

Sección Tres: Experiencias, reseñas, debates e informes

Teoría crítica y marxismo en las ciencias sociales y humanas: alcances, limitaciones y reconfiguraciones. Sociología transformadora.

La voce artificiale. Un'indagine media-archeologica sul computer parlante.¹

The artificial voice. A media-archaeological investigation on the talking computer.

Luciano Petullà
Investigador independiente y
bloguero
luciano@lucianopetulla.net

Reseña del libro *La voce artificiale. Un'indagine media-archeologica sul computer parlante*. Napoli: Edizione Scientifica. Domenico Napolitano (2022).

El sonido juega un papel muy importante en la vida social, pero sigue siendo poco debatido en ciencias sociales y en la comunicación digital. El libro de Domenico Napolitano quiere compensar esta carencia enfocándose en particular en las tecnologías informáticas de los asistentes de voz – Siri, Alexa, Google Home, etc. – para adentrarse en la dinámica constructiva y arqueológica de la voz; e interpretar cada acontecimiento sonoro como artefacto cultural y social.

El autor del libro es un joven investigador social y experimentado organizador de espectáculos sonoros, y artista él mismo en este campo. Por eso muestra una gran capacidad de abordar con tenacidad e ingenio un amplio espectro de cuestiones, examinándolas desde perspectivas de estudio diferentes pero fructíferamente interrelacionadas – filosóficas, sociológicas, técnicas, literarias, lingüísticas, informáticas, económicas, estéticas e incluso etnográficas, con entrevistas a profesionales que desarrollan estrategias y usos de la inteligencia artificial en el ámbito vocal. Desde este punto de vista, la investigación sobre el ordenador que ha aprendido a responder y hablar con una voz sintética – que ahora también puede modularse a voluntad con las características de una determinada voz humana (*cloning*) –, además de ser una descripción muy clara de todas las facetas del fenómeno, es un manual generoso e instructivo

¹ Recibido:10/01/2023 Evaluado: 12/02/2023 Aceptado: 15/03/2023

sobre cómo analizar un medio tecnológico en general, en su construcción y recepción como artefacto social y cultural.

Así, la investigación se realiza siempre en un doble nivel – sobre la especificidad de la nueva tecnología vocal y sobre las implicaciones más generales de los nuevos ensamblajes entre el hombre y la máquina. Como afirma el autor al comentar las tesis posthumanistas, "la voz artificial puede verse como la expresión de una condición que se da *entre* los seres humanos y la tecnología, una modulación de la frontera móvil que los relaciona" (p. 425). Este continuo desplazamiento entre lo particular y lo general es realmente útil en un contexto tan complejo como el abierto por la difusión de las tecnologías digitales en red y el desarrollo de productos informativos que – al informatizar y filtrar todo tipo de actividades y comportamientos – impregnán entornos y relaciones de todo tipo. A todos los efectos, se trata de una obra de *hacking* y desmitificación que consigue abrir e ilustrar las *black boxes* algorítmicas a las que estamos sometidos, aclarando sus principios de funcionamiento y los riesgos potenciales inherentes a cada nuevo entrelazamiento entre humanos y máquinas.

Entre las muchas cuestiones que se abordan está, por ejemplo, el análisis del nuevo software inteligente, del que aprendemos que detrás del buen funcionamiento de una inteligencia artificial hay cálculos estadísticos que no tienen que entender el significado de los signos, sino *sólo* identificar la probabilidad de que ciertos datos textuales se correlacionen con ciertos datos sonoros – lo que se consigue si los algoritmos pueden acceder a una increíble cantidad de muestras de hechos reales. De este modo las palabras son una simulación *data driven* de efectos vocales preadquiridos, operaciones poderosas que pueden permitirse el lujo de obviar las habilidades internas, semántica y hermenéutica, a la que el ser humano está ligado. Al mismo tiempo, podemos ver cómo la capacidad de externalizar funciones humanas tan sofisticadas – que luego se hibridan íntimamente con nuestras vidas afectando a la voluntad de decisión de las personas – queda en manos de empresas que responden a intereses que suelen eludir – a pesar de su poder efector – los principios de responsabilidad (ética y social) que normalmente rigen las actividades comunitarias.

Estas consideraciones nos hacen entender por qué el estudio de una tecnología de la comunicación abarca un espectro tan amplio de cuestiones. En primero lugar, hay que tener en cuenta un imaginario – el de las voces sin cuerpo – que es tan antiguo como la humanidad, así como los numerosos intentos de construir máquinas parlantes. Pero, sobre todo, el hecho de que las facultades comunicativas ligadas al lenguaje juegan un papel central para el ser humano, y la voz, también como marcador corporal único, es su medio por excelencia. Con los ordenadores que entienden y producen el lenguaje vocal mediante procesos computacionales a partir de datos acústicos, entramos en un nuevo escenario existencial en el que el primer supuesto que se desmorona es que toda voz supone un cuerpo humano. Esto ya pone en tela de juicio toda una serie de supuestos: la voz, tanto en la acción de hablar como en la de responder, siempre ha sido experimentada y entendida como un principio de individuación humana, y también como una indicación de una presencia plena y responsable – un concepto muy debatido en filosofía en términos de una "metafísica de la presencia" que tiende a criticar los privilegios atribuidos a la función fonética/performativa con respecto a otros signos de presencia. Para el autor, por tanto, la voz artificial es un "fenómeno sociotécnico que se mueve transversalmente entre tecnologías, conocimientos, historias, deseos, intereses e imaginarios, y que concierne a estructuras antropológicas, sociales y epistemológicas al mismo tiempo" (p. 19).

Se menciona así mismo la posible utilidad de la obra como guía instructiva para investigar las tecnologías de los medios de comunicación. De hecho, son varios los factores que juegan a favor de esta tesis: el enfoque metodológico; el amplio numero de referencias teórico-críticas a los maximos estudiosos del sector en el panorama internacional; la inteligibilidad discursiva. También es importante el esfuerzo de integrar las teorías tomadas como referencia – arqueología de los medios, materialismo y construcción (Science and Technology Studies, Actor-network Theory, Critical Algorithm Studies, Sound Studies) – para ilustrar cómo el proceso de "hacerse voz" es histórico y social, y cómo las tecnologías también tienen que ser sondeadas en su materialidad, así como en los contextos de su uso, siendo ellas mismas verdaderas "formas de pensar" en la medida en que las actualizan y las realizan "encarnando epistemologías, psicologías y capacidades operativas" (p. 70).

Para concluir, es interesante mencionar la utilidad de recurrir a las experiencias artísticas relacionadas con el sonido por su valor de contribución a las cuestiones de investigación. La voz y sus efectos, también en la intersección con los artefactos tecnológicos – "la voz *en la máquina*" y "la voz *de la máquina*" – están de hecho en el centro de las reflexiones y prácticas artísticas que, gracias a los impulsos experimentales, tienen la capacidad no sólo de representar nuevas formas de acontecimientos, sino también de anticipar innovaciones creativas que son incluso candidatas a convertirse en productos comerciales. En la práctica, los artistas sonoros desvelan al público – y aquí hay muchos ejemplos – lo que puede suceder en los laboratorios especializados de alguna empresa cuando persisten los escenarios técnicos y socioculturales actuales. En estas obras, apoyadas por el uso de los últimos programas informáticos – *machine learning* y redes neuronales hábilmente gestionadas – se puede, por ejemplo, participar en retos improvisando dúos en tiempo real en el escenario entre la propia voz encarnada y la misma voz clonada pero generada y cantada por un ordenador – en expresiones vocales imposibles para un ser humano. El tema aquí es la frontera entre la identidad corporal y la máquina: la voz es efectivamente la del artista, pero al mismo tiempo es también la de la máquina. También es posible encontrar prácticas artísticas que revelan tanto el poder como la imprevisibilidad del diseño de los procesos más generales de datafificación a los que todo parece estar sometido hoy en día, y esto no puede sino ser una fuente de inquietud. La datafificación de todo fenómeno y el uso de *software* capaz de relacionarlos – por el mero hecho de haberse convertido en datos puros – son capaces de generar objetos/fenómenos híbridos que combinan, a voluntad, algunas de sus características singulares. Así, nos enteramos de que existen síntesis de voz capaces de hibridar el timbre de la voz de alguien con el estilo prosódico de otra persona, y esto se aplica a cualquier otra característica paralingüística (conversión de la voz): varias empresas ya ofrecen como producto las llamadas *skin voices*.

El trabajo que realiza Napolitano con este libro, es de interés para todas las profesiones que tienen que ver con la voz, el sonido y la comunicación, aportando temáticas de interés poco usuales pero imprescindibles en la difusión científica.

