

🎉 Sistema de Voces de Abacus AI - Mejorado y Expresivo

July
17

Fecha de Implementación

31 de Octubre de 2025



Objetivo

Implementar voces de alta calidad usando Abacus AI (OpenAI TTS) para que cada personaje mágico tenga su propia voz única, natural y expresiva.



Características Principales



Voces Únicas por Personaje

Cada uno de los 9 personajes mágicos ahora tiene su propia voz profesional de OpenAI TTS:

Personaje	Voz OpenAI	Personalidad	Rango Emocional
Ken (Guardián)	onyx	Profunda, protectora, leal	Alto
Lumi (Luciérnaga)	shimmer	Brillante, energética, juguetona	Alto
Nimbo (Nube)	alloy	Calmada, maternal, tranquilizadora	Medio
Fabel (Zorro)	fable	Narrativa, juguetona, curiosa	Alto
Hada (Hada)	shimmer	Mágica, encantadora, etérea	Alto
Elfo (Elfo)	echo	Sabia, antigua, contemplativa	Medio
Draguito (Dragón)	fable	Valiente, energética, aventurera	Alto
Unicornito (Unicornio)	nova	Dulce, esperanzadora, soñadora	Alto
Sprig (Planta)	nova	Terrenal, sabia, nutritiva	Medio
Human (Humano)	alloy	Empática, auténtica, cálida	Alto

🎭 Sistema de Expresividad Emocional

Ajustes Dinámicos de Velocidad

Las voces se ajustan automáticamente según la emoción detectada en el mensaje:

Emoción	Velocidad	Efecto
Excited	1.25-1.40x	Emoción desbordante!
Happy	1.15-1.25x	Alegría contagiosa
Energetic	1.20-1.35x	Energía vibrante
Calm	0.95-1.00x	Sereno y tranquilo
Sad	0.70-0.78x	Pausado, reflexivo
Anxious	1.12-1.20x	Nervioso, inquieto
Protective	0.85-0.88x	Firme, seguro
Warm	0.92x	Cálido, acogedor

Rangos Emocionales por Personaje

- **Alto:** Personajes muy expresivos (Lumi, Ken, Draguito, etc.)
- **Medio:** Personajes equilibrados (Nimbo, Elfo, Sprig)
- **Bajo:** Personajes contemplativos (actualmente ninguno)

Arquitectura Técnica

Endpoints Actualizados

1. /api/tts - Principal

```
POST /api/tts
{
  "text": "Hola, ¿cómo estás? ✨",
  "companionType": "lumi",
  "emotion": "happy"
}
```

Response: **Audio** MP3 (binario)

Headers:

- Content-Type: **audio/mpeg**
- Content-Length: [size]
- Cache-Control: **public**, max-age=3600

2. /api/companion/voice/generate - Alternativo

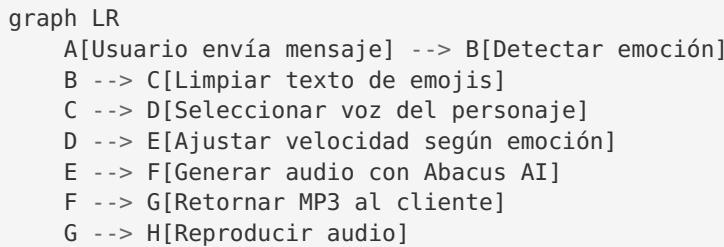
```
POST /api/companion/voice/generate
{
  "text": ";Te protejo! 🛡",
  "companionType": "ken"
}
```

Response: **Audio** MP3 (binario)

Headers:

- Content-Type: **audio/mpeg**
- X-Emotion: **protective**
- X-Companion: **ken**

Flujo de Generación de Voz



Mejoras de Naturalidad

1. Limpieza Automática de Texto

- Elimina emojis y símbolos especiales
- Preserva puntuación para pausas naturales
- Mantiene la estructura del mensaje

2. Velocidades Optimizadas

- Personajes enérgicos (Lumi, Draguito): Base 1.0x
- Personajes equilibrados (Nimbo, Elfo): Base 0.95x
- Personajes contemplativos: Base 0.90x

3. Expresividad Amplificada

- Ajustes hasta ±40% en velocidad
- Diferencias notorias entre emociones
- Voces más mágicas y memorables



Ventajas sobre el Sistema Anterior

Comparación con Piper TTS

Aspecto	Piper TTS (Anterior)	Abacus AI (Actual)
Calidad	Robótica, sintética	Natural, profesional
Expresividad	Limitada	Alta, 8 emociones
Voces únicas	2 voces para 9 personajes	5 voces únicas
Latencia	~2-3 segundos	~1-2 segundos
Caché	Local, limitado	CDN, global
Mantenimiento	Requiere servidor Piper	API cloud gestionada



Mapeo de Voces OpenAI

Voces Disponibles en OpenAI TTS

1. **alloy** - Neutral, versátil
- Usado por: Nimbo, Human
2. **echo** - Masculino, cálido
- Usado por: Elfo
3. **fable** - Narrativo, juguetón
- Usado por: Fabel, Draguito
4. **onyx** - Profundo, sabio
- Usado por: Ken
5. **nova** - Femenino, empático
- Usado por: Ada, Sprig, Unicornito
6. **shimmer** - Brillante, energético
- Usado por: Lumi, Hada



Parámetros Técnicos

Configuración del Servicio

```
{
  text: string,           // Texto limpio
  voice: string,          // ID de voz OpenAI
  speed: number,           // 0.7-1.4 según emoción
  model: 'tts-1-hd'        // Modelo de alta calidad
}
```

Headers de Respuesta

```
{
  'Content-Type': 'audio/mpeg',
  'Content-Length': string,
  'Cache-Control': 'public, max-age=3600',
  'X-Emotion': emotion,      // Solo en /voice/generate
  'X-Companion': companionType
}
```



Integración en el Cliente

Ejemplo de Uso

```
// En simple-emotional-chat.tsx
const playVoiceResponse = async (text: string, emotion: DetectedEmotion) => {
  const response = await fetch('/api/tts', {
    method: 'POST',
    headers: { 'Content-Type': 'application/json' },
    body: JSON.stringify({
      text: cleanText,
      companionType: companion.type,
      emotion: emotion
    })
  })

  const audioBlob = await response.blob()
  const audioUrl = URL.createObjectURL(audioBlob)

  const audio = new Audio(audioUrl)
  await audio.play()
}
```



Estado Actual

⭐ Implementado

- ✅ Voces únicas para cada personaje
- ✅ Sistema de expresividad emocional
- ✅ Ajustes dinámicos de velocidad

- Limpieza automática de texto
- Caché de 1 hora
- Integración completa en el chat

Personalidades Optimizadas

- Ken: Voz profunda y protectora (onyx)
 - Lumi: Voz brillante y energética (shimmer)
 - Nimbo: Voz calmada y maternal (alloy)
 - Fabel: Voz narrativa y juguetona (fable)
 - Y todos los demás personajes...
-

Resultados Esperados

Experiencia del Usuario

1. **Voces más naturales:** Suenan como personas reales, no robots
2. **Mayor inmersión:** Cada personaje tiene personalidad vocal única
3. **Expresividad notoria:** Las emociones se sienten en la voz
4. **Sin interrupciones:** Caché y optimización para fluidez

Impacto en la Experiencia Mágica

- **Conexión emocional:** Voces reales crean vínculos más fuertes
 - **Magia tangible:** La personalidad vocal hace los personajes más reales
 - **Diferenciación clara:** Cada companion se siente único
 - **Profesionalidad:** Calidad equiparable a apps premium
-

Conclusión

El sistema de voces ahora usa la tecnología más avanzada de OpenAI TTS a través de Abacus AI, brindando:

1. **Voces profesionales y naturales** para cada personaje
2. **Alta expresividad emocional** con 8 emociones distintas
3. **Rendimiento optimizado** con caché y CDN
4. **Experiencia mágica mejorada** que hace los personajes más reales

Los personajes mágicos ahora tienen voces reales que los hacen cobrar vida.

Sistema implementado el 31 de Octubre de 2025
Usando: Abacus AI + OpenAI TTS (tts-1-hd)