de *Sonic Pi* que procesa el sonido. La proyección de las palabras a gran escala, como una imagen en alto contraste, condición necesaria para que el OCR opere, dialoga visualmente con la estética de las fotocopias del proceso análogo, lo que alimenta la relación poética entre lenguaje y el sonido. En ambas presentaciones se trabajó con textos significativos para la comunidad, en una primera oportunidad (Foro de las Artes 2023) con discursos emblemáticos del Presidente Salvador Allende, y posteriormente con los textos de la propuesta de nueva Constitución emanada de la Asamblea Constituyente que fuera organizada con posterioridad al Estallido Social ocurrido en Chile el año 2019, y que tenía por objetivo reemplazar la actual constitución impuesta durante la dictadura de Augusto Pinochet, y que sin embargo fue rechazada en un proceso de elección popular. La apropiación de estos textos, para resignificarlos sonoramente, propone un acto de memoria y deconstrucción de los discursos, un acto de lectura consciente que mediante el *copy/paste* se resignifica, encontrando nuevos y diversos énfasis en ellos para su posteriormente reinterpretarlos mediante una codificación científica y matemática, que los ubica en un medio distinto al habla o la escritura, construyendo una nueva obra inesperada a partir de lo que podría ser “un simple” acto de lectura.

30

Por otra parte, la proyección del procesamiento y conversión de las letras en códigos en tiempo real le proporciona al espectador algunas señales que permitan entender o simplemente seguir el procesamiento de sonificación que está presenciando.

Con estos elementos la performance funciona como una cadena de acciones en diversos planos: en secuencia lineal, primero uno de los ejecutores corta y fotocopia los textos, extrayéndolos de su formato original, configurando pequeños poemas visuales los que pasa a un segundo ejecutor quien los captura y procesa ejecutando las programaciones necesarias para que los textos capturados de manera análoga sean codificados de manera digital y sonificados para componer la obra musical y finalmente ser emitidos a la audiencia.

3. Resultados

Los resultados de esta obra-performance evidencian cómo el proceso de sonificación y resignificación de textos permite la creación de una arquitectura sonora que desafía las convenciones tradicionales del lenguaje escrito llevándolo a lenguajes no codificados. A través del uso de tecnologías como el OCR y la notación MIDI, se ha logrado transformar los textos en composiciones sonoras, demostrando la versatilidad del diseño de instrumentos y la capacidad de cualquier texto para generar una experiencia sensorial única. Estos resultados no solo responden a las demandas de Avraamov (1923 en Zielinski, 2019) en cuanto a la creación de instrumentos experimentales, sino que también abren nuevas posibilidades en la interacción entre texto y sonido.

La obra logra una resignificación del texto escrito a través de su traducción a sonido. Este proceso muestra que cualquier texto, independientemente de su origen o contenido, puede ser sonificado

mediante los instrumentos creados específicamente para este fin. La capacidad de los dispositivos OCR y MIDI para convertir caracteres en sonidos ofrece un resultado que trasciende lo lingüístico, proporcionando una nueva forma de explorar el significado del lenguaje escrito a través del sonido.



Figura 3. Circulación de Obra. Foro de las Artes 2023 y Festival Soundtiago 2024.

La obra se presentó en dos instancias, a diferentes audiencias en diferentes contextos y con diferentes textos, una en el Foro de las Artes 2023, donde se sonificaron los discursos de Salvador Allende, y luego en el festival Soundtiago 2024 que se trabajó con extractos de la fallida propuesta constitucional. En ambos casos se utilizó la misma metodología y la intención artística y en ambos casos el resultado fue sorpresivo para la audiencia. El proceso de lectura colectiva mediante las pantallas y el seguimiento del sonido como un paisaje sonoro genera un atractivo que permite proyectar la obra a cualquier tipo de texto, demostrando la versatilidad de la conceptualización realizada y de los instrumentos creados.

Ambos casos sirvieron como apertura del proceso para probar e iterar los instrumentos, acciones y puesta en escena, realizando modificaciones entre iteraciones. En la primera oportunidad las secuencias de sonido resultaron ser muy largas lo que hizo monótono el sonido emitido. Gracias al control de BPM propio del script en Sonic Pi y el *Live Coding* esto fue modificado para la segunda apertura lo que permitió activar la escucha y hacer más comprensible el proceso que propone la obra. Por otra parte, en la primera puesta en escena se trabajó con dos pantallas, una con el registro del proceso de corte y pegado de textos y otra con el procesamiento de la cámara OCR. Al estar estas dos pantallas enfrentadas en el espacio, no sucedió una dialéctica entre ambas imágenes, generando confusión en el espectador. Para la segunda puesta en escena se trabajó con una sola pantalla la que emitía alternadamente los procesos de corte y pegado y escaneo mediante un switch de cámaras operado por los mismos ejecutores de la performance. Este recurso permitió marcar mayor intencionalidad en el proceso, ya que el traspaso de material (entregar la hoja impresa para su escaneo) era también marcado por el cambio de imagen en el momento exacto en que los ejecutores lo deseaban. En la segunda performance, también se incorporó un micrófono que registraba el sonido de la fotocopiadora (bip de botones, golpes del movimiento de la tapa, ruido de rodillos de la salida del papel) y que lo incorporaba en el diseño sonoro, este simple anexo permitió darle mayor relevancia al acto de fotocopiado, integrándolo a la performance de manera intencionada y agregando una capa de sonido activa y análoga.

La obra responde a las demandas de Avraamov (Zielinski, 2019) en cuanto a la creación de instrumentos experimentales que no solo tienen un propósito técnico, sino también un significado social. La construcción de estos instrumentos permite resignificar el contenido textual, no sólo como un medio de comunicación, sino como un material sonoro que transforma el texto en una experiencia sensorial. Los textos dejan de ser entendidos únicamente como un conjunto de palabras para ser interpretados como un paisaje sonoro a través del diseño específico de los instrumentos.

La inmediatez del proceso de captura y traducción sonora, como lo exige Avraamov (1923), se logra mediante el uso de tecnologías avanzadas como el OCR y Live Coding, que permiten un flujo continuo entre texto y sonido, y facilitan la interacción en tiempo real.

En términos de *performance*, la obra se apoya en una producción hábil y técnica, donde los artistas no solo manipulan los textos, sino también las herramientas tecnológicas para generar el sonido. La instalación de artefactos sonoros, la proyección de imágenes en tiempo real y el procesamiento de los textos reflejan un dominio técnico que requiere precisión en cada presentación.

La *performatividad* se manifiesta en la acción presente e irrepetible del proceso de sonificación en tiempo real. Cada presentación es única, ya que el proceso mecánico y apurado de cortar, copiar y escanear textos crea un acto efímero, donde la transformación de las palabras en sonidos es tanto un acontecimiento sensorial como conceptual. Este acto de performatividad activa la memoria del texto, recontextualizándolo a través del sonido.

4. Reflexiones

Este proyecto de investigación-creación explora los vínculos entre el concepto de experimentación sonora y las nociones de experimental en ciencia y creación artística. Ambas disciplinas comparten procesos de prueba, descubrimiento y exploración, aunque con fines distintos. En la ciencia, un experimento es un método sistemático para probar hipótesis, generar datos y obtener conclusiones verificables. El proceso implica rigor metodológico y control sobre las variables para llegar a resultados que confirmen o refuten una teoría.

32

En la creación artística, la experimentación es un proceso más abierto e intuitivo, donde las pruebas y los resultados no buscan necesariamente una verdad objetiva, sino la exploración de nuevas formas, significados y experiencias sensoriales. En la experimentación sonora, se pueden explorar las posibilidades acústicas a través de métodos no convencionales, permitiendo que el sonido se transforme en un fenómeno artístico, más allá de la música formal.

En ambos casos el rigor, científico o artístico, es necesario para poder emular y repetir resultados producto de la experimentación, ya que esta, sin el registro adecuado resulta inútil por la imposibilidad de aplicar los resultados obtenidos. Por eso la escritura de este artículo académico resulta ser un ejercicio necesario para, como una apertura más del proceso, nos permite reflexionar, contextualizar y establecer de manera formal y casi científica los resultados obtenidos a partir de un proceso de creación artística.

Las reflexiones finales de esta investigación se centran en la capacidad de la obra para resignificar el contenido de los textos mediante su traducción al sonido. A través del diseño de instrumentos específicos y la integración de tecnologías avanzadas, se establece una relación fluida entre escritura y sonido que trasciende su significado original. La obra, en su carácter performativo, genera un espacio para la experimentación donde los textos son transformados en un fenómeno autónomo. Asimismo, la distinción entre *performance* y *performatividad* permite analizar el proceso técnico y la acción en tiempo real como parte integral de la experiencia artística.

La obra, al permitir la sonificación de cualquier tipo de texto, abre una reflexión sobre el contenido de estos. Los textos, independientemente de su origen o propósito inicial, se resignifican al ser convertidos en un paisaje sonoro que desafía la comprensión tradicional del lenguaje. Esto subraya que la intención de los artistas no es meramente representar los textos de manera sonora, sino transformarlos, otorgándoles nuevas capas de significado a través del diseño de los instrumentos.

Dentro de las proyecciones en torno a esto, se abre la posibilidad de configurar un “alfabeto sonoro”, diseñando específicamente cada sonido emitido y dándole una correlación con su fuente gráfica. De esta manera podrían existir familias de sonidos que respondan a formas similares o emparentadas integrando los procesos de creación tipográfica y emulando sus metodologías. Así, letras similares como la *n*, *m*, *b*, *ñ* podrían tener sonidos similares a partir de que su construcción sucede desde un mismo módulo, al igual que las letras *d*, *p*, *q*, *b*. Mismo caso podría suceder con los sonidos asignados a mayúsculas o minúsculas que si bien son las mismas letras tienen énfasis distintos. Este ejercicio podría acercarse más a la construcción de un lenguaje codificado, alejándose de la sonificación explorada hasta ahora y podría ayudar determinar los límites de la obra en cuanto a su poesía, abstracción o capacidad comunicativa o funcional.

La capacidad de los dispositivos para traducir texto a sonido no se limita a una representación de los signos lingüísticos, sino que genera una experiencia sensorial que resalta la interacción entre lo visual, lo auditivo y lo conceptual. La obra enfatiza que, al ser sonificado, el texto se convierte en un fenómeno autónomo, libre de su significado original y abierto a nuevas interpretaciones.

En línea con las demandas de Avraamov sobre la creación de instrumentos con un *"colosal significado social"*, el diseño de los artefactos en esta obra permite una resignificación que va más allá de la mera función tecnológica. El uso de OCR y MIDI no solo facilita la traducción de texto a sonido, sino que actúa como un puente entre la tecnología y la creación artística, donde el contenido del texto se disocia de su contexto original y se transforma en una composición sonora única.

33

El proceso de sonificación, más que un simple método de traducción se convierte en una exploración creativa donde los textos se reconfiguran en un nuevo formato sensorial. La obra no solo busca representar datos o textos específicos, sino ofrecer un espacio para la experimentación sonora, donde cualquier texto puede ser recontextualizado a través del sonido.

La distinción entre *performance* y *performatividad* en la obra resulta clave para entender la dinámica entre los procesos técnicos y la acción en tiempo real. Las habilidades técnicas de los intérpretes se complementan con la inmediatez de la acción performativa, donde la sonificación se convierte en un acto presente que involucra tanto a los artistas como al público. La obra invita a una reflexión no solo sobre el contenido de los textos, sino sobre el acto mismo de su transformación en sonido, constituyendo un acto de lectura y escucha colectiva pero fuera de los códigos constituidos.

La obra, en cada una de sus presentaciones, genera una experiencia irrepetible que destaca la performatividad del momento, donde el texto se transforma en sonido ante los ojos del público, creando un acontecimiento sonoro que trasciende la simple lectura o audición. Este acto, resulta poético al operar como una evocación sonora de un texto, todo esto sucediendo en tiempo real. No es una composición que se realiza “inspirada en”, sino que es la representación sonora de textos, letras, palabras y códigos mediante nuevos instrumentos, es la transformación del lenguaje escrito que conocemos en una partitura, en un principio irreconocible y luego, en un paisaje sonoro amable y evocativo, una experiencia colectiva de deconstrucción y aprendizaje de nuestro propio lenguaje.

5. Agradecimientos

En la realización de este proyecto ha sido fundamental el apoyo, gestión y aporte de: la Dirección de Creación Artística (DiCREA) de la Vicerrectoría de Investigación y Desarrollo de la Universidad de Chile, a través de su director Fernando Gaspar y todo el equipo del Foro de las Artes 2023; el Director de Investigación y Creación (DIC) de la Facultad de Arquitectura y Urbanismo de la Universidad de Chile, Cristian Gómez-Moya quien orientó la realización de esta obra; Stephanie Orellana, encargada de Diseño y Producción en el Laboratorio Transversal DIC que estuvo a cargo del registro audiovisual y fotográfico de la performance; la Facultad de Arquitectura, Animación, Diseño y Construcción de la Universidad de las Américas (Chile) a través de su Secretaría Académica, Marisol Frugone y el equipo organizador del Festival Soundtiago 2024; y también el artista y poeta Martín Gubbins. A todos ellos y ellas, infinitas gracias por su apoyo, confianza y guía en la realización de este proyecto de investigación-creación.

6. Referencias bibliográficas

- Bal, M. (2016). *Lexicón para el Análisis Cultural* (1st ed.). Akal.
- Cage, J. (2005). *El futuro de la música. Silencio*. Árdora.
- Casati, R., y Dokic, J. (1994). *La philosophie du son*.
- Celedón Bórquez, G. (2023). *Filosofía y experimentación sonora* (1st ed.). Ediciones Metales Pesados.
- Fink, T., y Akdag Salah, A. A. (2023). *Extending the Visual Arts Experience: Sonifying Paintings with AI* (pp. 100–116). https://doi.org/10.1007/978-3-031-29956-8_7
- González-Castañeda, E. F., Torres-García, A. A., Reyes-García, C. A., y Villaseñor-Pineda, L. (2017). *Sonification and textification: Proposing methods for classifying unspoken words from EEG signals*. *Biomedical Signal Processing and Control*, 37, 82–91. <https://doi.org/10.1016/j.bspc.2016.10.012>
- Hyde, J. (2020). The new analogue. Media archaeology as creative practice in 21st-century audiovisual art. In A. Knight-Hill (Ed.), *Sound and Image: Aesthetics and Practices* (1st ed., pp. 188–205). Routledge.
- Kane, J. (2012). *Manual de Tipografía* (2nd ed.). Gustavo Gili.
- Lupton, E. (2015) *Tipografía en Pantalla: Una Guía Para Diseñadores, Editores, Tipógrafos, Blogueros y Estudiantes*. (1st ed.). Gustavo Gili.
- Mata, V. (2018) *Plague, Copie, Manipule, Robe, Reescriba este Libro*. (2023). Kikuyo Editorial.
- Schafer, R. M. (2013). *El paisaje sonoro y la afinación del mundo* (V. G. Cazorla, Trans.; 1st ed.). Intermedio.
- Sloterdijk, P. (2020). *El Imperativo Estético*. Akal.
- Tay, N. W., Liu, F., Wang, C., Zhang, H., Zhang, P., & Chen, Y. Z. (2021). *Protein music of enhanced musicality by music style guided exploration of diverse amino acid properties*. *Heliyon*, 7(9), e07933. <https://doi.org/10.1016/j.heliyon.2021.e07933>
- Temple, M. D. (2020). *Real-time audio and visual display of the Coronavirus genome*. *BMC Bioinformatics*, 21(1), 431. <https://doi.org/10.1186/s12859-020-03760-7>
- Temple, M. D. (2024, June 24). *DNA SONIFICATION USING 8-CHANNEL AUDIO FOR DATA ANALYSES AND MUSIC COMPOSITION*. The 29th International Conference on Auditory Display (ICAD 2024).
- Zielinski, S. (2019). *Variations on Media Thinking*. University of Minnesota Press.