

Manual de referencia: ElevenLabs

Datos de esta manual

Fase	MAN — Manuales de referencia
Módulo	MAN — Biblioteca de consulta

Objetivo práctico

Este manual reúne, en un único documento de consulta, todo lo que necesitas saber sobre ElevenLabs como pieza VOZ del sistema: planes y precios vigentes, modelos de voz y su cobertura en español, gestión y seguridad de la clave de API, límites de créditos, accesibilidad real del panel con NVDA, atajos disponibles y el catálogo completo de errores frecuentes de generación de voz con su diagnóstico. Úsalo como referencia cuando necesites resolver una duda puntual, no como una lección que debas leer de principio a fin.

Resultado que obtendrás

Un documento de consulta permanente al que volver cada vez que trabajes con ElevenLabs: desde la primera configuración de la cuenta hasta la resolución de un error concreto de generación de voz meses después.

Dónde encaja dentro del sistema final

Dentro del sistema TELÉFONO → AGENTE → CEREBRO → VOZ → AUTOMATIZACIÓN → CALENDARIO, ElevenLabs ocupa siempre la pieza VOZ: convierte el texto que decide el CEREBRO (el LLM) en el audio que escucha la persona que llama. No sustituye ni al AGENTE (Vapi o Retell, que orquestan la llamada completa) ni al CEREBRO (el LLM que decide qué decir); su responsabilidad se limita a cómo suena lo que ya se ha decidido decir. La clave de API de ElevenLabs es, además, uno de los tres o cuatro secretos que conectan literalmente las piezas del sistema entre sí: sin ella, Vapi o Retell no pueden generar audio con una voz de ElevenLabs durante la llamada real.

Conocimientos previos

- Haber completado el módulo 2.2 (voz con ElevenLabs) dentro de la progresión del curso.
- Conocer el concepto de TTS (texto a voz) del módulo 1.2.
- Protocolo de creación segura de cuentas y de gestión de claves de API del módulo 2.1.

Herramientas necesarias

- Una cuenta de ElevenLabs (gratuita como mínimo).
- Un navegador con NVDA activo.
- Un gestor de contraseñas para guardar la clave de API si la generas.

Posible coste

Plan gratuito (verificado el 30 de agosto de 2026): el plan Free de ElevenLabs no tiene coste e incluye 10.000 créditos mensuales, aproximadamente equivalentes a 10 minutos de generación de voz con el modelo Multilingual v2, o a unos 15 minutos de uso de un agente conversacional. El plan Free no incluye derechos de uso comercial: cualquier contenido generado debe atribuirse a ElevenLabs y no puede monetizarse legalmente sin pasar a un plan de pago. Tampoco incluye clonación instantánea de voz ni clonación profesional de voz (PVC); el catálogo de voces se limita a la biblioteca de voces de stock.

Planes de pago (verificados el 30 de agosto de 2026 contra la página oficial de precios):

- **Starter**, 6 USD/mes: primer escalón de pago, pensado para pruebas más allá del plan gratuito.
- **Creator**, 22 USD/mes (11 USD el primer mes en la oferta vigente en la fecha de consulta): 100.000 créditos mensuales, clonación de voz profesional (PVC) y salida de audio a 192 kbps. Es el plan orientado a creación de contenido y a las primeras pruebas serias de un proyecto de agente telefónico.
- **Pro**, 99 USD/mes.
- **Scale**, 299 USD/mes.
- **Business**, 990 USD/mes: 6.000.000 de créditos mensuales y 10 puestos, con TTS de baja latencia desde 0,05 USD por minuto en el escalón más alto de volumen.
- **Enterprise**: precio a medida, con acuerdo de protección de datos (DPA) y de nivel de servicio (SLA) personalizados, BAA para clientes sujetos a HIPAA, SSO a medida y soporte prioritario.

La facturación anual aplica un descuento aproximado del 17 % (dos meses gratis) sobre el precio mensual equivalente.

Consumo por unidad (overage), verificado el 30 de agosto de 2026: el sistema de créditos es unificado, no por «caracteres» separados de otros productos: un crédito equivale aproximadamente a un carácter de texto en TTS estándar, y cada producto (transcripción, doblaje, etc.) consume créditos a su propia tasa. Como referencia de coste por unidad fuera de plan: TTS con el modelo Flash, 0,05 USD por cada 1.000 caracteres; TTS con el modelo Multilingual, 0,10

USD por cada 1.000 caracteres; transcripción, desde 0,22 USD por hora; doblaje, entre 0,33 y 2,20 USD por minuto. Verifica siempre estas cifras en elevenlabs.io/pricing antes de presupuestar un proyecto, porque cambian con más frecuencia que otras partes de la documentación.

Advertencia de seguridad

Protección de la clave de API: según la documentación oficial de seguridad de ElevenLabs, la clave de API es un secreto: no debe compartirse ni exponerse nunca en código que se ejecute en el navegador o en una aplicación cliente (el usuario final de esa aplicación podría leerla). Todas las peticiones a la API deben incluir la clave en la cabecera HTTP `xi-api-key`. Para escenarios donde el cliente sí necesita autenticarse (por ejemplo, un widget en una página web), ElevenLabs ofrece tokens de un solo uso, válidos durante un tiempo limitado, que permiten conectar con la API sin exponer la clave principal.

Tipos de clave y su alcance: una clave de API de usuario hereda el acceso de la persona que la crea, admite fecha de caducidad y deja de funcionar si esa persona pierde el acceso al espacio de trabajo o lo abandona; una clave de cuenta de servicio pertenece a un rol de servicio, no a una persona concreta, y sigue funcionando aunque cambie la composición del equipo. Cada clave admite además restricción de alcance (a qué endpoints concretos puede acceder), una cuota de créditos propia independiente del resto de la cuenta, y una lista blanca de direcciones IP o rangos CIDR autorizados; una petición desde una IP no autorizada se rechaza con un error 403.

Filtración accidental: si una clave de API de ElevenLabs se publica por error en un repositorio público de GitHub, GitHub notifica a ElevenLabs y la clave se desactiva automáticamente para evitar un uso no autorizado. Aun así, no dependas de esta protección: sigue el protocolo del módulo 2.1 (variables de entorno, nunca la clave en el código fuente) y rota la clave (crea una nueva, actualiza tus aplicaciones, borra la antigua) si sospechas que se ha filtrado.

Modelo mental

Piensa en ElevenLabs como en las cuerdas vocales del telefonista. El agente (Vapi o Retell) decide qué decir y cuándo; el LLM redacta las palabras exactas; pero el sonido final —el timbre, la entonación, el acento— depende de las cuerdas vocales que se le hayan asignado. ElevenLabs es el «taller» donde eliges, afinas y proteges esa voz antes de entregársela al resto del sistema.

Explicación

Según la documentación oficial de ElevenLabs verificada el 30 de agosto de 2026, la plataforma se organiza en torno a tres motores de texto a voz vigentes —Eleven v3, Multilingual v2 y Flash v2.5—, además de funciones adicionales como diálogo multihablante (Text to Dialogue), diseño de voz (Voice Design), cambio de voz sobre audio ya grabado (Voice Changer), efectos de sonido (Sound Effects v2) y generación musical (Eleven Music v2).

Elección de modelo según el caso de uso: Eleven v3 está recomendado quiere expresividad interpretativa y ofrece la cobertura de idiomas más amplia (74 idiomas, en una única versión multilingüe); Multilingual v2 está pensado para narración larga y pulida; Flash v2.5 está optimizado para volumen alto y uso en tiempo real, que es exactamente el escenario de un agente telefónico, donde la latencia importa tanto como la calidad.

Cobertura de español: tanto el español de España como el español de México figuran de forma explícita entre los idiomas soportados por los modelos multilingües de ElevenLabs, junto a otros más de veinte idiomas y variantes regionales (inglés en varias variantes, portugués de Brasil y Portugal, francés de Francia y Canadá, árabe de Arabia Saudí y EAU, entre otros). Al elegir una voz para un agente en España, confirma en la ficha de la voz qué variante de español tiene marcada, porque una voz etiquetada como español de México puede sonar notablemente distinta a la esperada por un cliente español.

Estructura del panel: el panel principal se organiza, de forma estable a través de las distintas versiones de interfaz, en una biblioteca de voces (propias, de la comunidad o clonadas), un generador de audio a partir de texto, un historial de generaciones anteriores, y una sección de administración con las claves de API del espacio de trabajo. La ubicación exacta de la sección de API varía según la versión del panel: a veces aparece como entrada independiente del menú principal, y a veces dentro de «Ajustes» o del menú de perfil.

Vocabulario nuevo

Crédito (ElevenLabs)

Piénsalo así: La unidad de consumo del contador, parecida a los minutos de una tarifa telefónica, pero repartida entre varios servicios distintos.

Unidad de consumo unificada de ElevenLabs. Un crédito equivale aproximadamente a un carácter de texto en TTS estándar; otros productos (transcripción, doblaje, agentes conversacionales) consumen créditos a tasas propias distintas.

Ejemplo: El plan Free incluye 10.000 créditos al mes, suficientes para unos 10 minutos de TTS con Multilingual v2.

Voice ID

Piénsalo así: El número de referencia interno de una voz concreta del catálogo, como el código de un artículo en un almacén.

Identificador único de una voz dentro de ElevenLabs, necesario para seleccionarla desde la API o desde una plataforma externa como Vapi o Retell.

Ejemplo: Al conectar tu voz de ElevenLabs con Vapi, se te pedirá el Voice ID exacto, no el nombre visible de la voz.

Voice Design

Piénsalo así: Describir de palabra a un actor de doblaje el tipo de voz que necesitas, en lugar de elegir una ya grabada.

Función de ElevenLabs que genera una voz sintética nueva a partir de una descripción textual de sus características, sin necesidad de una grabación de referencia.

Ejemplo: Describir «voz femenina adulta, cálida, ritmo pausado, acento de España» para generar una voz candidata.

Clonación de voz (Voice Cloning)

Piénsalo así: Grabar a una persona real para que el sistema aprenda a imitar su timbre con precisión.

Función que crea una voz sintética a partir de grabaciones de una voz real. Existen dos niveles: clonación instantánea (Instant Voice Cloning), disponible desde planes de pago básicos, y clonación profesional (Professional Voice Cloning, PVC), de mayor fidelidad y disponible desde el plan Creator.

Ejemplo: Un negocio que quiere que su agente telefónico use una voz reconocible ya asociada a su marca.

Token de un solo uso

Piénsalo así: Una entrada válida para un único uso, distinta de la llave maestra del

edificio.

Credencial temporal de corta duración que permite autenticar una petición a la API de ElevenLabs sin exponer la clave de API principal, pensada para código que se ejecuta en el lado del cliente (navegador).

Ejemplo: Un widget de voz embebido en una página web pública, que no debe llevar la clave de API completa.

Preparación

Ten a mano tu correo profesional, tu gestor de contraseñas y, si vas a gestionar claves de API, un lugar seguro para guardarlas (variables de entorno de tu proyecto, nunca un documento de texto sin cifrar).

Procedimiento paso a paso

1. **Contexto:** Vas a generar tu primera clave de API para conectar ElevenLabs con Vapi, Retell o cualquier script propio.

Acción de teclado: Dentro del panel de ElevenLabs, recorre encabezados con **H** hasta la sección de administración del espacio de trabajo, o busca «API Keys» dentro de «Ajustes»/«Settings». Activa el botón «Create API Key» (o equivalente) con **Enter**.

Respuesta esperada de NVDA: NVDA debería anunciar un texto similar a «Create API Key, botón» y, tras activarlo, un formulario con campos de nombre, alcance (scope) y cuota de créditos.

Qué significa: Estás creando una credencial nueva, no reutilizando la de otra persona del equipo: cada clave debe tener un propósito identificable.

Acción siguiente: Ponle un nombre descriptivo (por ejemplo, «Vapi — Cliente Fontanería Ejemplo»), restringe el alcance a los endpoints de TTS si no necesitas más, y guarda la clave generada en tu gestor de contraseñas inmediatamente: en la mayoría de plataformas, la clave completa solo se muestra una vez.

2. **Contexto:** Vas a elegir una voz en español de España desde la biblioteca de voces.

Acción de teclado: Recorre encabezados con **H** hasta «Voices»/«Biblioteca de

voces». Usa el campo de filtro de idioma con **Tab** y selecciona «Spanish (Spain)» o equivalente.

Respuesta esperada de NVDA: NVDA debería anunciar la lista filtrada de voces disponibles junto con su nombre y, en muchos casos, una breve descripción.

Qué significa: El filtro por idioma no garantiza que la voz suene con acento de España: algunas voces multilingües mantienen un acento neutro o de otra región aunque hablen español correctamente.

Acción siguiente: Reproduce la muestra de audio de dos o tres voces candidatas con **Enter** sobre el botón de reproducción, y anota el Voice ID de la que elijas finalmente.

3. **Contexto:** Vas a generar un audio de prueba desde el generador de texto a voz.

Acción de teclado: Recorre encabezados con **H** hasta «Text to Speech». Escribe o pega el texto de prueba en el campo principal con **Tab** y **Ctrl + V**, elige el modelo (Flash v2.5 para pruebas rápidas de latencia, Multilingual v2 para calidad de narración) y activa «Generate» con **Enter**.

Respuesta esperada de NVDA: NVDA debería anunciar el progreso de la generación y, al finalizar, un reproductor de audio con controles de reproducción y descarga.

Qué significa: El audio generado consume créditos de tu cuenta en el acto; revisa la longitud del texto antes de generar si tu plan tiene pocos créditos disponibles.

Acción siguiente: Descarga el audio o cópialo a tu historial de pruebas antes de cerrar la sesión.

Posibles diferencias de interfaz

ElevenLabs actualiza su interfaz con cierta frecuencia. Los conceptos (biblioteca de voces, generación, historial, claves de API, modelos) son estables aunque el nombre exacto de una sección o de un botón cambie ligeramente entre versiones del panel.

Problema de accesibilidad y alternativa

ElevenLabs ha realizado mejoras de accesibilidad documentadas en su propio registro de cambios (navegación por teclado en todos los elementos interactivos,

controles de la barra de progreso enfocables, mejor visibilidad del foco), pero, según el propio panel de comentarios oficial de ElevenLabs (feedback.elevenlabs.io), existen incidencias de accesibilidad activas y reconocidas por la propia plataforma: la sección de historial puede bloquear la navegación tras el botón «Manage History», impidiendo interactuar con elementos posteriores; algunos iconos del reproductor de audio (me gusta, no me gusta, descargar, minimizar) no se anuncian con una etiqueta comprensible con lector de pantalla; y varios botones de la función «Create Projects» (reproducir, reproducción continua, regenerar, insertar separador, encabezados, menú de descarga) carecen de una etiqueta accesible adecuada.

Qué hacer si te encuentras con alguna de estas barreras: usa la navegación por objetos o por elementos de formulario de NVDA (en lugar de solo Tab) para localizar botones sin etiqueta clara, ya que a menudo siguen siendo activables aunque su anuncio sea incompleto o críptico; si necesitas certeza total sobre qué botón estás activando, prueba primero en un audio de descarte antes de aplicar la acción a un proyecto importante; y, para las tareas críticas de generación de voz (texto a voz simple, selección de voz, gestión de claves de API), usa la API directamente mediante un script propio en lugar del panel visual, ya que la API no depende en absoluto de la accesibilidad de la interfaz web.

Errores frecuentes y diagnóstico

La generación de audio falla o se detiene sin descargar nada.

Cómo localizarlo: Comprueba, en el historial o en el mensaje de error mostrado, si el motivo es haber agotado los créditos del periodo (plan Free o de pago) o haber superado el límite de caracteres por petición del modelo elegido.

Solución: Si se han agotado los créditos, espera a la renovación mensual o cambia de plan; si el texto es demasiado largo para una única petición, divídelo en fragmentos más cortos y concaténalos después.

La voz suena con un acento distinto al esperado pese a haber filtrado por «Spanish (Spain)».

Cómo localizarlo: Algunas voces multilingües mantienen rasgos de acento de su locutor original, con independencia del idioma filtrado; el filtro de idioma indica qué idiomas puede pronunciar la voz, no garantiza un acento regional concreto.

Solución: Escucha la muestra de audio de cada voz candidata antes de elegirla para producción, y considera probar Voice Design describiendo explícitamente el acento deseado si ninguna voz de catálogo encaja.

No encuentras la sección de claves de API en el menú principal del panel.

Cómo localizarlo: En varias versiones de la interfaz, la gestión de claves de API está dentro de «Ajustes» o del menú de tu perfil de espacio de trabajo, no como sección independiente del menú principal.

Solución: Busca por el término «API» o «Developers» recorriendo encabezados y botones, o consulta directamente elevenlabs.io/docs/overview/administration/workspaces/api-keys.

Una petición a la API devuelve un error 403 (prohibido).

Cómo localizarlo: Suele deberse a una clave de API restringida por lista blanca de IP que no incluye la dirección desde la que estás realizando la petición, o a un alcance (scope) de la clave que no cubre el endpoint que estás llamando.

Solución: Revisa la configuración de la clave en el panel: amplía la lista de IP autorizadas o el alcance de endpoints si corresponde, o genera una clave nueva con los permisos correctos.

Una clave de API deja de funcionar de forma repentina, sin haberla borrado tú mismo.

Cómo localizarlo: Puede deberse a que la clave se filtró públicamente (por ejemplo, en un repositorio de GitHub) y ElevenLabs la desactivó automáticamente como medida de protección, o a que caducó si se creó con fecha de expiración.

Solución: Genera una clave nueva, actualiza inmediatamente todas las aplicaciones que dependían de la antigua, y revisa tu historial de commits o de mensajes para confirmar si la clave llegó a exponerse en algún sitio público.

Comprobación final

Confirma que puedes, recorriendo el panel de ElevenLabs solo con NVDA: generar un audio de prueba en español con la voz que hayas elegido, localizar el Voice ID de esa voz, y encontrar la sección donde se gestionan las claves de API. Si consigues las tres cosas, la comprobación es positiva.

Qué acabas de conseguir

Tienes una referencia completa a la que volver cada vez que necesites resolver una duda de ElevenLabs sin depender de memorizar toda la plataforma de antemano:

precios vigentes, elección de voz en español, gestión segura de claves de API y el catálogo de errores más frecuentes con su solución.

Ejercicio práctico

Genera el mismo texto de prueba con los tres modelos (Eleven v3, Multilingual v2 y Flash v2.5) y compara, de oído, la diferencia de expresividad y de tiempo de generación entre ellos. Anota cuál usarías para un agente telefónico en producción y por qué.

Resumen de lo imprescindible

- ElevenLabs es la pieza VOZ (TTS) del sistema: convierte en audio lo que decide el CEREBRO (LLM), sin decidir el contenido.
- El plan Free (10.000 créditos/mes, sin coste) no incluye uso comercial ni clonación de voz; los planes de pago van desde 6 USD/mes (Starter) hasta 990 USD/mes (Business), con Enterprise a medida.
- Español de España y español de México son variantes distintas dentro de los idiomas soportados: escucha siempre la muestra antes de elegir una voz para producción.
- La clave de API nunca debe exponerse en código de cliente; usa tokens de un solo uso para escenarios donde el navegador necesita autenticarse.
- ElevenLabs reconoce oficialmente incidencias de accesibilidad activas en su panel (historial, iconos del reproductor, botones de Create Projects); la API es la vía alternativa más fiable para tareas críticas.

Estoy preparado para continuar si puedo...

- Elegir una voz en español adecuada para un agente telefónico, verificando su acento por muestra de audio y no solo por el filtro de idioma.
- Generar y proteger correctamente una clave de API de ElevenLabs, con el alcance y la cuota adecuados.
- Diagnosticar los errores más frecuentes de generación de voz y de autenticación con la API.
- Aplicar una alternativa (navegación por objetos de NVDA, uso directo de la API) cuando el panel visual presente una barrera de accesibilidad conocida.

Fuentes

- [ElevenLabs Pricing — ElevenLabs \(consultado el 2026-08-30\)](#)
- [API Keys — ElevenLabs \(consultado el 2026-08-30\)](#)
- [Secure by design — ElevenLabs \(consultado el 2026-08-30\)](#)

- [Models — ElevenLabs \(consultado el 2026-08-30\)](#)
- [Voices — ElevenLabs \(consultado el 2026-08-30\)](#)
- [How do I authorize myself using an API key? — ElevenLabs \(centro de ayuda oficial\) \(consultado el 2026-08-30\)](#)
- [Changelog — mejoras de navegación por teclado y foco \(26 de mayo de 2025\) — ElevenLabs \(consultado el 2026-08-30\)](#)
- [Screen Reader/Accessibility Issues \(panel de comentarios oficial\) — ElevenLabs \(consultado el 2026-08-30\)](#)

Última actualización de este contenido: **2026-08-30**.