

Casos límite: preparar al agente para lo inesperado

Datos de esta lección

Fase	F4 — Diagnóstico, calidad y contingencias
Módulo	M4.2 — Casos límite y transferencias
Puntos de profesionalidad al superar el test	45

Objetivo práctico

Al terminar esta lección sabrás preparar y probar deliberadamente tu agente frente a veinte situaciones difíciles habituales en una llamada real, y reforzar el Prompt Maestro para que responda con seguridad ante cada una, sin inventar información.

Resultado que obtendrás

Una batería de pruebas de casos límite ya ejecutada contra tu propio agente, con las mejoras de prompt correspondientes aplicadas donde hiciera falta.

Dónde encaja dentro del sistema final

Esta lección refuerza directamente la pieza CEREBRO del sistema (el prompt), preparando al agente para el «Protocolo final de pruebas de estrés» de la Fase 6, que exige exactamente este tipo de casos.

Conocimientos previos

- Construcción del Prompt Maestro del recepcionista (módulo 1.3).
- El hito de Tool Calling ya completado (módulo 3.4).

Herramientas necesarias

- Tu agente ya funcional (Fases 2 y 3).
- Un navegador con NVDA activo.

Posible coste

Cada llamada de prueba adicional tiene el coste habitual explicado en el módulo

Modelo mental

Piensa en la formación de un telefonista humano nuevo: no basta con enseñarle el guion de una llamada perfecta; hace falta prepararlo para el cliente que se enfada, la línea con ruido de fondo, o la pregunta que nadie esperaba. Esta lección es exactamente esa formación, aplicada al agente.

Explicación

Un agente probado únicamente con conversaciones «perfectas» puede funcionar de maravilla en una demostración y fallar en la primera llamada real. Esta lección organiza veinte situaciones límite en cuatro familias:

Problemas de la señal de voz: silencio prolongado, ruido de fondo, eco, interrupciones (el cliente habla antes de que el agente termine), acentos marcados, necesidad de repetición.

Problemas de contenido de la conversación: preguntas fuera de contexto, cliente enfadado, urgencia real, palabra desconocida o mal transcrita, información que el agente no tiene, llamadas anormalmente largas, cambio de idioma a mitad de conversación.

Problemas de datos ambiguos: fecha ambigua («el jueves que viene» sin especificar cuál si hay dos jueves razonables), hora ambigua («por la tarde»), correo electrónico difícil de dictar, nombre propio poco común o difícil de transcribir.

Problemas de gestión de la cita: cancelación, modificación de una cita ya reservada, solicitud explícita de hablar con una persona.

La regla de oro para todos los casos es la misma que aprendiste en el módulo 1.3: el agente debe reconocer explícitamente que no sabe algo, en lugar de inventar una respuesta. Un agente que dice «no tengo ese dato, pero puedo transferirte con una persona» es mucho más profesional que uno que inventa una respuesta plausible pero falsa.

Preparación

Ten a mano tu Prompt Maestro (módulo 1.3) para poder editarlo si hace falta, y descarga la «Hoja de pruebas» de Recursos si ya está disponible en tu sesión de trabajo.

Procedimiento paso a paso

1. **Contexto:** Vas a probar deliberadamente un caso de fecha ambigua.

Acción de teclado: Realiza una llamada de prueba y pide una cita diciendo «para el jueves que viene» sin más precisión, en un momento en que esa expresión sea razonablemente ambigua.

Respuesta esperada de NVDA: No aplica directamente; escucha la respuesta del agente.

Qué significa: Un agente bien preparado debería pedir una confirmación explícita de la fecha exacta (por ejemplo, «¿te refieres al jueves día 15?») en lugar de asumir una fecha sin confirmar.

Acción siguiente: Si el agente no pide esa confirmación, añade una regla explícita al Prompt Maestro: «Ante cualquier fecha u hora ambigua, confirma la fecha exacta en formato día/mes antes de continuar».

2. **Contexto:** Vas a probar un caso de información ausente.

Acción de teclado: Realiza otra llamada de prueba y pregunta al agente algo que, deliberadamente, no está en su información (por ejemplo, un servicio que el negocio no ofrece).

Respuesta esperada de NVDA: No aplica directamente; escucha la respuesta del agente.

Qué significa: Un agente bien preparado debería decir explícitamente que no dispone de esa información, sin inventar una respuesta plausible.

Acción siguiente: Si el agente inventa una respuesta, refuerza en el Prompt Maestro la prohibición explícita de inventar datos, con un ejemplo concreto similar al que acabas de probar.

Errores frecuentes y diagnóstico

El agente inventa una respuesta convincente pero falsa ante una pregunta fuera de su información.

Cómo localizarlo: Es una alucinación (módulo 1.2), casi siempre resultado de un prompt sin una prohibición suficientemente explícita.

Solución: Refuerza el prompt con una regla explícita y un ejemplo concreto del tipo de pregunta que debe rechazar con honestidad.

Comprobación final

Prueba al menos cinco de los veinte casos límite de esta lección contra tu propio agente, anota el resultado de cada uno, y ajusta el prompt donde haga falta. Si consigues que el agente maneje correctamente los cinco casos probados, la comprobación es positiva.

Qué acabas de conseguir

Tienes un agente probado deliberadamente contra situaciones difíciles reales, no sólo contra conversaciones perfectas, preparándote para el módulo 4.2 (transferencias) y el protocolo de pruebas de estrés de la Fase 6.

Ejercicio práctico

Elige cinco casos límite más de la lista de esta lección (distintos de los que ya probaste) y repite el proceso completo: probar, observar, y reforzar el prompt si hace falta.

Resumen de lo imprescindible

- Un agente probado sólo con conversaciones perfectas puede fallar en la primera llamada real.
- Los veinte casos límite de esta lección se agrupan en cuatro familias: señal de voz, contenido de la conversación, datos ambiguos y gestión de la cita.
- La regla de oro es siempre la misma: reconocer no saber algo, en lugar de inventar una respuesta.
- Cada caso límite fallido se corrige reforzando el Prompt Maestro con una regla explícita y un ejemplo concreto.

No necesitas aprender todavía

Todavía no necesitas configurar la transferencia real de llamadas: es el contenido de la siguiente lección.

Estoy preparado para continuar si puedo...

- Probar deliberadamente un agente contra casos límite reales, no sólo contra conversaciones perfectas.
- Reconocer una alucinación del agente y corregirla reforzando el prompt.
- Explicar por qué reconocer no saber algo es preferible a inventar una respuesta plausible.

Test de comprobación

Para registrar el resultado y obtener Puntos de profesionalidad, realiza este test desde la versión web de la lección. Las opciones se muestran aquí sin señalar la respuesta correcta.

1. Un cliente pide una cita «para el jueves que viene» y el agente reserva automáticamente la fecha más próxima sin confirmar nada, aunque hay ambigüedad razonable. ¿Qué debería haber hecho el agente, según esta lección?
 - a. Confirmar explícitamente la fecha exacta antes de proceder con la reserva.
 - b. Colgar la llamada inmediatamente ante cualquier ambigüedad.
 - c. Transferir automáticamente cualquier llamada con una fecha mencionada.
2. Un cliente pregunta por un servicio que el negocio no ofrece, y el agente responde con un precio y una disponibilidad inventados. ¿Qué está ocurriendo, y cómo se corrige?
 - a. Es una alucinación; se corrige reforzando el Prompt Maestro con una prohibición explícita de inventar datos y un ejemplo concreto.
 - b. Es un fallo del STT; se corrige revisando la transcripción de la llamada.
 - c. Es un fallo de Twilio; se corrige revisando el número de teléfono.
3. ¿Por qué esta lección insiste en probar el agente contra casos límite antes de considerarlo terminado?
 - a. Porque un agente que sólo se ha probado con conversaciones perfectas puede comportarse de forma poco fiable en la primera situación real inesperada.
 - b. Porque es un requisito legal obligatorio en España para cualquier agente de IA.
 - c. Porque las plataformas de voz cobran menos si se prueban casos límite.

Fuentes

- [Voice AI Prompting Guide — Vapi \(consultado el 2026-08-30\)](#)
- [Prompt engineering — OpenAI \(consultado el 2026-08-30\)](#)

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