



ReconVox es nuestro motor **Robusto de Reconocimiento del Habla** de alta eficiencia, capaz de ejecutarse en entornos empotrados con estrictos requisitos hardware bajo condiciones ruidosas y de utilizar vocabularios 100% personalizados. La solución está diseñada para ser incorporada en prácticamente cualquier entorno, ya que está disponible en forma de **SDK** abierto que exporta su funcionalidad a través de un sencillo pero potente **API** (*Application Programming Interface*).

ReconVox soporta tres modos de reconocimiento: **lenguaje natural** (para obtener la transcripción completa en habla conversacional), **vocabulario definido a medida** (para reconocimiento de órdenes de voz en entornos empotrados o forzar un vocabulario específico en call-centers) y modo **Word-Spotting** (para detectar palabras clave en tiempo real en condiciones muy desfavorables: entornos acústicos ruidosos, habla desestructurada, múltiples idiomas entremezclados, etc).

Clave para la robustez en entornos acústicos problemáticos es nuestra exclusiva función de **AutoLearn**. Con esta tecnología el motor de reconocimiento es capaz de adaptarse sobre la marcha a condiciones específicas del audio para así mejorar enormemente la precisión en los resultados: un entorno acústico ruidoso, un locutor con características vocales muy particulares o incluso una región dialéctica con una pronunciación muy marcada.

Su flexibilidad permite una gran variedad de **aplicaciones**:

- **Transcripción automática de lenguaje natural**: en reuniones o para subtítulo de TV o cine.
- **IVR en call-centers**.
- **Monitorización en tiempo real de llamadas y búsqueda por contenido**: detección de palabras clave o frases cortas definidas sobre la marcha y lanzamiento de alarmas.
- **Asistentes de voz en vehículos**: sistemas de *infotainment* controlados mediante la voz, robustos al entorno acústico ruidoso presente en un coche.
- **Domótica y alarmas**: control de dispositivos electrónicos mediante la voz, activación de alarmas...



Por otra parte, si las necesidades de seguridad así lo requieren, un sistema de Reconocimiento del Habla basado en **ReconVox** puede trabajar conjuntamente con nuestro sistema de Biometría de Voz **BioVox**. De este modo es posible la verificación continua de la identidad del locutor en segundo plano y de manera

transparente, para así lograr transacciones seguras a lo largo de toda la interacción con el sistema.

PRODUCTO

- Sistema de Reconocimiento del Habla en Lenguaje Natural robusto frente a ruido.

CARACTERÍSTICAS CLAVE

- **Tres modos de reconocimiento:**
 1. **Lenguaje Natural:** transcripción de habla conversacional.
 2. **Vocabularios definidos desde cero:** dos tipos de modelo de lenguaje soportados.
 - Abierto: habla continua dentro de un vocabulario específico.
 - Cerrado: gramáticas basadas en reglas fijas.
 3. **WordSpotting:** detección de palabras o frases clave definidas sobre la marcha sin necesidad de un modelo de lenguaje.
- **AutoLearn:** adaptación automática a un locutor determinado, región dialéctica o entorno ruidoso.
- **Idiomas disponibles:** Afrikaans, Arabic, Armenian, Azerbaijani, Belarusian, Bosnian, Bulgarian, Catalan, Chinese, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Greek, Hebrew, Hindi, Hungarian, Icelandic, Indonesian, Italian, Japanese, Kannada, Kazakh, Korean, Latvian, Lithuanian, Macedonian, Malay, Marathi, Maori, Nepali, Norwegian, Persian, Polish, Portuguese, Romanian, Russian, Serbian, Slovak, Slovenian, Spanish, Swahili, Swedish, Tagalog, Tamil, Thai, Turkish, Ukrainian, Urdu, Vietnamese, and Welsh.
- **Frecuencias de muestreo soportadas:** modelos acústicos disponibles para **16 KHz** y **8 KHz**.
- **Motor de reconocimiento eficiente en C++:** apto para entornos empujados.



ESPECIFICACIONES TÉCNICAS

- Preprocesado de la señal: detección de actividad, *dithering*, compensación de la componente continua, filtrado LMS.
- Formatos de audio: PCM lineal 16 bits 8/16 KHZ (recomendado), Ley-A, Ley- μ , MP3.
- Consumo de HW según el modo de reconocimiento:
 - Modo 1: 153 MB – 3.3 GB RAM, 75 MB – 3 GB HDD (todos los idiomas, tamaño de modelos acústicos configurable).
 - Modos 2 y 3: 35 MB RAM, 5 MB HDD (por idioma).

SISTEMAS OPERATIVOS

- Windows® 7, 8, 10, 11.
- Linux, diferentes distribuciones.

