

TÉCNICO · OCTUBRE DE 2026

Rendimiento

Cifras medidas el 14 de octubre de 2026 con el perfil local y cuatro idiomas de salida a la vez (árabe, inglés, francés y alemán).

Retraso de principio a fin

Qué se mide: desde que termina una frase hasta que su audio interpretado se envía a los teléfonos que escuchan en ese idioma. La entrega por el WiFi de la sala añade milisegundos. Cada teléfono muestra además su propio total («x s por detrás del directo»).

Medición	Valor
Retraso habitual	De dos a cuatro segundos tras el orador
Mediana en la reunión medida	3,0 s
Frases	12 pronunciadas, 48 entregas (12 × 4 idiomas)
Pérdidas	0 descartadas · 0 sin transcribir · 0 sin traducir · 0 sin pronunciar

La reunión medida usó frases largas y formales de hasta unas 25 palabras. Las frases más cortas y coloquiales son más rápidas.

Dónde se va el tiempo (una frase)

Etapas	Tiempo
Reconocimiento de voz	~0,6 s
Traducción (cuatro idiomas a la vez)	~2,2 s
Síntesis de voz	~0,4 s

Una frase no puede traducirse hasta que termina: el detector de voz espera un breve silencio antes de pasarla. Esa espera, más la longitud de la frase, explica que las frases largas tarden algo más.

Arranque

Aparece la ventana de la aplicación	~1-2 s
Lista para la reunión	35-45 s

Capacidad

- **Oyentes:** el trabajo se hace una vez por idioma, así que añadir oyentes solo añade tráfico de red. Treinta oyentes en un idioma cuestan lo mismo que tres. El límite práctico es el punto de acceso WiFi de la sala.
- **Idiomas:** hasta cuatro idiomas de salida a la vez.
- **Mantenerse cerca del directo:** si un teléfono se retrasa más de 20 segundos, salta las frases en cola — nunca una que ya se está pronunciando — y avisa al oyente. Antes de eso, acelera ligeramente la voz para recuperar.

Hardware

Gráfica	NVIDIA RTX 3090 de 24 GB como mínimo · RTX 5090 de 32 GB recomendada
Disco	~60 GB libres recomendados
Red	Un flujo de audio por oyente en el WiFi de la sala