

200 milisegundos

La espera es lo menos humano del servicio

Ernesto Matamoros

Director de Ventas · Omilia Natural Language Solutions

La pregunta que nos sigue reuniendo

¿Cómo hacemos que la atención
al cliente se sienta más humana?

Llevamos décadas respondiéndola con guiones, con capacitación, con encuestas de satisfacción.

La respuesta también es un número.

200

milisegundos

Es lo que tarda una persona en tomar el turno de palabra en una conversación.

El estudio comparó diez lenguas de todo el mundo y encontró el mismo patrón en todas: evitamos hablar encima y evitamos el silencio.

El presupuesto de la naturalidad

< 300 ms

Se percibe natural

La conversación fluye

300 – 500 ms

Aceptable

Se nota, pero no molesta

> 500 ms

Se rompe

El interlocutor repite o abandona

“Que se sienta humano” deja de ser una aspiración y se vuelve una meta de ingeniería medible.

Crterios de referencia de uso corriente en la industria de voice AI. No provienen de investigación revisada por pares.

PRIMERA PARTE

**Lo menos humano del servicio
no es la máquina. Es la espera.**

Medimos una espera. Son dos.

01

La cola

Todo el mundo la mide. Tiene indicador, tablero y meta mensual.

02

El interrogatorio

Número de cuenta, fecha de nacimiento, los últimos cuatro dígitos. Casi nadie la mide por separado.

El cliente ya esperó en la cola. Ahora tiene que probar quién es, antes de poder decir para qué llamó.

El único control de seguridad que resta fricción

Normalmente seguridad y experiencia se pelean: cada capa de protección agrega pasos.

**La identidad puede viajar en el mismo audio que ya trae la petición.
Cero turnos adicionales.**

Y la contraparte honesta:

- Clonar una voz hoy cuesta muy poco. La biometría sola ya no alcanza.
- Hace falta detección de ataques de presentación (PAD) — norma ISO/IEC 30107-3.
- El NIST es explícito: un biométrico no es autenticador por sí solo.
- Es una defensa que se reentrena, no una casilla que se marca una vez.

SEGUNDA PARTE

**Si el problema está claro,
¿por qué sigue sin resolverse?**

El dato incómodo

+40%

de los proyectos de IA agéntica se
cancelarán antes de terminar 2027

Las tres causas que señala Gartner:

Costos crecientes

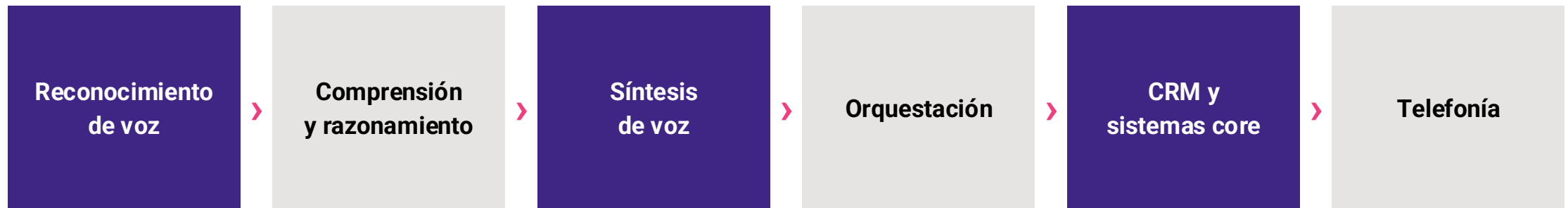
Valor de negocio poco claro

Controles de riesgo insuficientes

Ninguna de las tres es un problema de modelo. Las tres son de arquitectura y de gobierno.

Una lectura posible: cada frontera cuesta

Una conversación atraviesa varios proveedores antes de responder. Cada salto consume parte del presupuesto y diluye la responsabilidad.



La pregunta operativa no es cuántos proveedores tengo, sino quién responde cuando la conversación se rompe.

Diagnóstico para mañana

Cuatro preguntas que cualquiera puede aplicar a su propia operación:

01

¿Cuántas fronteras entre proveedores atraviesa una conversación?

02

¿Cuánto tarda mi autenticación, medida como indicador propio?

03

¿Quién es el dueño del SLA de la conversación completa?

04

¿Qué indicadores dejan de servirme si la capacidad deja de ser escasa?

TERCERA PARTE

**Y hacia 2030,
dos cosas cambian de raíz.**

Uno: el tablero pierde sentido

Nivel de servicio, tiempo medio de operación y tasa de abandono existen porque el tiempo de agente era escaso.

Escasez

Métricas de racionamiento:
cuánto esperó, cuánto duró,
cuántos colgaron.

Capacidad elástica

Métricas de resolución:
¿quedó resuelto, a la primera,
sin que volviera a llamar?

La pregunta para esta industria: ¿qué van a medir cuando el tiempo deje de ser el recurso escaso?

Dos: quien llama puede no ser una persona

El 19 de mayo de 2026 Google anunció que un usuario puede pedirle que llame a un negocio en su nombre. Despliegue inicial en Estados Unidos.

Lo que eso abre para un centro de contacto:

- Hoy verificamos que quien llama es nuestro cliente.
- Mañana habrá que verificar que el agente que llama representa legítimamente a ese cliente.
- El centro de contacto se vuelve punto de control de identidad también para los no humanos.

¿Su operación está lista para atender al agente de su cliente?

La división que importa

La automatización se queda con la espera

Nadie renunció a un empleo para leerle un menú de opciones a alguien.

Las personas se quedan con el juicio

91% de los líderes de CX espera que los agentes humanos sigan siendo críticos dentro de tres años.

Automatizar la espera no es reemplazar a nadie. Es dejar de gastar personas en tareas que nunca necesitaron una.

**“Mi madre no va a poder
en su vida resolver algo
con su banco apretando
botones en un teléfono.”**

Dimitris Vassos, fundador de Omilia

200

milisegundos

**No necesitaba un menú mejor.
Necesitaba poder hablar.**

Y hablar tiene un reloj.

Que la próxima vez que alguien necesite ser atendido, sea de inmediato,
y su duda quede resuelta en poco tiempo.

Fuentes

- Stivers et al., «Universals and cultural variation in turn-taking in conversation», PNAS 106(26), 2009.
- NIST SP 800-63B, Digital Identity Guidelines, §5.2.3 — uso de biométricos y resistencia a ataques de presentación.
- ISO/IEC 30107-3, Biometric presentation attack detection — Part 3: Testing and reporting.
- Gartner, comunicado de prensa del 25 de junio de 2025 sobre cancelación de proyectos de IA agéntica.
- Genesys, The State of Customer Experience 2026 (quinta edición).
- Google, «A new era for AI Search», blog oficial, 19 de mayo de 2026.
- Cita de Dimitris Vassos: Expedition Growth Capital, «Omilia and Expedition's partnership», 25 de julio de 2022.

Las cifras citadas provienen de las fuentes primarias listadas. Ninguna lámina contiene datos de clientes.