

**Formas de adaptabilidad: literatura y transmedialidad**  
Una muestra particular  
editado por Alessandro Mistrorigo y Enric Bou

# Voces inventadas

Felipe Cussen

Cuando Vicente Huidobro escribió «ia ia ia» en el Canto VII de *Altazor* no imaginó que casi cien años después un software Text-to-speech lo leería en voz alta como «Intelligence artificielle Intelligence artificielle Intelligence artificielle». Pero eso fue lo que sonó en mi computador cuando realicé el ejercicio de colocar ese canto ininteligible en una aplicación para que lo pronunciara en francés: el aullido se convirtió en una sigla.

Llevo muchos años obsesionado con este espacio límite de la experimentación de Huidobro, y particularmente con algunas versiones sonoras a partir de ellas, como la musicalización de Patricio Wang para el grupo Quilapayún y la grabación de Jaap Blonk con su propia voz. Me interesaba cómo en cada caso se ponían en relieve distintas dimensiones: en la primera, surgía como un canto tribal,

Este ensayo forma parte de mi proyecto *Música mística: Javiera Mena y Caterina Barbieri* (032591CA-REG). Proyecto Dicyt Regular - USACH, 2025-2028. Está basado en dos presentaciones previas: «Glosolalia en el canto VII de *Altazor* a través de una improvisación sonora» (IV Congreso Internacional de Investigación en Artes, Universidad Academia de Humanismo Cristiano, 18 de abril de 2024) y «Voces inventadas» (Simpósio Territorio y Escucha, Espacios Resonantes, II Festival de Arquitectura y Escucha. Galpón UDLA, 16 de abril de 2025). Agradezco los comentarios de Marcela Labraña, Megumi Andrade Kobayashi y Rodrigo Robledo.



**Biblioteca di Rassegna iberistica 46**

e-ISSN 2610-9360 | ISSN 2610-8844

ISBN [ebook] 978-88-6969-996-2 | ISBN 979-12-5742-008-6

**Open access**

Submitted 2025-10-01 | Published 2025-12-18

© 2025 Cussen | CC-BY 4.0

DOI 10.30687/978-88-6969-996-2/008

---

chamánico, mientras que en la segunda parecía más bien un poema fonético dadaísta.<sup>1</sup>

Me quedó rondando la posibilidad de hacer mi propia versión. Esto ocurrió cuando Rainer Krause me invitó a colaborar en una selección de poemas sonoros, bajo el título «we are one world», para la convocatoria «soundMIGRATION. soundart in a global context». Me pareció sugerente, pues me hizo pensar en los futuristas rusos y el «zaum», aquel lenguaje irracional que pretendía ir más allá de las lenguas nacionales, y, también en el espíritu utópico y pacifista de muchos poetas experimentales y artistas postales de la segunda mitad del siglo XX. Al mismo tiempo, dicho espíritu resonaba con los bienintencionados proyectos de artistas pop en los ochenta bajo la consigna «We are the world».

Para responder utilicé, como ya señalé, la vocalización computacional de este texto en diversos idiomas, es decir, con diversos acentos, como si fueran distintos colores. El resultado debería ser igualmente impenetrable para auditores de distintas culturas, pero intenté producir una gran conversación absurda entre robots de todo el mundo a través de 48 tracks que aparecían y desaparecían aleatoriamente. Lo titulé «we are not one world».<sup>2</sup>

Poco después Jaap Blonk vino a Chile y se organizó una presentación en la sala SCD con varios poetas y músicos. Le propuse que leyera su versión del canto VII, y posteriormente yo hice una improvisación a partir de mis voces robotizadas. Éste fue un cambio importante respecto a la versión previa, donde solamente había armado esa conversación sin alterar las voces. Acá opté por ponerlas muy rápido y manipularlas hasta estrujarlas lo más posible, mediante efectos granulares, de saturación y stutter (literalmente «tartamudez»). Ocupé varios controladores MIDI, más un sensor del movimiento de las manos, para que se percibiera en mis propios gestos las acciones que estaba ejerciendo. Fue, creo, una respuesta más violenta a esa utopía de lenguaje universal, como una forma de subrayar su fracaso o quizás la angustia ante la imposibilidad de comunicarse.<sup>3</sup>

Algunos años después estas disquisiciones continúan formando parte de mis reflexiones sobre las palabras, el sonido, las tecnologías digitales y sus formas de recepción. Son varias las líneas que aún quisiera profundizar. Por una parte, está la glosolalia, esta gran zona de lenguas inventadas que responden a iluminaciones

---

<sup>1</sup> Ver mi artículo «Canto VII de Altazor: lecturas críticas a través del sonido». *Confluencia*, vol. 29, n. 2, Spring 2014, 81-91.

<sup>2</sup> Ver «We are not one world», soundMIGRATION Chile 2018. *The W:OW Project*: <https://wow.nmartproject.net/soundmigration-selection-2018/>. El audio está disponible en <https://felipecussen.bandcamp.com/track/we-are-not-one-world>.

<sup>3</sup> Ver «We are not one world», Jaap Blonk en vivo, Sala SCD Bellavista, 11 de noviembre de 2017. El video está disponible en <https://www.youtube.com/watch?v=3UyTkFJWMhc>.

místicas, patologías, ficciones cinematográficas o búsquedas de una comunicación universal. Considero que hasta ahora ha primado un estudio de sus características como sistemas lingüísticos con elementos y reglas propias, pero sería interesante desarrollar más la relación con su dimensión visual y sonora. Tengo en mente las «otras escrituras» que propone Megumi Andrade Kobayashi en su estudio sobre Cy Twombly, Henri Michaux, Inoue Yuichi, Mirtha Dermisache y Mira Schendel. Se trata de escrituras que no corresponden a un código por una decisión estratégica, y que «aparentan un impulso lingüístico que pareciera querer comunicar pero no lo hacen bajo los mecanismos habituales del lenguaje», pues además están realizadas con herramientas y materiales propios de la escritura, y así se vinculan al campo del lenguaje verbal. Del mismo modo, me gustaría pensar, a partir de ese planteamiento, en una perspectiva paralela respecto a la tradición de la poesía sonora, en la que sus múltiples experimentos de distorsión y fragmentación puedan leerse también como una pantomima de una expresión oral, que también encarna la posibilidad de otra forma de comunicación. A diferencia de la glosolalia puntual de Huidobro, o de otras más complejas como la «*Lingua ignota*» de Hildegard von Bingen y tantas otras que vinieron después, me parece interesante que en este tipo de casos las unidades mínimas ya no son, respectivamente, grafemas y fonemas que aún podríamos reconocer (pues suelen corresponder a alfabetos ya disponibles), sino una serie de elementos inclasificables en un sistema específico, y de los cuales resulta imposible determinar una sintaxis particular. Creo que eso ocurría, precisamente, en mi manipulación del Canto VII, en el que aquellos sonidos inicialmente distinguibles eran reemplazados por otro tipo de texturas que se acercaban mucho más al ruido.

Otra ruta que me interesaría explorar es aquella que emerge a través del uso de la ya mentada Inteligencia Artificial, que cuando realicé mi versión del Canto VII se veía mucho más lejana que hoy, en que pareciera ser al mismo tiempo la mayor solución y el mayor problema de la humanidad. Recientemente hice la prueba de pedirle a ChatGPT que inventara una lengua nueva. Tras responder, como es de costumbre, con frases de buena crianza como «*Creating a new language can be a fascinating and creative Endeavor*», me propuso el «*Lumara*», que correspondería a una civilización de marinos, influenciada por las fases lunares y las mareas, cuyo folklore y poesía suelen referir a los cuerpos celestes y al océano. He aquí un ejemplo: «*Lhunara shola indrasi frondi, iluna tayalara shola timala. Solara yuven tayalara, indrasi shola cethani dilara. Lumara thalaindara lumo. Indrasi cethani lumara, iluna thivara, solara rethalara*» (su traducción al inglés sería: «*Underneath the swaying trees, the moon's light dances upon the waves. The sun bids farewell, and stars join*

the night's chorus. Lumara's heart glows softly. The trees whisper in Lumara, the moon sings, and the sun rests»).

Mucho más atractivo me produjo el lenguaje asémico creado a través de Inteligencia Artificial por So Kanno, quien solicitó a varias personas que escribieran en sus respectivos idiomas, para que la máquina luego reprodujera una nueva caligrafía creada a partir de las formas específicas de cada sistema de caracteres, pero sin un significado. Como explicó el artista, «the generated lines are written as if they meant something important, also look like trying to deceive».⁴ También son divertidos los resultados que obtuvo Jörg Piringer al colocar largas cadenas de vocales («aaaaaaaa...», «eeeeeee...», etc.) en el Text-To-Speech-Playground de OpenAI, ya que, a pesar de lo que se supondría, en vez de un largo sonido continuo, se generan muchísimas variaciones.<sup>5</sup> Por último, las recientes indagaciones de Rodrigo Robledo destacan por una mirada crítica al modo en que la Inteligencia Artificial trabaja con las identidades culturales. En su video «LATIN AMERICAN HUMANS IN A LATIN AMERICAN LATENT SPACE» trabaja a partir de imágenes de la Colección de Fotografía Patrimonial del Museo Histórico Nacional de Chile, pero además crea voces a partir de Stable Audio, un modelo creado originalmente para generar música, con un prompt que pedía que fuera una persona hablando como chileno. Esta intención de utilizar una herramienta con un sentido diferente al original (o «Databending con modelos de IA generativa», como lo llama) da como resultado unas fantasmáticas «voces que no dicen nada realmente».<sup>6</sup> En efecto, si escuchamos de manera distraída podremos percibir con claridad un acento e inflexiones típicamente nacionales, pero al aguzar el oído no conseguiremos decodificar el mensaje. Estos tres casos coinciden, a mi parecer, en una intención decididamente experimental de la IA, atenta a sus errores y desvaríos como posibilidades de exploración.

Dentro de mi propia práctica me he querido enfocar en la interacción de ciertas tecnologías musicales análogas y digitales más antiguas que se relacionan con la voz. Existen, por ejemplo, algunos efectos para guitarra eléctrica (como el Wah Wah y el Talkbox) que, a través del manejo de las frecuencias crea un efecto similar al de las formantes en el habla humana. De manera paralela, el Vocoder puede captar el espectro de frecuencias de la voz y de ese modo filtrar el sonido resultante de un sintetizador. Pero hay otras técnicas que permiten sonar de manera aún más realista, que habitualmente se engloban bajo el concepto de «speech synthesis», como el software

---

4 Disponible en [https://www.kanno.so/project/asemic\\_languages](https://www.kanno.so/project/asemic_languages).

5 Disponible en <https://soundcloud.com/jpiringer/sets/vowels-artificial-auditory-hallucinations>.

6 Disponible en <https://www.youtube.com/watch?v=x0wB246HKL8>.

que ocupé para reproducir los gemidos de *Altazor*. Existen dos modalidades: la primera es la síntesis concatenativa, que se basa en la unión de samples grabados de voces, y la síntesis de formantes, que intenta replicar el habla a través de síntesis aditiva y modelación. Algunas de estas últimas, en sus versiones más antiguas, son muy reconocibles debido a su uso en juguetes (como Speak & Spell de Texas Instruments) o en juegos de computador de las décadas de los '70 y '80. A diferencia de aquellos modelos actuales que buscan la mayor naturalidad e inteligibilidad posibles, éstas me resultan más atractivas no sólo por los recuerdos de infancia que evocan, sino justamente porque no suenan naturales, porque suenan robóticas.

Quise comenzar a experimentar, entonces, con este tipo de materiales retrofuturistas, para indagar en la invención de voces que sugirieran un lenguaje aún no existente. Para ello me basé en dos sintetizadores, el Arturia Microfreak a través de su máquina «Speech», y «Plaids» de Mutable Instruments, manejado desde un Poly Beebo, que remiten precisamente a los modelos vintage que recién indiqué. Ambos funcionan a partir de la generación de sus propios sonidos, con formantes que marcan las vocales y sonidos percusivos que corresponden a las consonantes; a veces también emiten algunas palabras completas. Mi modo de trabajo fue muy sencillo: fui moviendo aleatoriamente las perillas de ambos sintetizadores, que controlaban variables como el timbre, los armónicos o el tipo de formante, y grabé los sonidos que se producían. A partir de estos samples quise generar una estructura que sugiriera una posible conversación entre distintas voces. Para provocar un ritmo parecido al del lenguaje hablado no quise secuenciar estos fragmentos de acuerdo a un pulso fijo, sino que ocupé las grabaciones de las charlas de escritores famosos como Marguerite Duras y Gilles Deleuze y las traspasé a lenguaje MIDI. Así, cada «nota» lanzaría un sample distinto de las voces sintetizadas en una secuencia más cercana al habla habitual. Quise sugerir, también, a través de ecos desajustados, un espacio extraño, en el que no se supiera bien de dónde surgían estas voces.<sup>7</sup> Finalmente, a partir de los mismos samples, desarrollé una estructura más rítmica a través de una secuencia que, primero, operaba de manera fija, y posteriormente se volvía aleatoria: como un coro que juega a inventar sonidos y que cada vez se vuelve más caótico.<sup>8</sup>

Tanto en mi versión del Canto VII de *Altazor* como en estos ejercicios recientes, he buscado, a través de tecnologías de

<sup>7</sup> Disponible en <https://felipecussen.bandcamp.com/track/voces-inventadas-fragmento-1>.

<sup>8</sup> Disponible en <https://felipecussen.bandcamp.com/track/voces-inventadas-fragmento-2>.

síntesis creadas para producir un habla comprensible, explorar las posibilidades de encontrar lenguajes desconocidos a través de su manipulación. Más aún, y al igual que en los ejemplos que me han inspirado, he pretendido indagar en el efecto que pueden producir estos sonidos, pues me imagino que en algunos momentos podrían sugerir un habla oculta y, en otros, simplemente se convertirían en una masa informe de sonidos indiscernibles. Ese tipo de conexión o desconexión respecto a la posibilidad de comprender, quizás un umbral que no necesariamente todos calibrámos de la misma manera, me resulta muy atractivo porque pone al desnudo nuestras expectativas frente a una potencial comunicación imposible. Brandon Labelle describe la forma de escucha que busca como una posición de «no saber», y creo que las distintas formas de glosolalia, de prácticas de poesía sonora, así como también aquellas caligrafías que desafían nuestras formas de lectura, nos exigen precisamente esa disposición, y nos desafían con una nueva pregunta: ¿qué se entiende cuando no se entiende?