

El sonido de los espacios invisibles: un ensayo sobre la inteligencia (artificial) de la música

Dusan Kotoras Straub¹

Introducción

La generación de música mediante técnicas de inteligencia artificial (IA) ha demostrado un potencial significativo para descomponer, modelar y reconfigurar las características del sonido en escalas y resoluciones, hasta hace poco, inaccesibles. Modelos como *MusicLM* (Agostinelli et al., 2023), *Riffusion* (Forsgren, Martiros, 2022), o *MusicVAE* (Roberts et al., 2019) convierten entradas auditivas, textuales o simbólicas en espacios vectoriales de alta dimensionalidad, reorganizando las estructuras musicales sin necesidad de recurrir a los parámetros de duración, altura o timbre. Estas arquitecturas no replican una gramática musical convencional, sino que condensan regularidades estadísticas derivadas del entrenamiento sobre grandes corpus de datos.

En el marco de las vanguardias de fines del siglo XX, el compositor Iannis Xenakis (1922–2001) definió la música como una estructura capaz de materializar las múltiples dimensiones de la inteligencia humana, en un sentido que anticipa buena parte de los desarrollos actuales en IA. El conjunto de su trabajo buscaba impulsar la síntesis de ciencia y arte a través de experimentos

¹ Sociólogo de la Universidad Diego Portales y doctorando en Estudios Mediales en la Universidad Alberto Hurtado, se desempeña como asistente de investigación del Núcleo Milenio Futures of Artificial Intelligence Research (FAIR). dcotoras@gmail.com

formales. En particular, la serie de instalaciones que mezclaban sonido, proyecciones visuales y arquitectura bajo el concepto de politopo, pueden ofrecer una perspectiva diferente sobre la dimensión material de las múltiples aristas que actualmente se sintetizan en las técnicas de IA.

Este ensayo busca explorar las técnicas empleadas para formalizar la inteligencia en la generación de música, mediante aprendizaje profundo, en comparación al conjunto de arreglos materiales que se despliegan en las instalaciones de Xenakis. La hipótesis que guía esta comparación es que, mientras los sistemas actuales de IA tienden a disociar los resultados obtenidos de sus soportes materiales, la obra de Xenakis ofrece un modelo alternativo que visibiliza dimensiones latentes en espacios concretos. Desde una perspectiva basada en el análisis de los medios requeridos para su funcionamiento, se contrastan las arquitecturas de los modelos de IA generativa, con aspectos equivalentes en el marco de la exhibición de los politopos. En base a su comparación, se espera promover una discusión más amplia con respecto a las condiciones que hacen posible la inteligencia de la música en la implementación de IA.

¿Cómo piensa la música?

*... la cualificación de belleza o fealdad carece de sentido respecto al sonido,
así como respecto a la música que de él se deriva;
la cantidad de inteligencia que los sonidos portan consigo
debe ser la verdadera medida de validez de una música en particular.*

Iannis Xenakis (1992)

Las técnicas de reconocimiento de patrones o espacios vectoriales multidimensionales que propicia la IA, han sido recientemente reconocidas como técnicas transversales que articulan las condiciones de producción, circulación y percepción del patrimonio cultural en el presente (Krämer, 2024; Pasquinelli, Joler, 2021). Estos procedimientos pueden ser comprendidos como técnicas culturales que modulan la inteligibilidad del mundo y

configuran modos de operación sobre lo real (Siegert, 2015). En este marco, el concepto de técnicas culturales permite superar una comprensión instrumental para relevar en su lugar los «procesos operativos que permiten trabajar con cosas y símbolos», más allá de los marcos interpretativos tradicionales centrados en el sonido o la representación. En contra de la clásica escisión moderna entre sujeto y objeto, este enfoque «promueve los logros de la inteligencia a través de los sentidos y externalización operacional de los procesos de pensamiento» (Krämer, Bredekamp, 2013: 27), permitiendo reinscribir el concepto de inteligencia en el conjunto de técnicas desplegadas culturalmente.

En el ámbito de los estudios musicales, esta perspectiva ha servido para cuestionar el sesgo representacional que ha dominado la comprensión occidental de la música. Sybille Krämer (2020), ha identificado un «dogma fonográfico» en la tendencia a centrar la atención, exclusivamente, en la fijación tonal o en el producto acústico final, descuidando las operaciones de visualización gráfica y formalización escrita que lo constituyen. Este sesgo ha tenido como consecuencia una desatención de las dimensiones técnicas involucradas en la composición, que tienden a quedar relegadas al estatuto de sistemas de registro con funciones subordinadas de almacenamiento y transmisión de sonido (Kittler, 1999). Investigaciones recientes en torno al rol de la notación musical han demostrado su potencial generativo (Fuhrmann, 2011), que actúa directamente sobre los procesos de invención musical, de manera análoga a los esquemas geométricos en las matemáticas o los diagramas en las ciencias naturales: como espacios operativos que permiten experimentar en función de estructuras abstractas (Celestini et al., 2019; Urbanek, 2024).

La emergencia de modelos de IA basados en aprendizaje profundo representa, en este escenario, un importante desafío para el análisis de las técnicas culturales relacionadas con el ejercicio de la música. En primer lugar, los parámetros empleados por estos modelos para la síntesis de datos, prescinden de las convenciones de la notación musical, al tiempo que introducen nuevas dimensiones

operativas (Kollinger, 2018). En lugar de trabajar con categorías discretas organizadas simbólicamente, pueden trabajar sobre distribuciones en espacios latentes de alta dimensionalidad (Pasquinelli, Kornweitz, 2023). Esto cuestiona los fundamentos mismos de la identificación de un objeto musical. Según una formulación clásica, la notación cumplía una función epistémica al fragmentar lo continuo en gradaciones finitas, permitiendo distinguir cada signo y reconocer una obra como tal (Goodman, 1976). Cuando se abandona esta base simbólica, se vuelve necesario preguntar qué condiciones permiten aún hablar de generación musical, y cómo es posible identificar eventos sonoros cuya estructura ya no se rige por reglas compositivas.

Esta compresión a gran escala obliga, igualmente, a reconsiderar la espacialidad en la que estas técnicas operan. Sybille Krämer (2023) ha sostenido que la inteligencia moderna se constituye al «aplanar artificialmente» la tridimensionalidad de la experiencia cotidiana, proyectándola sobre superficies bidimensionales aptas para la formalización visual. En estricto sentido, la notación musical occidental responde a esta lógica, al inscribir relaciones de frecuencia y duración sobre un plano gráfico. Esta operación permite estabilizar lo efímero del sonido, visualizar su organización y construir estructuras compositivas a partir de esa forma espacial plana (Krämer, 2017). En cambio, los modelos actuales parecen operar en espacios que exceden las posibilidades de visualización bidimensional. Su lógica no prolonga, necesariamente, la inscripción gráfica tradicional, sino que trabaja mediante relaciones distribuidas no directamente accesibles a la percepción (Zylinska, 2024). La pregunta que se abre a partir de estos condicionamientos es si efectivamente tales modelos reproducen, desplazan o reconfiguran las condiciones espaciales que históricamente hicieron posible pensar las técnicas culturales relacionadas a la música.

Es necesario, además, considerar la dimensión operativa de los objetos experimentales que se despliegan en el ejercicio de la música, históricamente caracterizados por su capacidad de admitir variación, iteración y reversibilidad dentro de marcos formales bien

definidos y desde una materialidad técnica determinada (Kittler, 2017; Krämer, 2006). Estas operaciones han permitido no solo la exploración de estructuras musicales, sino también, el control consciente sobre los procedimientos de transformación. En contraste, los modelos de aprendizaje profundo parecen tensionar o desplazar esas condiciones al operar en representaciones latentes que no se presentan bajo una forma manifiesta. La manipulación que realizan se sustenta en cálculos estadísticos distribuidos, cuya trazabilidad resulta limitada, incluso, para sus propios desarrolladores. Esta situación ha contribuido a consolidar una imagen ilusoria de autonomía que tiende a oscurecer las condiciones materiales y operativas que hacen posible su funcionamiento (Pasquinelli et al., 2024). Frente a ello, cabe preguntarse cómo se expresa la inteligencia en operaciones específicas efectuadas sobre estos modelos, y en qué medida la composición musical queda subordinada a procedimientos automáticos que restringen la intervención directa del compositor.

En lo que sigue, se analizan algunos de los principales modelos de IA para la generación de música desde la perspectiva de las técnicas culturales aquí esbozada (Krämer, Bredekamp, 2013). Se pondrá especial énfasis en las dimensiones de espacialidad, operatividad y simbolización que estructuran, tanto el funcionamiento interno de estos sistemas, como su proyección en las prácticas compositivas. El objetivo es evaluar en qué medida estas tecnologías logran expandir las fronteras tradicionales de nuestra comprensión de la música, y al mismo tiempo, qué límites introducen como parte de su concepción técnica, especialmente en relación con los desplazamientos estéticos, formales y materiales propuestos por las vanguardias musicales de fines del siglo XX.

El sonoscopio de la IA

Los modelos de generación musical basados en IA se estructuran mediante codificadores y decodificadores que transforman la información sonora en representaciones intermedias de alta dimensionalidad. Estas representaciones no son accesibles a la interpretación humana directa, pero permiten la interpolación

y recombinación eficiente de materiales. En lugar de manipular eventos musicales discretos, los sistemas como MusicLM, Riffusion, o MusicVAE trabajan sobre vectores que condensan información estadística distribuida. Proponemos aquí el término *sonoscopio* para designar un modo de observación que desplaza el foco desde las superficies audibles hacia las arquitecturas que procesan, codifican y transforman recursos heterogéneos. En su raíz etimológica (*sonos*, sonido; *skopein*, observar, examinar), constituye un neologismo inspirado en el concepto de *nooscopio*, revisitado por Pasquinelli y Joler (2021), como una herramienta que permite cartografiar esta técnica cultural en donde el pensamiento es modelado.

Un primer aspecto que considerar es la forma en que estos modelos sintetizan el sonido. Modelos como *MusicLM* (Agostinelli et al., 2023) permiten observar con claridad cómo la generación musical no opera sobre parámetros musicales explícitos (altura, ritmo, duración), sino a partir de representaciones latentes jerárquicas extraídas directamente del audio. Su arquitectura se basa en una cascada de codificadores, en la que, primero, se transforma un archivo de audio en una secuencia de tokens discretos mediante un modelo autosupervisado de codificación acústica. Estos tokens no codifican eventos musicales individuales, sino que resumen información espectral y temporal de forma compresiva. Luego, a partir de descripciones textuales, se emplea un modelo generativo autoregresivo, entrenado para predecir secuencias plausibles a partir de descripciones textuales, estableciendo así una correspondencia estadística entre lenguaje natural y contenido acústico. Finalmente, un decodificador reconstruye el audio a partir de estos tokens. El procedimiento evita toda representación simbólica intermedia: no hay partituras, notas, ni escalas codificadas, solo distribuciones cuya geometría se aprende del corpus de datos (Figura 1).

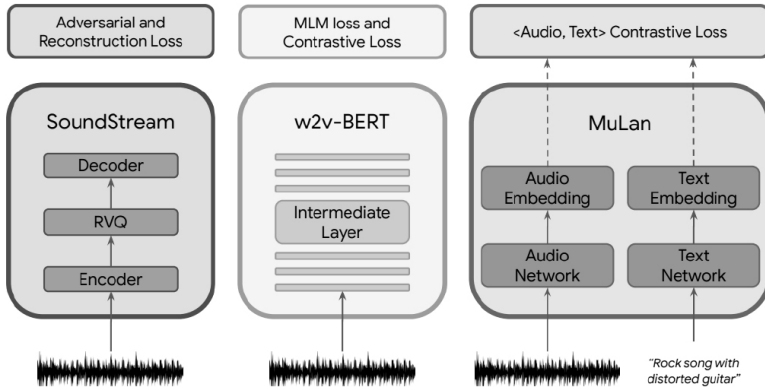


Figura 1: Modelos pre-entrenados con representaciones independientes (MusicLM).

Por otro lado, el modelo *Riffusion* (Forsgren, Martiros, 2022) permite observar cómo ciertas arquitecturas reintroducen una espacialidad visual para intervenir lo sonoro desde una lógica no necesariamente musical. Su funcionamiento se basa en el uso de modelos de difusión originalmente diseñados para generación de imágenes, adaptados para producir espectrogramas –representaciones bidimensionales de frecuencia en el tiempo–. En lugar de trabajar sobre archivos MIDI, *Riffusion* genera imágenes espectrales a partir de ruido mediante un proceso iterativo de refinamiento, condicionado por un texto o características de estilo. El espectrograma es luego convertido en audio mediante una técnica de síntesis de Fourier de onda corta. A diferencia de los modelos que operan en espacios latentes no interpretables, *Riffusion* permite visualizar la estructura espectral del material generado, lo que habilita cierta trazabilidad de las operaciones. Sin embargo, este tipo de espacialidad remite a densidades gráficas donde los eventos están inscritos en una matriz manipulada mediante operaciones de interpolación visual.

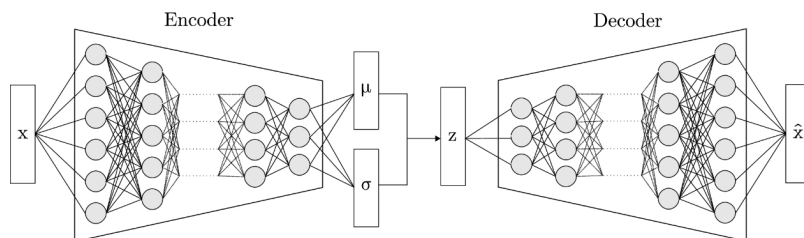


Figura 2: Variational AutoEncoder (VAE).

Finalmente, *MusicVAE* ofrece un ejemplo sobre cómo la interpolación en espacios latentes puede emplearse como operación compositiva. A diferencia de los modelos anteriores, *MusicVAE* trabaja sobre datos simbólicos (archivos MIDI) que son codificados en secuencias de tokens, representando parámetros musicales como pitch, duración, velocidad, silencios y desplazamientos temporales. La arquitectura del modelo se basa en un autoencoder variacional (Figura 2), compuesto por un codificador bidireccional que transforma una secuencia en una distribución gaussiana que comprende una media (μ) y desviación estándar (σ) dentro de un espacio latente de 512 dimensiones. Este vector (z) representa una posición geométrica que codifica propiedades de la secuencia original (x). Posteriormente, un decodificador genera una nueva secuencia a partir de ese vector, optimizando, simultáneamente, la fidelidad de reconstrucción y la regularidad del espacio para garantizar que cualquier punto intermedio sea musicalmente plausible ($\hat{x}$). De tal manera, este espacio latente continuo permite realizar interpolaciones entre secuencias –como entre una melodía barroca y un solo de jazz– mediante la navegación lineal entre vectores, como lo muestra la Figura 3. Dichas operaciones no dependen de reglas musicales explícitas, sino de la geometría interna aprendida por el modelo a partir del corpus de datos, donde los desplazamientos vectoriales permiten alterar rasgos específicos de una secuencia (como aumentar su densidad rítmica o su cromatismo), y construir variaciones estilísticas sin codificación simbólica directa.

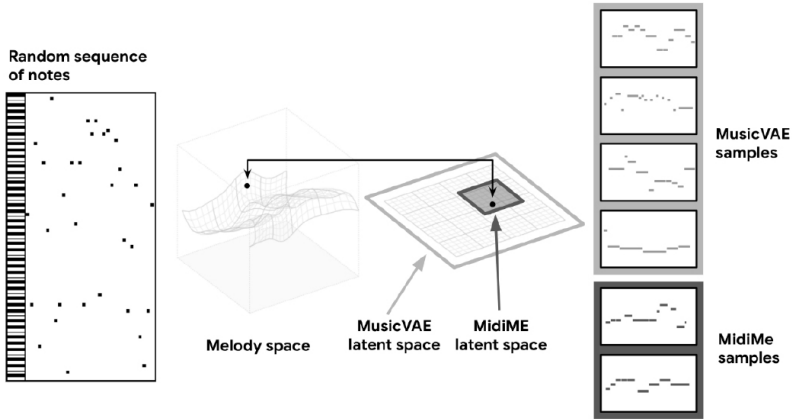


Figura 3: Espacio Latente (MusicVAE).

Estos tres casos ilustran desde distintos ángulos las dimensiones fundamentales que se quisieran destacar en el *sonoscopio*: el trabajo con representaciones latentes jerárquicas no interpretables (MusicLM), la proyección espectral en espacios visuales bidimensionales (Riffusion), y la interpolación operativa en espacios vectoriales continuos (MusicVAE). En conjunto, permiten constatar que lo que antes se pensaba como «composición», ahora puede entenderse como una operación de posicionamiento, interpolación o desplazamiento dentro de una topología técnica donde la música se configura, no a partir de reglas, sino de geometrías aprendidas. En vez de articularse en una superficie de notación estable, la música emerge aquí desde espacios de alta dimensionalidad en los que las regularidades son estadísticas y las transformaciones, geométricas.

El tipo de inteligencia que promueven estos modelos puede definirse como una inteligencia topológica-estadística. Esta lógica reemplaza la composición simbólica por una navegación continua entre configuraciones probabilísticas. Sin embargo, resulta necesario desmitificar la idea de una ruptura total con el pensamiento musical precedente. Más que una discontinuidad, estas técnicas reactualizan aspiraciones de larga data: la espacialización del sonido, la abstracción

formal y la automatización del gesto compositivo. En lo que sigue, se pondrá en cuestión la novedad de estas formulaciones, proponiendo un antecedente clave que permita reconsiderar esta técnica cultural desde la perspectiva de las vanguardias electroacústicas de fines del siglo XX.

Las utopías concretas de Xenakis

En una reciente intervención en la Bienal de Música de Venecia, Matteo Pasquinelli y Arif Kornweitz (2023), extendieron una provocadora tesis con respecto a que ciertas técnicas de IA actual relacionadas con espacios multidimensionales, podrían encontrar un antecedente en las vanguardia del siglo XX. En particular, destacaron la serie de instalaciones inmersivas desarrolladas por Iannis Xenakis entre las décadas de 1967 y 1978, agrupadas bajo el concepto de politopo –del griego poli (muchos) y topos (espacio)–, en las que se integraban música electroacústica, proyecciones visuales y sistemas lumínicos.

El uso del término politopo admite múltiples interpretaciones. En primer lugar, el concepto de topos puede entenderse como «punto en el espacio», ya que estos espectáculos son apreciables desde diversas perspectivas y posiciones, variando según el movimiento del espectador. En segundo lugar, cada politopo se configura como un espectáculo totalizador, un «espacio sensorial» en el que todos los sentidos se activan en una experiencia global e inmersiva. Asimismo, la noción de topos remite a la convergencia interdisciplinaria de las disciplinas artísticas convocadas por Xenakis, incluso, extendiendo esta integración a los ámbitos científico y tecnológico, evidenciando el carácter utópico de síntesis entre arte y ciencia que atraviesa toda su obra (Xenakis, 1985).

Un caso especialmente revelador de la concepción espacial en los politopos de Xenakis es el Politopo de Cluny, desarrollado entre 1972 y 1974 en las antiguas termas galo-romanas de París. La instalación articulaba una compleja arquitectura de medios en la que el sonido era distribuido mediante una red de altavoces monta-

dos sobre una estructura de andamiaje que reproducía en forma de retícula ortogonal, la curvatura interna de las bóvedas góticas del edificio. Esta disposición permitía transformar el espacio arquitectónico en una matriz acústica activa, donde cada pista sonora era trasladada mediante fundidos entre distintos puntos de emisión a velocidades constantes o variables, generando trayectorias dinámicas y envolventes. De tal manera, la experiencia auditiva operaba sobre desplazamientos físicos del sonido, con efectos cinéticos que transformaban la percepción en tiempo real. Así entendida, la espacialidad no era el soporte pasivo de la obra, sino una superficie que moldeaba los parámetros compositivos, estructurados a través de relaciones entre materia sonora, geometría arquitectónica y movimiento corporal del oyente.

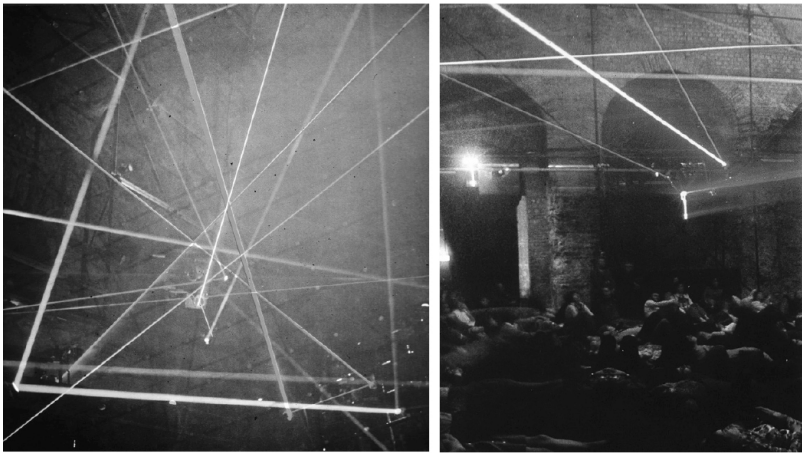


Figura 4: Espacios geométricos formados por haces de luces (Politopo de Cluny).

La dimensión visual se integraba en la instalación mediante un sistema de rayos láser, flashes electrónicos y espejos, cuyo control estaba regido por funciones matemáticas que traducían transiciones de números complejos en distribuciones de probabilidad, determinando parámetros como intensidad, color, duración y trayectoria de los haces (Figura 4). Esta configuración permitía proyectar formas elementales –puntos, líneas– que interactuaban con los espejos fijos, para producir ilusiones geométricas y alterar la percepción del espacio a través de reflejos, multiplicaciones o desplazamientos lumínicos. En palabras de Xenakis, se trataba de materializar una «música visual, abstracta», donde el comportamiento físico de la luz funcionara como un equivalente formal del sonido. En conjunto, la instalación articulaba una espacialidad multimedial construida mediante la coordinación precisa entre fuentes visuales y sonoras, proponiendo un modelo técnico-operativo de inteligencia que contrasta con los actuales enfoques de IA musical.

El *Diatope*, instalado en 1978 frente al Centro Pompidou, representa uno de los dispositivos compositivos más avanzados dentro de la serie de politopos. En la obra *La Légende d'Eer*, compuesta especialmente para esta instalación, la espacialización sonora fue tratada como principio constructivo fundamental de la infraestructura de la instalación. En lugar de una secuencia de eventos fijos en el tiempo, Xenakis diseñó trayectorias móviles, registradas en una partitura técnica que esquematiza los desplazamientos de cada pista a través de una red de altavoces distribuidos estratégicamente (Figura 5). Según el análisis de Kiourtsoglou (2017), esta notación –similar a una tablatura espacial– no registra notas individuales, sino recorridos físicos del sonido en el espacio, permitiendo controlar velocidad, dirección y posición. Esta organización implica una reconceptualización operativa del acto compositivo: el sonido ya no es producido desde un punto fijo, sino gestionado como entidad móvil en un entorno tridimensional.

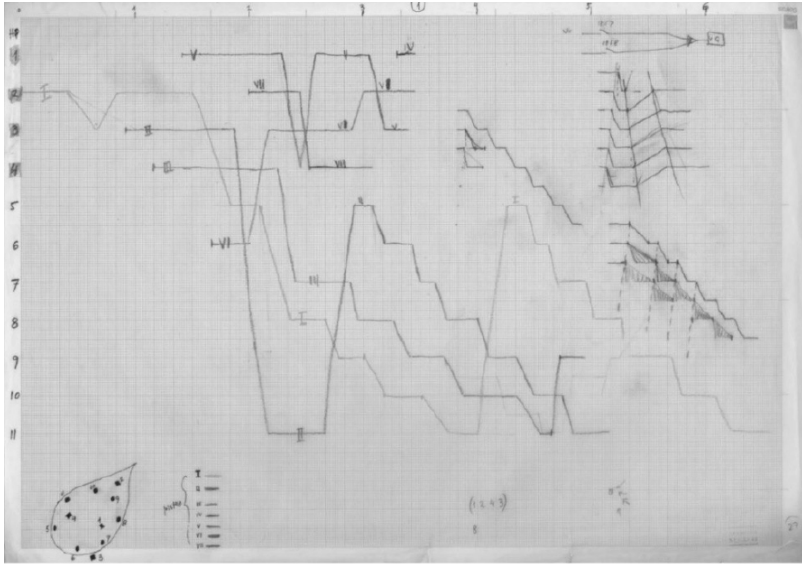


Figura 5: «Partitura» de *La Légende d'Eer*.

En esta composición se emplearon procedimientos estocásticos desarrollados por Xenakis como alternativa a los enfoques armónicos clásicos. La técnica se basa en el modelado directo de la curva presión-tiempo mediante trayectorias aleatorias –paseos brownianos– reguladas por funciones logísticas y distribuciones probabilísticas, sin recurrir a procesos de descomposición espectral. Como él mismo plantea: «para invertir el principio de análisis armónico que supone una progresión desde la simplicidad hasta mayor complejidad [...] se comienza en su lugar con paseos aleatorios (*random walks*) a partir de funciones estocásticas» (1996: 153). Esta inversión prioriza la generación de alta granularidad desde el comportamiento de sistemas físicos no deterministas. El Diatope extiende esta lógica al plano arquitectónico. En sus bocetos (Figura 6), Xenakis señala que la estructura está pensada para permanecer permeable a las ondas que circulan en el entorno espacial, incluyendo propagación acústica, emisión lumínica o variaciones térmicas.

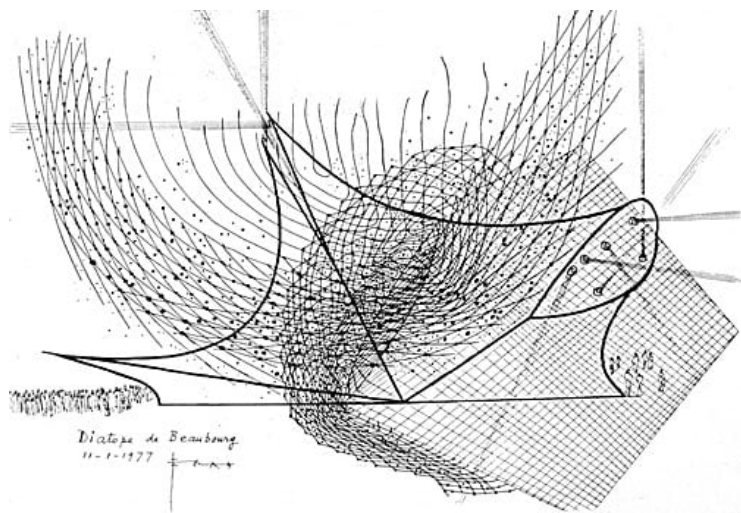


Figura 6: Boceto de estructura permeable en El Diatopo.

El Polítopo de Micenas (*Mycènes Alpha*, 1978), presentado en el sitio arqueológico de Micenas, fue la primera obra compuesta íntegramente mediante el sistema UPIC (*Unité Polyagogique Informatique du CEMAMu*), una interfaz gráfica digital desarrollada por Xenakis que permite dibujar directamente sobre una superficie y convertir esas formas en estructuras sonoras (Figura 7). Esta característica imprime a la obra una lógica de continuidad morfológica basada en transiciones visuales que se traducen en interpolaciones sonoras. En lugar de operar con unidades discretas o eventos musicales puntuales, la composición se organiza como un espacio de transformaciones progresivas donde cada línea o curva dibujada en el UPIC, define una evolución paramétrica continua –en frecuencia, densidad, intensidad o timbre– dentro de un espacio de representación bidimensional. El resultado es una textura dinámica en la que las transformaciones no responden a un modelo fijo, sino a trayectorias gráficas que definen relaciones topológicas entre estados sonoros, moduladas en tiempo real durante la ejecución, revelando una concepción operativa de la composición que emerge directamente del tratamiento gráfico del sonido.

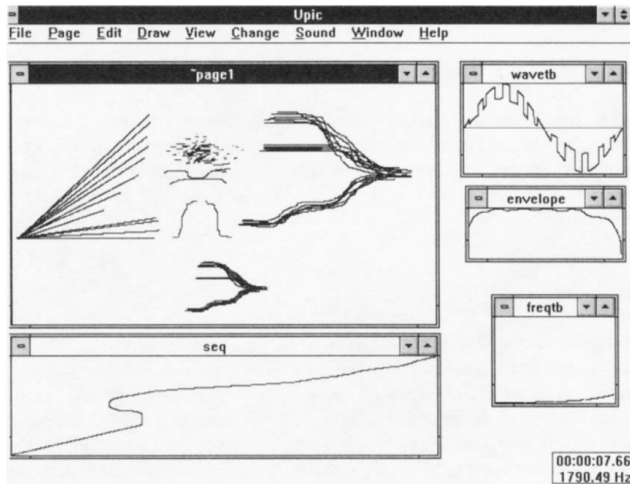


Figura 7: Diagrama para composición diseñado en UPIC (*Mycènes Alpha*, 1978).

En marzo de 1974, Xenakis propuso el Politopo Mundial, un proyecto inacabado que aspiraba a conectar poblaciones dispersas mediante experiencias sensoriales compartidas a escala global. La propuesta consistía en la instalación de politopos –dispositivos de luz, sonido y control computacional– en múltiples ciudades del mundo, transformando espacios públicos en nodos receptores de señales transmitidas por redes técnicas de alta complejidad. Estas transmisiones, gestionadas por computadoras o cintas digitales programadas, se articularían mediante haces de rayos láser reflejados sobre largas distancias y sistemas de comunicación vía satélite geoestacionarios o móviles. La idea era que las señales moduladas desde los centros de emisión alteraran en tiempo real los parámetros de los espectáculos en los centros de recepción, conformando una red sincrónica de composición y percepción distribuida. Xenakis concebía este sistema como una forma de contacto directo entre humanos a través de estructuras energéticas compartidas, sin recurrir a sistemas simbólicos intermediarios como el lenguaje o la representación. Esta visión proyecta con claridad la cosmología que subyace a toda su obra: una reorganización de las dimensiones del espacio como medio de sincronización a escala planetaria.

Otra IA es posible

La comparación entre los modelos contemporáneos de IA y los politopos de Iannis Xenakis permite visibilizar las condiciones materiales, formales y perceptivas que configuran la práctica compositiva en la forma histórica de las técnicas culturales descritas. La divergencia entre ambos enfoques no radica exclusivamente en los medios empleados –ya sea redes neuronales profundas o sistemas electroacústicos–, sino en el modelo de inteligencia que articula los eventos musicales como técnicas concretas. Organizar su discusión a partir de tres ejes –síntesis paramétrica, espacialidad y operatividad– permite observar cómo cada modelo se relaciona con el problema de la dimensionalidad del espacio y cómo modula la inteligibilidad de la experiencia sonora.

Dimensión	Sonoscopio	Politopo
Espacialidad	Espacio latente abstracto, multidimensional, no perceptible directamente. Se organiza por similitud estadística entre vectores.	Espacio físico intervenido mediante arquitectura, altavoces, proyecciones. Manipulación directa de la percepción espacial a través del diseño estructural.
Síntesis	No se parte de parámetros musicales tradicionales, sino de codificaciones comprimidas. La interpolación ocurre entre vectores en espacios latentes.	Empleo de síntesis estocástica, interpolación visual-sonora a través de UPIC. Relación directa entre gráfica, parámetro y resultado sonoro.
Operatividad	Basada en redes neuronales profundas y entrenamiento estadístico. Procesos de generación mediados por sampling estocástico y aprendizaje supervisado/no supervisado.	Control explícito mediante esquemas gráficos, diagramas matemáticos y dispositivos electromecánicos programables. Integración entre concepto y ejecución técnica.

En términos de síntesis paramétrica, los modelos de IA musical actuales construyen espacios de representación a partir

de datos de entrenamiento, generando interpolaciones en regiones latentes que permiten producir transiciones estilísticas o tímbricas. En contraste, las técnicas desarrolladas por Xenakis conservan una inscripción material visible: el gesto gráfico sobre una superficie es inmediatamente traducido en sonido, y mantiene su condición operativa durante todo el proceso compositivo. No hay aquí separación entre representación y ejecución, sino una continuidad entre trazo, parámetro y percepción. La interpolación se realiza sobre superficies perceptualmente accesibles y materialmente modulables, lo que implica una concepción distinta del lugar del compositor, quien desempeña la función de interfaz en lugar de ser completamente sustituido.

Los espacios vectoriales abstractos encuentran en los politopos un equivalente funcional radicalmente material. No se trata simplemente del lugar donde la obra ocurre, sino de una dimensión estructurante que modula la parametrización de los medios involucrados. Espacios como el de la capilla de Cluny o la estructura geodésica del *Diatope* no funcionan como escenarios pasivos, sino como partituras tridimensionales desde donde se articula la obra en base a posibilidades y restricciones materiales explícitas. En este sentido, los politopos funcionan como puntos clave de una tecnoestética, en el sentido dado por Gilbert Simondon (2017), donde se reorganizan los vínculos entre cultura y técnica en un diagrama de comportamiento colectivo.

La operatividad, por su parte, permite interrogar las condiciones bajo las cuales, estas formas de organización sonora se inscriben en estructuras técnicas efectivas. En los modelos de IA, las operaciones compositivas quedan distribuidas en infraestructuras de cómputo cuya escala y complejidad dificultan su inspección directa. El entrenamiento, la parametrización y la inferencia, se realizan en marcos estadísticos altamente optimizados, pero cuyo funcionamiento es opaco desde la perspectiva del usuario. En los politopos, la operatividad no es un proceso oculto, sino una práctica explícita que se realiza mediante dispositivos visibles, materiales y programables. Cada elemento –ya sea una fuente sonora móvil, un retardo espacial o un algoritmo estocástico– responde a una

lógica de control trazable. Esta diferencia no debe entenderse como una oposición binaria entre transparencia y opacidad, sino como una diferencia en la distribución contextual de los medios técnicos empleados. Aquí puede ser útil recurrir a la noción de policontextualidad formulada por Gotthard Günther (1979), que describe sistemas en los que múltiples lógicas coexisten sin poder ser reducidas unas a otras. Los politopos ejemplifican esta coexistencia: gráficos, arquitecturas, sonidos y movimientos se integran sin colapsar en un único marco representacional. En los sistemas de IA, por el contrario, la tendencia es hacia la unificación de estas dimensiones en una arquitectura que denominamos como sonoscopio, con ventajas en términos de eficiencia, pero una pérdida significativa de intermediación material y sensible.

Desde una lectura estrictamente técnica, los politopos de Xenakis no constituyen una anticipación directa de los modelos actuales de IA musical, sino una exploración alternativa sobre cómo distribuir inteligencia compositiva en estructuras materiales. En lugar de reducir la complejidad a través de abstracciones funcionales, proponen desplegarla como red de relaciones operativas entre medios heterogéneos. La inteligencia, en este contexto, no se define por su capacidad de predicción ni por la optimización de un modelo, sino por su aptitud para construir condiciones de integración multiescalar entre percepción, forma y técnica. Esta diferencia, más que oponer dos paradigmas, señala la necesidad de ampliar el marco analítico con el que se abordan las infraestructuras musicales contemporáneas: ni los politopos pueden leerse como simples obras multimedia, ni los sistemas de IA como sonoscopios invisibles. Ambos casos exigen modelos de análisis que reconozcan las configuraciones materiales, operativas y perceptuales que los constituyen, así como las dimensiones –visibles o no– que hacen posible su funcionamiento.

Conclusión

Los resultados del análisis propuesto en el presente ensayo sugieren que la manera en que se estructura la relación entre parámetros compositivos, soporte técnico y entorno material, influye en el tipo de inteligencia que se modela o se operacionaliza en cada sistema. En los modelos de IA musical, la eficiencia de generación depende de la capacidad de capturar regularidades estadísticas y condensarlas en representaciones de baja interpretabilidad. Esta lógica favorece la escalabilidad y la automatización, pero al mismo tiempo, restringe la intervención directa sobre las condiciones físicas del sonido y su despliegue en el espacio. En contraste, los polítopos de Xenakis demuestran que es posible diseñar sistemas donde las transformaciones acústicas, visuales y espaciales están directamente integradas en la arquitectura de control, lo que permite una trazabilidad más precisa entre intención compositiva, medio técnico y experiencia perceptiva.

Estas diferencias abren líneas de trabajo relevantes tanto para la crítica de las infraestructuras actuales como para el desarrollo de sistemas híbridos. Por un lado, permiten cuestionar la invisibilización de las condiciones materiales en los procesos de síntesis musical por IA, y por otro, ofrecen un modelo de composición en el que los medios no son un soporte neutral, sino parte activa del sistema. Esto implica que futuras investigaciones y desarrollos podrían orientarse no solo a mejorar la calidad de salida de los modelos generativos, sino también, a rediseñar sus interfaces, flujos operativos y entornos de ejecución para restituir la dimensión material como variable compositiva. Lejos de oponer ambos enfoques, el análisis invita a considerar cómo su comparación puede enriquecer la comprensión de las condiciones técnicas que hacen posible, inteligible y modulable la generación musical en contextos computacionales avanzados.

Referencias bibliográficas

- Agostinelli, A., Denk, T. I., Borsos, Z., Engel, J., Verzetti, M., Cailion, A., Huang, Q., Jansen, A., Roberts, A., Tagliasacchi, M., Sharifi, M., Zeghidour, N., & Frank, C. (2023). *MusLM: Generating Music From Text* (Versión 1). arXiv.
- Celestini, F., Nanni, M., Obert, S., & Urbanek, N. (2019). *Zu einer Theorie der musikalischen Schrift*. Brill. https://doi.org/10.30965/9783846763537_002
- Fuhrmann, W. (2011). Notation als Denkform. Zu einer Mediengeschichte der musikalischen Schrift. *Musiken. Festschrift für Christian Kaden*, 114–135.
- Goodman, N. (1976). *Los lenguajes del arte: Aproximación a la teoría de los símbolos* (1. ed). Editorial Seix Barral.
- Günther, G. (1979). Life as Poly-Contextuality. En *Beiträge zur Grundlegung einer operationsfähigen Dialektik* (Vol. 2).
- Kiourtsoglou, E. (2017). An Architect Draws Sound and Light: New Perspectives on Iannis Xenakis's *Diatope* and *La Légende d'Eer* (1978). *Computer Music Journal*, 41(4), 8–31. https://doi.org/10.1162/comj_a_00437
- Kittler, F. A. (1999). *Gramophone, Film, Typewriter*. Stanford University Press.
- Kittler, F. A. (2017). Real Time Analysis, Time Axis Manipulation. *Cultural Politics*, 13(1), 1–18. <https://doi.org/10.1215/17432197-3755144>
- Kollinger. (2018). Wenn aus Zahlen Töne werden... überlegungen zu computergenerierter musik und komposition. En C. Engemann & A. Sudmann (Eds.), *Machine Learning – Medien, Infrastrukturen und Technologien der Künstlichen Intelligenz* (pp. 291–308). transcript Verlag.
- Krämer, S. (2006). The Cultural Techniques of Time Axis Manipulation: On Friedrich Kittler's Conception of Media. *Theory, Culture & Society*, 23(7–8), 93–109. <https://doi.org/10.1177/0263276406069885>

- Krämer, S. (2017). Flattening As Cultural Technique: Epistemic and Aesthetic Functions of Inscribed Surfaces. *Journal of the American Musicological Society*, 70(1), 239–244. <https://doi.org/10.1525/jams.2017.70.1.221>
- Krämer, S. (2020). Schrift, Schriftbildlichkeit, Musik. En C. Ratzinger, N. Urbanek, & S. Zehetmayer (Eds.), *Musik und Schrift: Interdisziplinäre Perspektiven auf musikalische Notationen* (pp. 67–86). Brill, Wilhelm Fink.
- Krämer, S. (2023). Planitud artificial: Reflexiones acerca de una técnica cultural. En R. Cedeño & J. Arango (Eds.), *Proyecto, imagen, algoritmo* (pp. 21–42). Universidad Nacional de Colombia; Universidad de Antioquia.
- Krämer, S. (2024). Die Nicht-Vernunft der Chatbots: Was macht auf Large Language Models beruhende Künstliche Intelligenz philosophisch interessant? En R. Adolphi, S. Alpsancar, S. Hahn, & M. Kettner (Eds.), *Philosophische Digitalisierungsforschung: Verantwortung, Verständigung, Vernunft, Macht* (pp. 297–314). transcript Verlag.
- Krämer, S., & Bredekamp, H. (2013). Culture, Technology, Cultural Techniques – Moving Beyond Text. *Theory, Culture & Society*, 30(6), 20–29. <https://doi.org/10.1177/0263276413496287>
- Pasquinelli, M., Alaimo, C., & Gandini, A. (2024). AI at Work: Automation, Distributed Cognition, and Cultural Embeddedness. *Tecnoscienza – Italian Journal of Science & Technology Studies*, 15(1), 99–131. <https://doi.org/10.6092/ISSN.2038-3460/20010>
- Pasquinelli, M., & Joler, V. (2021). El Nooscopio de manifiesto. *la Fuga*.
- Pasquinelli, M., & Kornweitz, A. (2023). *The Sound of Multidimensional Space: How Avant-Garde Music Foreshadowed Artificial Intelligence*.
- Roberts, A., Engel, J., Raffel, C., Hawthorne, C., & Eck, D. (2019). *A Hierarchical Latent Vector Model for Learning Long-Term Structure in Music* (No. arXiv:1803.05428). arXiv. <https://doi.org/10.48550/arXiv.1803.05428>

- Siegert, B. (2015). *Cultural Techniques: Grids, Filters, Doors, and Other Articulations of the Real*. Fordham University Press.
- Simondon, G. (2017). Reflexiones sobre la tecnoestética (1982). En *Sobre la técnica: 1953-1983* (pp. 365–382). Cactus.
- Urbanek, N. (2024). Writing Electroacoustic Music. Xenakis's Œuvre as a Theoretical Challenge. En R. Friedl, T. Grill, N. Urbanek, & M. Ziegler (Eds.), *Xenakis—Back to the Roots: Philological Approaches to Electroacoustic Music* (pp. 11–38). transcript Verlag.
- Xenakis, I. (1985). *Arts-sciences, alloys: The thesis defense of Iannis Xenakis before Olivier Messiaen, Michel Ragon, Olivier Revault d'Allonnes, Michel Serres, and Bernard Teyssède*. Pendragon Press.
- Xenakis, I. (1992). *Formalized music: Thought and mathematics in composition* (Revised edition). Pendragon Press.
- Xenakis, I. (1996). Determinacy and indeterminacy. *Organised Sound*, 1(3), 143–155. Cambridge Core. <https://doi.org/10.1017/S1355771896000210>
- Zylinska, J. (2024). *La máquina de percepción: Nuestro futuro fotográfico entre la visión Y la IA*. Akal.