

TÉCNICO · OUTUBRO DE 2026

# Desempenho

Valores medidos a 14 de outubro de 2026 no perfil local, com quatro idiomas de saída ao mesmo tempo (árabe, inglês, francês e alemão).

## Atraso de ponta a ponta

*O que é medido:* desde o momento em que uma frase termina até o áudio interpretado ser enviado para os telemóveis que ouvem nesse idioma. A entrega pelo WiFi da sala acrescenta milissegundos. Cada telemóvel mostra também o seu próprio total («x s atrás do direto»).

| Medição                   | Valor                                                                |
|---------------------------|----------------------------------------------------------------------|
| Atraso habitual           | Dois a quatro segundos depois do orador                              |
| Mediana na reunião medida | 3,0 s                                                                |
| Frases                    | 12 ditas, 48 entregas (12 × 4 idiomas)                               |
| Perdas                    | 0 descartadas · 0 não transcritas · 0 não traduzidas · 0 nunca ditas |

A reunião medida usou frases longas e formais com até cerca de 25 palavras. Frases mais curtas e coloquiais são mais rápidas.

## Para onde vai o tempo (uma frase)

| Etapas                                   | Tempo  |
|------------------------------------------|--------|
| Reconhecimento de voz                    | ~0,6 s |
| Tradução (quatro idiomas ao mesmo tempo) | ~2,2 s |
| Síntese de voz                           | ~0,4 s |

Uma frase não pode ser traduzida antes de terminar: o detetor de voz espera por um breve silêncio antes de a passar. Essa espera, somada ao comprimento da frase, explica porque as frases longas demoram um pouco mais.

### Arranque

|                               |                |
|-------------------------------|----------------|
| A janela da aplicação aparece | ~1-2 s         |
| Pronto para a reunião         | <b>35-45 s</b> |

### Capacidade

- **Ouvintes:** o trabalho é feito uma vez por idioma, por isso mais ouvintes só acrescentam tráfego de rede. Trinta ouvintes num idioma custam o mesmo que três. O limite prático é o ponto de acesso WiFi da sala.
- **Idiomas:** até quatro idiomas de saída ao mesmo tempo.
- **Ficar perto do direto:** se um telemóvel ficar mais de 20 segundos atrás, salta as frases em fila — nunca uma que já esteja a ser dita — e avisa o ouvinte. Antes disso, acelera ligeiramente a voz para recuperar.

### Hardware

|               |                                                                    |
|---------------|--------------------------------------------------------------------|
| Placa gráfica | NVIDIA RTX 3090 de 24 GB no mínimo · RTX 5090 de 32 GB recomendada |
| Disco         | ~60 GB livres recomendados                                         |
| Rede          | Um fluxo de áudio por ouvinte no WiFi da sala                      |