

FORMATO DE PROYECTO

1. Título.

gla 1.0b (generador de lenguaje asémico)

2. Descripción breve del proyecto. Se trata de una descripción general y sintética de la obra (Máximo 500 palabras).

El proyecto consiste en una instalación interactiva que invita a escuchar los idiomas del mundo, apelar a nuestro imaginario para lograr entender algo que en principio no hará sentido, pero interpelará nuestras subjetividades.

Quienes experimenten la obra se encontrarán con un dispositivo pequeño en un pedestal con un dos auriculares colgados al frente. El aparato es un controlador con perillas rotuladas que representan cada una un idioma distinto. Al interactuar con el sistema se escucharán monólogos teniendo la posibilidad de mezclar cada lengua, descubriendo qué aparece en su interpolación.

3. Fundamentación / Marco conceptual (Máximo 1000 palabras).

El sonido en el habla comenzó a interesarme a partir de un viaje en el que recorrió varios países. Escuchar a la gente a mi alrededor despertó una inquietud acerca de cómo sonaban los distintos idiomas. Grabando material y mediante una escucha reducida comencé a establecer ciertas relaciones entre lenguas que, en un plano geográfico y/o cultural, no tenían relación alguna. ¿Qué sucedería si se plantea un mapa donde las distancias entre territorios sean acordes a la similitud de sus fonemas? Entendiéndose por fonemas a la unidad mínima de sonido para formar palabras.

Resulta curioso cuando las personas pueden identificar los idiomas sin saber hablarlos; quizás escuchando música de algún lugar, TV o cualquier material donde se escuche gente hablando, hay quienes logran identificar correctamente qué idioma está escuchando (sería interesante conocer una estadística de identificaciones correctas en un grupo). ¿Por qué sucede esto? ¿Qué características de una lengua sintetizan su identidad? ¿Cuáles son los aspectos que los vuelven familiares?

Me interesa en particular el habla que no hace sentido, así como existe la escritura asémica, por qué no pensar en habla asémica, como un balbuceo. Entre estas preguntas e inquietudes, más el uso de tecnologías de inteligencia artificial surge la idea de crear un sintetizador de habla asémica, es decir, habla que carezca de sentido, lo cual lo diferencia de un TTS (texto a habla) en el que el usuario escribe un texto y luego es reproducido en una voz sintetizada. Este software sería capaz de generar habla que nos remita a cierta lengua que el usuario pueda elegir, incluso la posibilidad de mezclar varias, sin ingresar contenido semántico. La mezcla consistiría en la interpolación dentro de la red neuronal generadora buscando lo que hay dentro del espacio latente luego de haber sido entrenada con grandes datasets de audio

clasificado en cada lengua. De esta manera podríamos establecer la forma conocida, o familiar, y la forma desconocida, que viene de esta interpolación.

Quizás esta tecnología pueda dar al menos un acercamiento a conocer los aspectos formales sonoros que identifican a un lenguaje particular, y además descubrir qué hay entre, por ejemplo, algunos en los que intuitivamente podemos hallar similitudes como fonemas coincidentes o aquellos que sean completamente ajenos entre sí. ¿Qué sucede entre lenguas? ¿Esos resultados nos serán familiares? ¿Por qué? ¿Qué otras preguntas pueden llegar a surgir a partir de escuchar estos nuevos sonidos?

El proyecto pretende investigar este aspecto y así generar espacios para resultados concretos como el software de habla propuesto anteriormente, recreación de voces como familiares y la propia, performances de sono-lenguas, piezas sonoras, instalaciones, personajes virtuales que las hablen, entre otros. Esta exploración propone un recorrido donde sus ramas se retroalimentan con estos resultados para continuar profundizando, generando un cartografía, o mejor aún, una red de conocimiento.

4. Descripción detallada del proyecto (Máximo 1000 palabras).

El proyecto del generador de habla asémica se enmarca en mi trabajo final de grado de la carrera de Lic. en Artes Electrónicas tutoreado por Leandro Garber. Esta instalación en particular es uno de los productos de la investigación centrada en el software generador de lenguajes. La idea es observar qué ocurre en los cruces de lenguas, haciendo foco en la cuestión morfológica de los sonidos de los idiomas, eliminando su contenido semántico. Cada lengua generada se asemejará a una ya existente, apelando al recuerdo e imaginario del espectador, pero sin dar ningún contenido en su mensaje.

Quienes experimenten la pieza se encontrarán con un dispositivo con perillas en su frente y auriculares para usar en los laterales, de manera que también invite a realizarlo en duplas, compartiendo la experiencia de escucha.

El audio de habla sonará constantemente, a menos que todas las perillas estén en cero.

5. Información técnica del proyecto (Máximo 1000 palabras).

El desarrollo técnico de la pieza consiste en un dispositivo controlador con perillas montado sobre un pedestal negro cuyo frente tiene impreso un texto con una breve descripción e instrucciones de uso. En el interior del pedestal estará una computadora corriendo el software generador de habla. La salida de audio de la computadora será dividida en 2 con un splitter pasivo para conectar el par de auriculares.

El controlador se asemeja a un sintetizador de audio modular, mostrando un frente con perillas rotuladas por cada idioma, cada una controlando la *cantidad* de características de la lengua que suena. A medida que se interactúe con estos controles se podrá mezclar e interpolar entre los sonidos de los idiomas, dando a conocer nuevas hablas generados a partir de esta combinación.

El software se desarrollará generando modelos de inteligencia artificial basados en autoencoders por cada lengua elegida. Para entrenar estos modelos se utilizarán datasets de audio de libre acceso. Combinando estos modelos se podría entrenar otro para la exploración del espacio latente de estos auto-encoders y así lograr interpolar y descubrir qué hay entre cada idioma. Se realizará en Jupyter Notebooks basando el código en PyTorch (librería de Python para IA). Para explorar estas combinaciones el dispositivo físico hará de interfaz MIDI que enviará los valores de cada potenciómetro al sistema.

El habla se escuchará a través de los auriculares siempre y cuando alguna de las perillas esté en un valor mayor a cero. De manera que sólo estará en silencio si todas están en cero.

Los elementos requeridos para el montaje son el pedestal negro y 2 auriculares.

6. Cronograma de trabajo (Máximo 1000 palabras).

Plan de trabajo para los 3 meses de desarrollo.

Mes 1:

- Recolección de datasets de cada idioma.
- Entrenamiento de modelos IA de cada idioma.
- Interpolación de modelos, generación de cruces.
- Diseño y construcción de controlador físico para la instalación.

Mes 2:

- Aplicación MIDI para navegar y testear el modelo combinado.
- Fine tuning de modelos, más entrenamiento.

Mes 3:

- Experimentación con el software, ajustes finales.
- Pruebas de montaje con el hardware requerido.

7. Presupuesto de Máxima y de Mínima.

Presupuesto Mínimo

Producto	Cantidad	Precio	Descripción
Suscripción Google Colab Pro+	1	\$ 30.000	500 compute units para entrenamiento de modelos
Raspberry Pi 4 Kit	1	\$ 103.219	Computadora para el montaje que correrá el software. MercadoLibre
Splitter audio	1	\$ 672	Salida de ambos auriculares MercadoLibre
Construcción de dispositivo físico	1	\$ 100.000	Potenciómetros, perillas, construcción de estructura metálica, rotulados y textos.
Honorarios de artista	1	\$ 320.000	

Total de materiales: \$ 553.891.-

Presupuesto Máximo

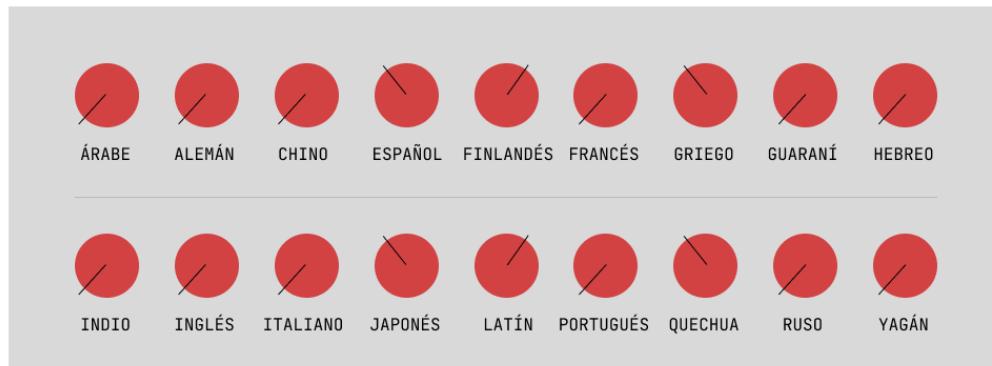
Producto	Cantidad	Precio	Descripción
Placa de video Nvidia Asus TUF Gaming GeForce RTX 40 Series RTX 4070 Ti TUF-RTX4070TI-12G-GAMING 12GB	1	\$ 600.000	Entrenamiento de modelos
Raspberry Pi 4 Kit	1	\$ 103.219	Computadora para el montaje que correrá el software. MercadoLibre
Splitter audio	1	\$ 672	Salida de ambos auriculares MercadoLibre
Construcción de dispositivo físico	1	\$ 100.000	Potenciómetros, perillas, construcción de estructura metálica, rotulados y textos.

Tota de materiales: \$ 1.103.891.-

La diferencia entre el monto otorgado y el total del costo será cubierto por el artista.

8. Imágenes / Registro audiovisual.





9. Autor/es/as (Máximo 1000 palabras).

Soy Juan Martín Sesali Maydana, músico, programador y artista electrónico. Comencé mis estudios universitarios en 2010 en Ingeniería en Informática, pasando por el ITBA y la UADE. Mis inquietudes artísticas me llevaron a virar hacia la Licenciatura en Artes Electrónicas en la UNTREF en 2014. Empecé a formar parte de Intercambios Transorgánicos como becario de investigación formando parte del equipo técnico desde 2014 hasta 2018, y luego mantuve un vínculo como consultor. Mientras tanto, desarrolle profesionalmente mi carrera como diseñador gráfico y web.

Hago música desde que tengo uso de razón. Desde 2017 publico música en plataformas digitales. Hacia 2018, comencé a realizar trabajos de forma freelance para productoras audiovisuales, enfocándome en montajes y desarrollo de software dedicado para experiencias interactivas y acciones en instalaciones. En 2019, lancé un álbum solista bajo el nombre MAY NADA, titulado "casa", donde recopilé creaciones sonoras desde mi adolescencia hasta el momento. Fue un trabajo realizado de forma individual con mis propias herramientas, abarcando todos los frentes que implican un lanzamiento: composición, producción, mezcla, masterización, diseño gráfico y gestión.

En 2021, fui contratado por Raven, una empresa ubicada en Houston, TX, como desarrollador frontend, pero también desempeñé funciones de desarrollo de bajo nivel, soporte y comunicación, y escribí gran parte de la documentación oficial. Raven ofrece productos y servicios para la digitalización de documentación, como scanners y soluciones en la nube para la gestión de archivos.

En 2022 me volqué a la escena Club como DJ y productor musical, realizando viajes y trabajos freelance en sistemas embebidos y automatizaciones. Actualmente, también investigo acerca de interfaces para la creación sonora, participando y dialogando en proyectos tecnológicos open source con fines musicales.