



Universidad
de La Laguna

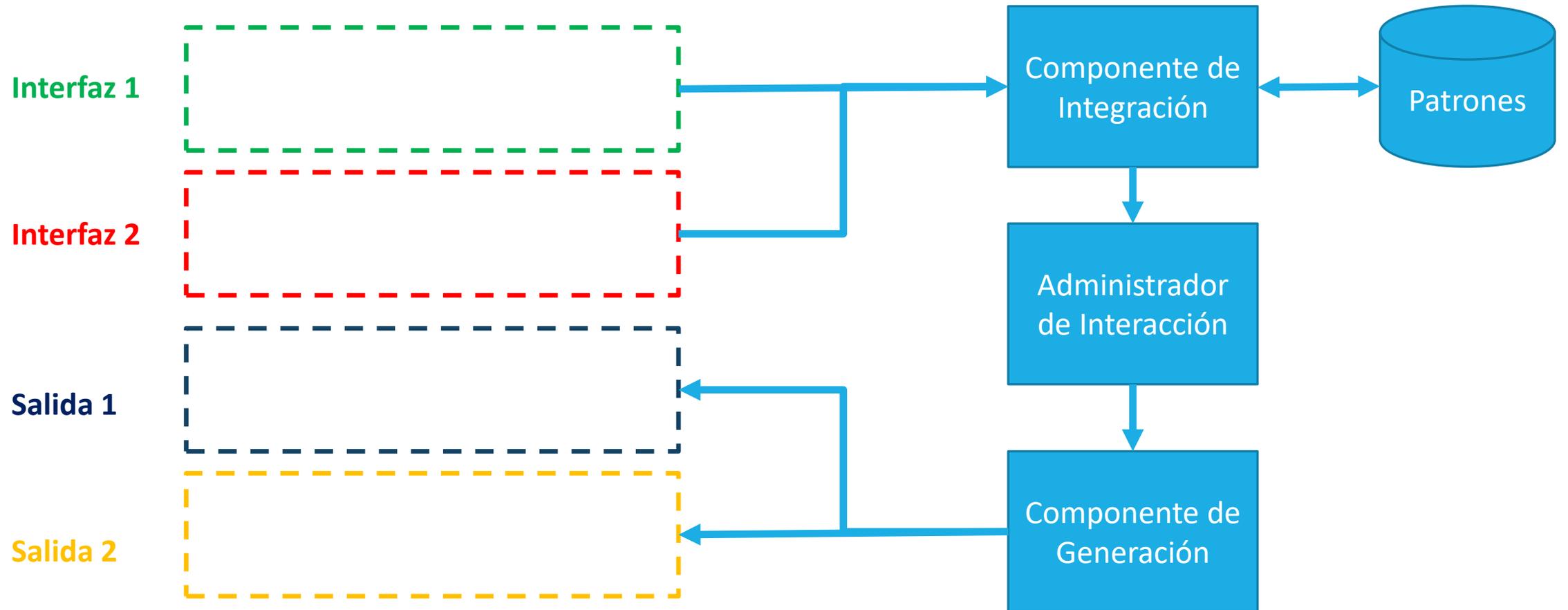
Interfaces Multimodales

INTERFACES INTELIGENTES

Multimodalidad (W3C)

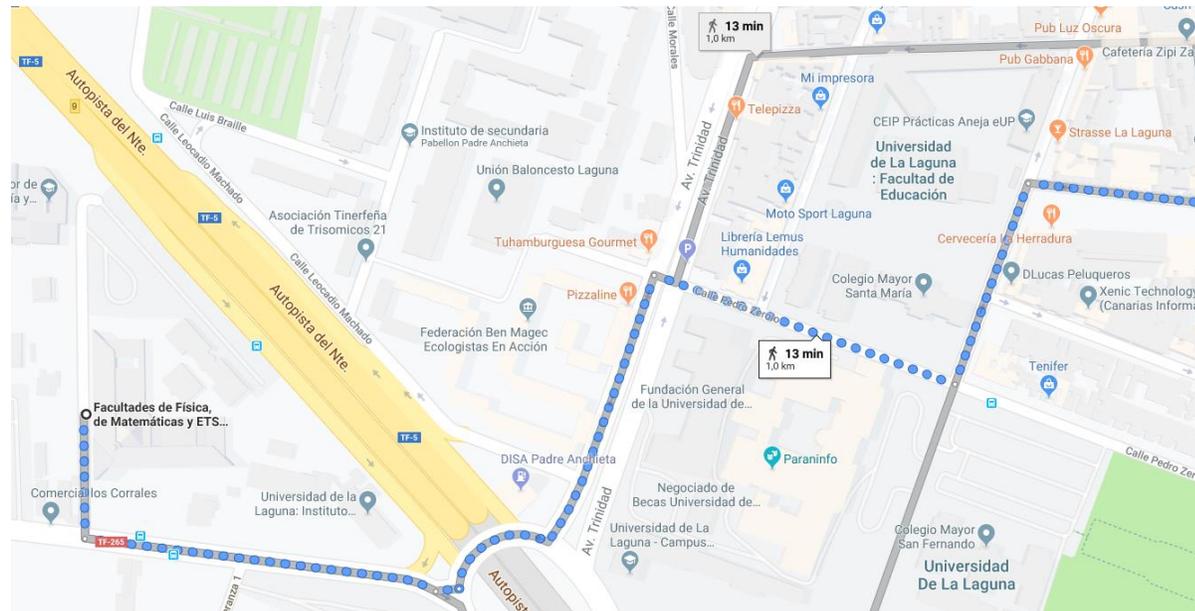
- *Proceso en el cual diversos dispositivos y personas son capaces de llevar a cabo una interacción conjunta desde cualquier sitio, en cualquier momento, utilizando cualquier dispositivo y de forma accesible, incrementando así la interacción entre personas, y entre dispositivos y personas.*
- **En la práctica:** La capacidad de comunicarse entre personas y dispositivos de mas de una forma al mismo tiempo.

Multimodalidad (W3C)

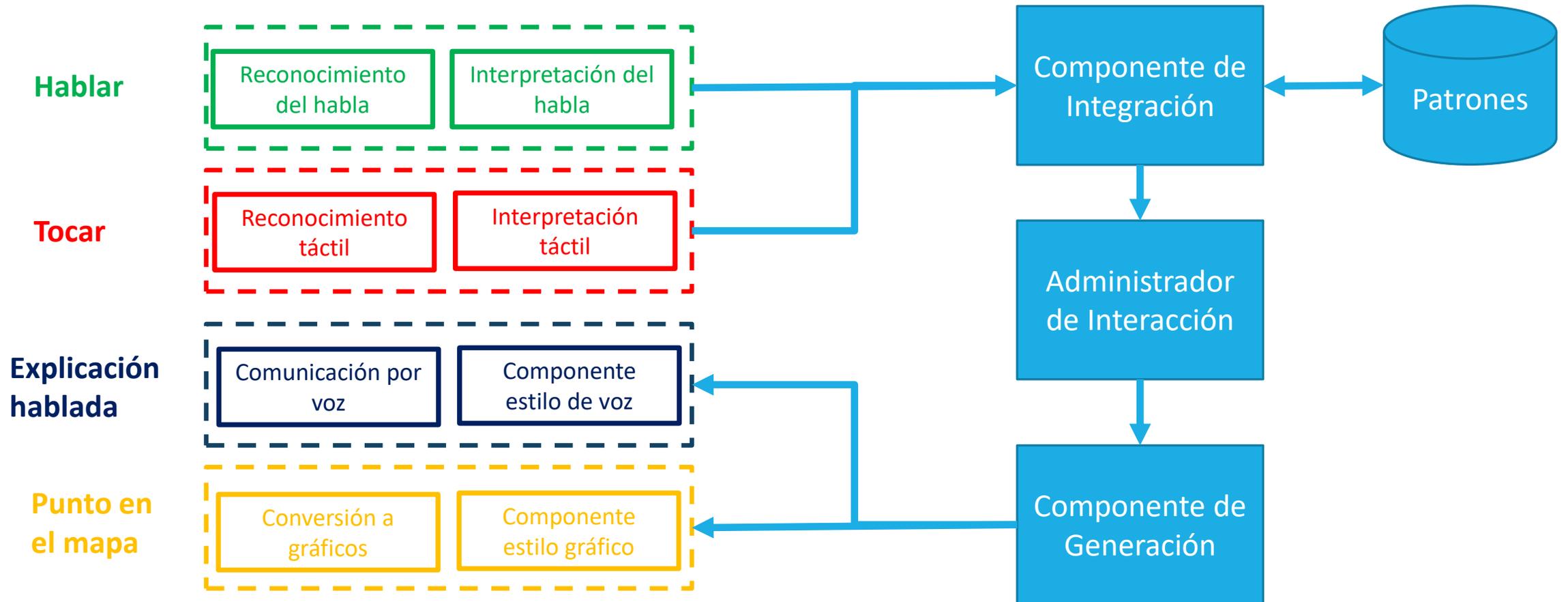


Multimodalidad. Ejemplo

- **Mapa Turístico Interactivo:** El usuario consulta un puesto turístico con mapa.
- **Entradas:** Táctil y Voz
- **Salida:** Punto en el mapa y descripción hablada



Multimodalidad. Ejemplo



Multimodalidad. Ejemplo

- **Usuario:** Señala un lugar en el mapa y dice, "¿Cómo se llama este sitio?".
- **Componente de reconocimiento del habla:** Reconoce las palabras "¿Cómo se llama este sitio?".
- **Componente de reconocimiento táctil:** Reconoce las coordenadas x-y que el usuario ha señalado en el mapa.

Multimodalidad. Ejemplo

- **Componente de interpretación del habla:** Convierte las palabras "¿Cómo se llama este sitio?" en una notación interna.
- **Componente de interpretación de indicación:** Convierte las coordenadas x-y del lugar indicado por el usuario en una notación interna.
- **Componente de integración:** Integra la notación interna de las palabras "¿Cómo se llama este sitio?", con la notación interna de las coordenadas x-y.

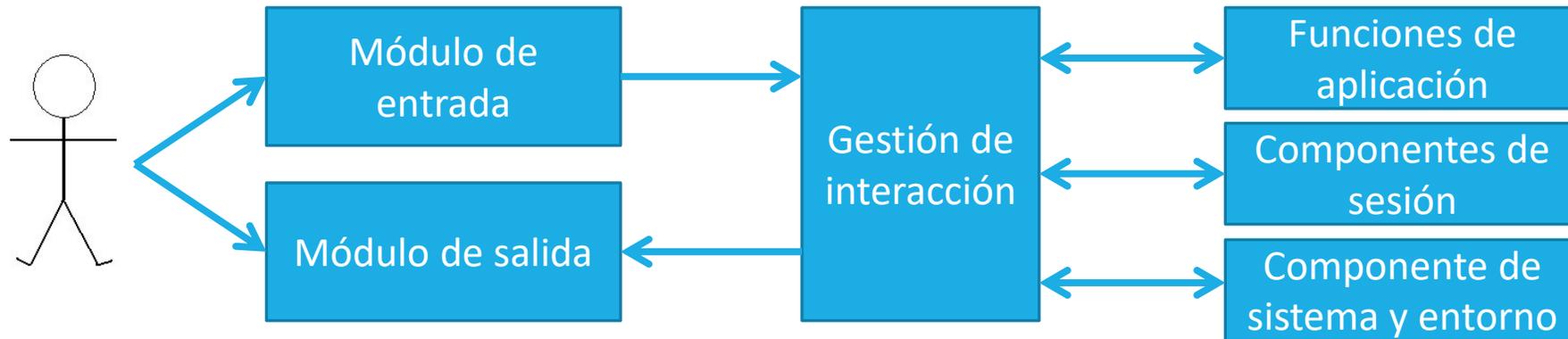
Multimodalidad. Ejemplo

- **Componente del administrador de la interacción:** Utiliza la notación interna para realizar un procesamiento y obtener una conclusión en notación interna. El administrador de la interacción convierte la respuesta en una notación interna y envía la respuesta al componente de generación.
- **Componente de generación:** Procesa el resultado en notación interna, obtiene las posibles salidas y los datos a los componentes de las distintas salidas.

Multimodalidad. Ejemplo

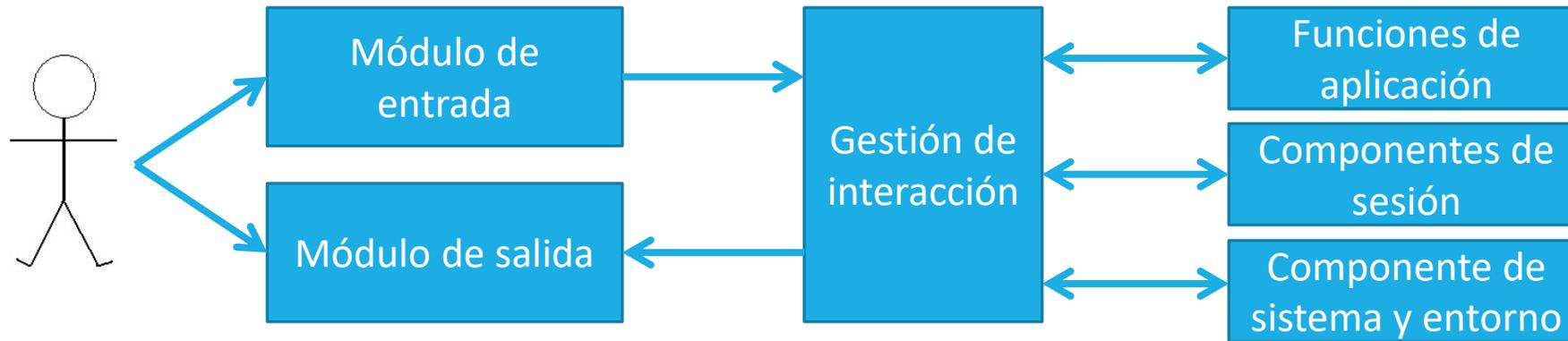
- **Componente de estilo de voz:** Convierte en SSML la notación interna que representa el resultado.
- **Componente de estilo de gráficos:** Convierte la notación interna que representa la ubicación en un mapa, usando notación HTML.
- **Componente de conversión de voz:** Convierte la notación SSML en voz acústica para que el usuario pueda oírla.
- **Componente de estilo de gráficos:** Convierte la notación HTML en gráficos visuales para que el usuario pueda verlos.

Multimodalidad. Arquitectura



- **Usuario:** Interacciona con el sistema a través de eventos y acciones. Recibe información a través de los diferentes modos de presentación.
- **Módulo de entrada:** Presenta al usuario los modos posibles de acción hacia el sistema (voz, escritura, puntero, etc.)
- **Módulo de salida:** El usuario recibe información en formato de texto, gráficos, voz, audio, animación, etc.

Multimodalidad. Arquitectura



- **Gestor de interacción:** Coordina la información y gestiona el flujo de ejecución a partir de los componentes de los objetos de interfaz asociados a los diferentes modos de entrada y salida.
- **Componente de sistema y entorno:** Permite que el gestor de interacción conozca y responda a los cambios que se producen en las capacidades y propiedades del terminal o dispositivo, a las preferencias del usuarios, o a las condiciones ambientales.
- **Componente de sesión:** Proporciona una interfaz al gestor de interacción, para soportar la gestión de estados, así como las sesiones temporales y persistentes en las aplicaciones multimodales. Necesario para escenarios de operación como son las aplicaciones que se ejecuten en varios terminales o necesiten de sesión (juegos multijugador, chats multimodales, etc.)

Interfaz multimodal. Coordinación

- Una interfaz multimodal proporciona una interfaz de usuario común dentro de un entorno heterogéneo.
- Se requiere una coordinación o un sistema de fusión de entrada para sincronizar los datos.
 - Se debe conseguir una toma de decisión ante datos confusos.
- El control de interacción multimodal exige desarrollar nuevos paradigmas de gestión de la interacción.

Interfaz multimodal. Coordinación

- En la multimodalidad se establecen establece tres niveles de sincronismo:
 - **Multimodalidad secuencial:** Cada interacción sólo es posible un modo, no varios simultáneos.
 - **Multimodalidad simultánea no coordinada:** Contempla la activación de más de una modalidad por interacción, pero no contempla un procesado integrado, sino un procesado independiente en orden aleatorio.
 - **Multimodalidad simultánea coordinada:** Contempla la activación simultánea de varios modos y un procesado complementario de información procedente de diferentes modos. Exige establecer procesos de sincronización complejos.

Interfaz multimodal. Coordinación

- Interfaces que permiten la entrada de información a los sistemas de cómputo de diferentes formas:
 - **Natural:**
 - Temperatura
 - Localización
 - Humedad
 - Habla
 - Gestos
 - Expresiones faciales (detección de sentimientos)
 - **Múltiple:**
 - Combinación de elementos naturales (auditivo, táctil, gestual, etc.)

Interfaz multimodal. Ventajas

- Mayor posibilidad de recuperación de errores:
 - Mas modos de obtener información, menos errores
- Se realiza una interacción mas simple

Interfaz multimodal. Desventajas

- Necesidades de cómputo elevadas que demandan algunos de los modos de interacción.
- Dificultad de gestionar la interacción multimodal.
- La complejidad de integrar las tecnologías que dan soporte a los diferentes modos de interacción con el usuario (reconocedor de habla, reconocedor de idioma, conversor texto-voz, etc.)
- La fuerte presión actual en innovación tecnológica que exige disponer de resultados industrializables a corto plazo.