

## Giorgio Assistant

Giorgio Assistant es un Proyecto con la finalidad de reconocer las habilidades del reconocimiento de voz utilizadas en los navegadores de hoy en día para capturar la voz de un usuario o usar el dictado de un texto mediante una gramática definida por el navegador y la web en general.

Esta página puede ser visitada mediante la siguiente url:

<https://jorgitoayala.github.io/GiorgioAssistant/>



# Giorgio Assistant

Idioma del asistente: Español ▼ Tema de la aplicación: Claro ▼

 Guardar cambios

Que 🤖 lo repita

## Convertidor de voz a texto



Speech to Text

## Convertidor de texto a voz



Text to Speech



## Giorgio



En lo que respecta al sitio web, esta tiene las siguientes funcionalidades:

- Reconocer la voz en dos diferentes idiomas: inglés y español
- Reconoce la voz por dictado y acto seguido la escribe automáticamente (Speech to Text)
- Reconoce el texto escrito y acto seguido la repite por voz (Speech to Text)
- Si se desea dictarle algo a Giorgio y deseas que lo repita con una voz entonces puedes hacerle click en “Que Giorgio lo repita” para que se haga la funcionalidad.

Utilidades de la aplicación:

- Mostrar con propósito educacional la funcionalidad de la aplicación web y demostrar las ventajas que esto trae en la implementación de otros sitios web.
- Comprobar la dicción de un usuario mediante el uso de un dispositivo de entrada (micrófono) al momento de recibir su voz mediante este medio.
- Analizar como funciona internamente la aplicación mediante la observación de la construcción de las oraciones durante el dictado.

Esta página web utiliza las propiedades de la API de Web Speech implementada años anteriores en la mayoría de los navegadores de ahora en la cual se detallará a continuación:

### **Web Speech API**

La Web Speech API le permite incorporar datos de voz en aplicaciones web. La Web Speech API tiene dos partes: SpeechSynthesis (Text-to-Speech) y SpeechRecognition (Reconocimiento de voz asíncrono).

#### **Conceptos y uso de Speech Web API**

La Web Speech API permite que las aplicaciones web manejen datos de voz. Hay dos componentes en esta API:

Se accede al reconocimiento de voz a través de la interfaz SpeechRecognition, que brinda la capacidad de reconocer el contexto de voz desde una entrada de audio (normalmente a través del servicio de reconocimiento de voz predeterminado del dispositivo) y responder adecuadamente. Por lo general, usará el constructor de la interfaz para crear un nuevo objeto SpeechRecognition, que tiene varios controladores de eventos disponibles para detectar cuándo se ingresa voz a través del micrófono del dispositivo. La interfaz de SpeechGrammar representa un contenedor para un conjunto particular de gramática que su aplicación debe reconocer. La gramática se define usando JSpeech Grammar Format (JSGF).

Se accede a la síntesis de voz a través de la interfaz SpeechSynthesis, un componente de texto a voz que permite que los programas lean su contenido de texto (normalmente a través del sintetizador de voz predeterminado del dispositivo). Los objetos SpeechSynthesisVoice representan diferentes tipos de voz y diferentes partes del texto que desea que se hablen están representados por objetos SpeechSynthesisUtterance. Puede hacer que se hablen pasándolos al método SpeechSynthesis.speak().

## Reconocimiento de voz

El reconocimiento de voz implica recibir el habla a través del micrófono de un dispositivo, que luego un servicio de reconocimiento de voz verifica con una lista de gramática (básicamente, el vocabulario que desea que se reconozca en una aplicación en particular). Cuando una palabra o frase se reconoce con éxito, se devuelve como resultado (o lista de resultados) como una cadena de texto, y como resultado se pueden iniciar más acciones.

Web Speech API tiene una interfaz de controlador principal para esto, SpeechRecognition, además de una serie de interfaces estrechamente relacionadas para representar gramática, resultados, etc. En general, el sistema de reconocimiento de voz predeterminado disponible en el dispositivo se utilizará para el reconocimiento de voz, la mayoría. Los sistemas operativos modernos tienen un sistema de reconocimiento de voz para emitir comandos de voz. Piense en Dictado en macOS, Siri en iOS, Cortana en Windows 10, Android Speech, etc.

## Aplicaciones del reconocimiento de voz

- **Sistemas de coche:** Para controlar algunas operaciones dentro del coche. Como, por ejemplo, activar el GPS, realizar una llamada, o reproducir una canción.
- **Dictado de voz:** El dictado de voz es un sistema que escribe automáticamente todo lo que dictamos. Es la aplicación más común en el reconocimiento de voz.
- **Control por comandos:** Consta de una serie de órdenes para controlar el ordenador. La cantidad de comandos que identifica es baja.
- **Ámbito telefónico:** Permite ejecutar órdenes mediante el habla en una conversación telefónica a un agente no humano, en lugar de pulsar botones físicos o táctiles.
- **Dispositivos móviles:** Puede utilizarse, también, en dispositivos de tamaño reducido y móvil (relojes inteligentes, teléfonos móviles, acelerómetros, etc.) para emitir órdenes y evitar manipularlos en situaciones incómodas.
- **Sistemas para personas con discapacidad:** Dependiendo del tipo de discapacidad que una persona posea, utilizará un dispositivo u otro, y de acuerdo a este dispositivo, se puede utilizar un sistema de reconocimiento de voz para ayudar a manejar su problema.

## Características del reconocimiento de voz

- Para perfeccionar los sistemas de reconocimiento del habla, **han tenido que intervenir varias disciplinas:** inteligencia artificial, ciencias de la computación, procesamiento de señales, lingüística, acústica y fisiología.
- Los primeros indicios de esta disciplina se encuentran en el año 1950.
- Un elemento clave para el gran avance que ha tenido el reconocimiento de voz en los últimos años, ha sido gracias al machine learning.
- Las empresas más relevantes en este ámbito, son: Apple (Siri), Microsoft (Cortana), y Amazon (Alexa). Siendo estas las grandes corporaciones que han traído las tecnologías más notables a la industria.