

Departamento de Computación, FCEyN, UBA

Procesamiento del Habla

Agustín Gravano

1er Cuatrimestre 2017

Sistemas de diálogo hablado



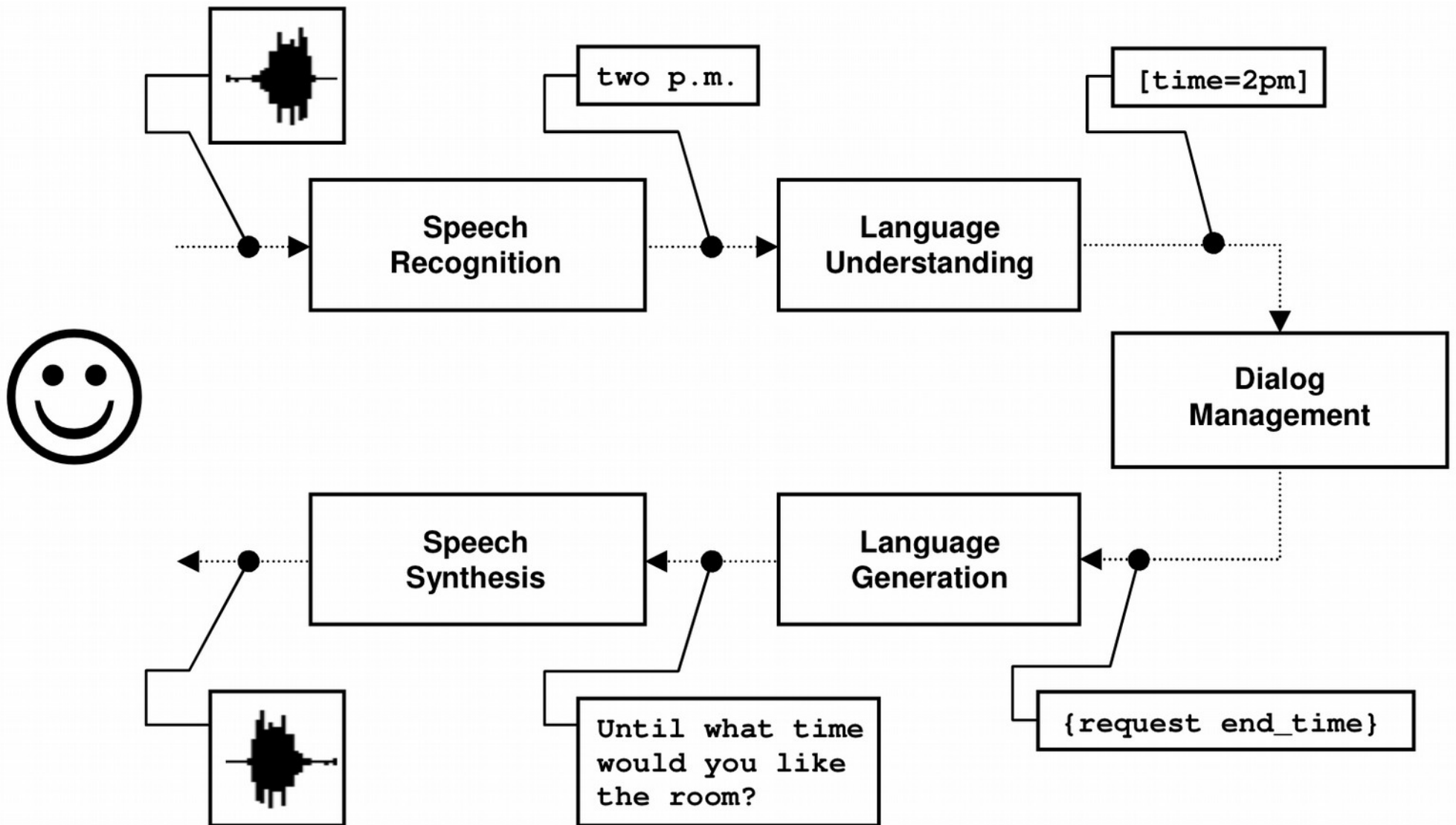
Sistemas de Diálogo Hablado (SDH)

- También conocidos como:
 - *Interactive Voice Response Systems (IVR Systems)*
 - *Spoken dialogue systems (SDS)*
 - Agentes Conversacionales
- Aplicaciones:
 - Preparación de viajes (Amtrak, United Airlines)
 - Ruteo de llamadas telefónicas
 - Enseñanza
 - Comunicación con robots
 - Cualquier dispositivo con teclado o pantalla limitados

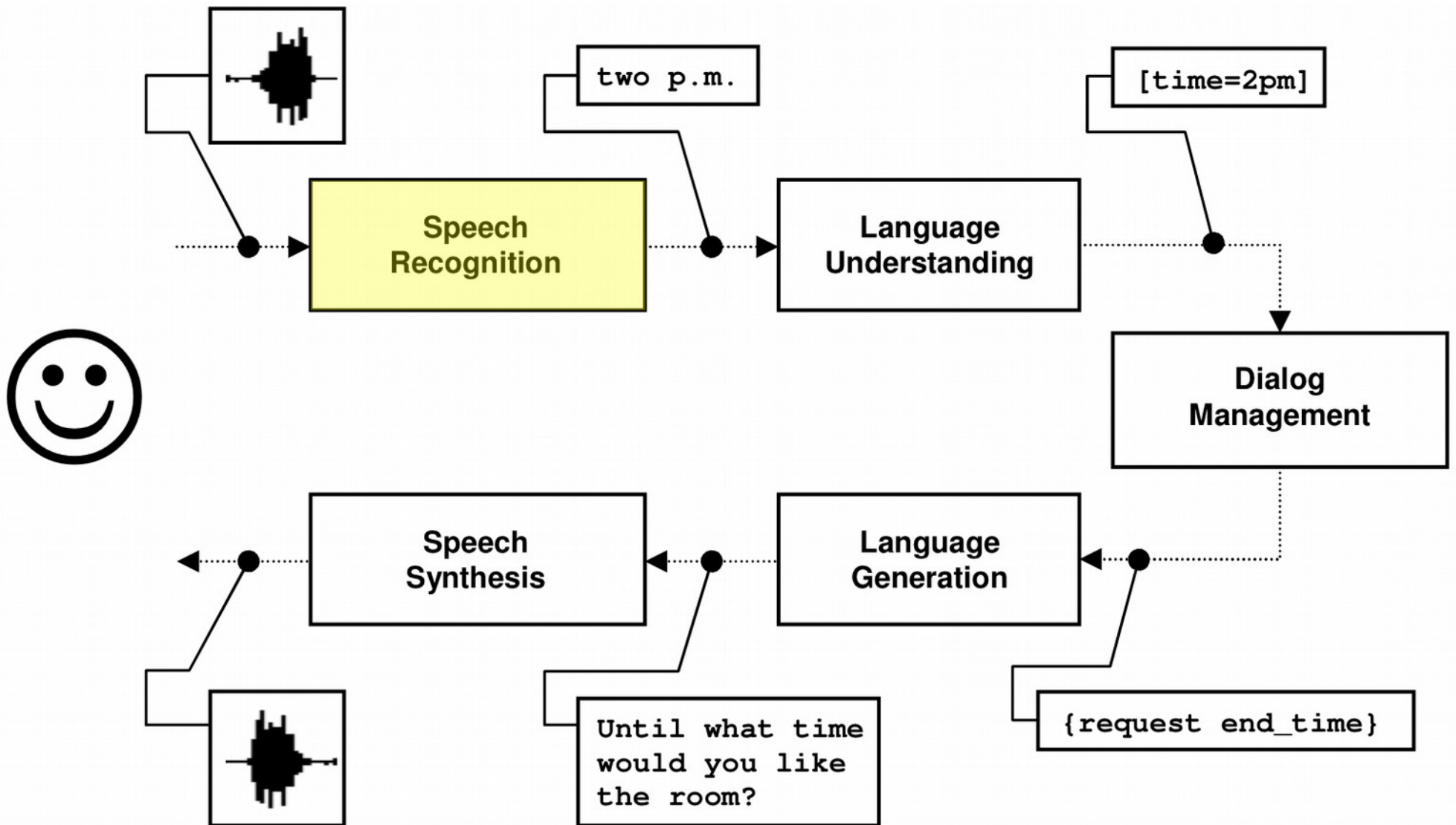
SDH actuales

- *Information-seeking, transaction-based.*
- La mayoría de los sistemas en producción hoy están diseñados para **acceder a BD** y/o **rutear llamadas**.
- Otros ejemplos:
 - CMU: *Let's Go!* Info del transporte de Pittsburgh.
 - MIT: *Jupiter*. Pronóstico del tiempo.
 - Columbia University: *CheckItOut*. Bibliotecario virtual.
 - Asistentes: *Siri* (Apple), *Cortana* (Microsoft), *Google Now*.

Arquitectura de un SDH



Arquitectura de un SDH



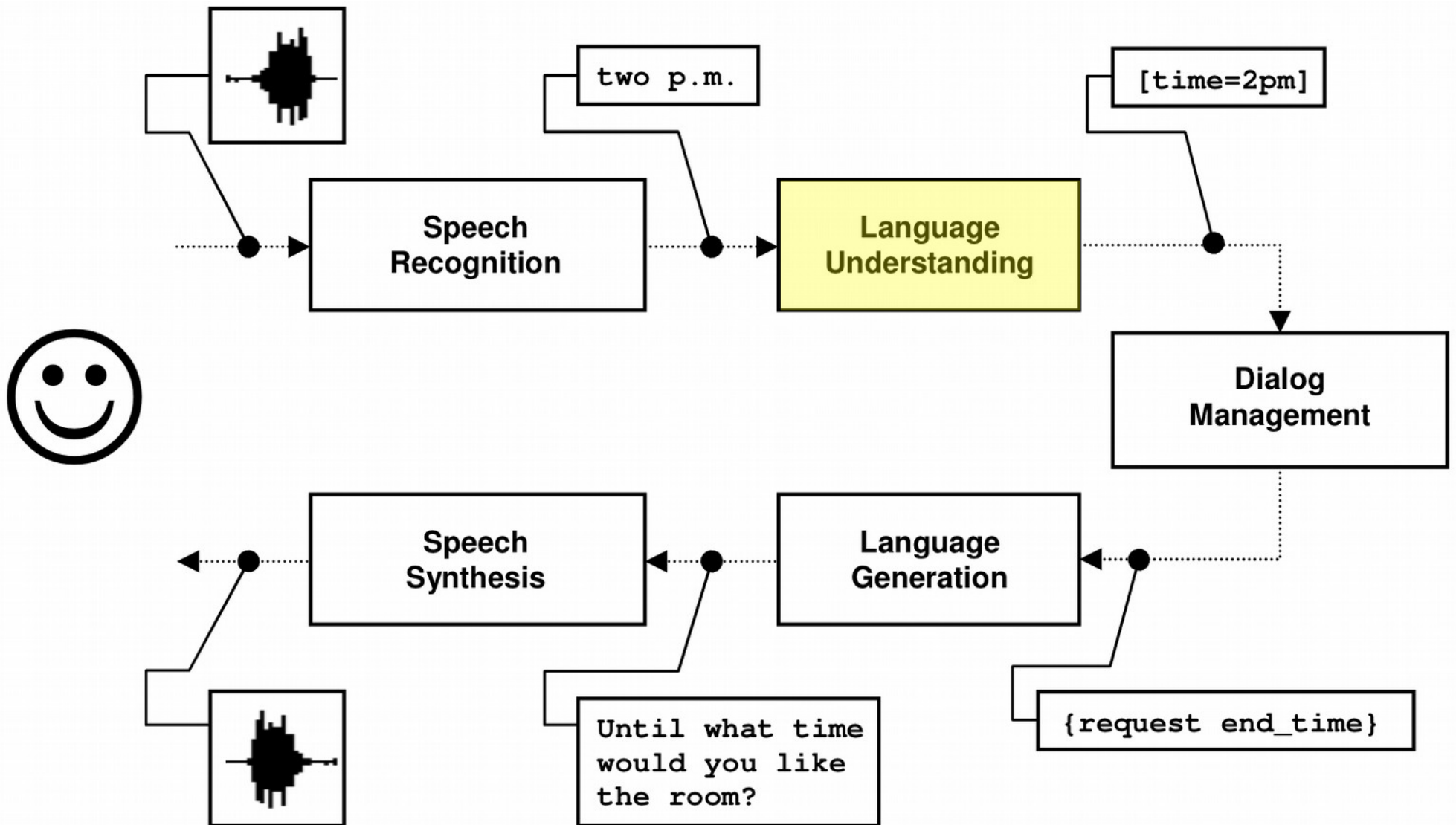
ASR para SDH es difícil...

- Un SDH debe contemplar variabilidad en:
 - Medio ambiente
 - Ruido de fondo, interferencia de celulares, etc.
 - Producción del habla
 - Disfluencias, acento, edad, sexo.
 - Familiaridad del usuario con la tecnología.
 - Los usuarios menos expertos usan palabras fuera del vocabulario/dominio.

ASR introduce incertidumbre

- La arquitectura de un SDH siempre opera con **información parcial**.
 - ASR output con nivel de confianza:
*quiero[.4] saber[.7] el[.5] código[.4] de[.5] área[.8]
para[.5] Chicago[.7]*
- Desafío: manejar ese nivel de incertidumbre.

Arquitectura de un SDH



Comprensión del Lenguaje Natural

- Extracción de significado de las oraciones.
 - *Natural Language Understanding* (NLU).
 - Semántica computacional.
- Formas de **representar el significado** de las oraciones.
 - Lógicas de predicados, lógicas modales.
- Sistemas de diálogo hablado:
 - Dialog acts
 - Frame/Slot Semantics

NLU: Actos de Diálogo

- **Acto de diálogo (*dialog act*):** Captura la acción que se pretende realizar al formular una oración en un diálogo.
- Ejemplos:
 - *Saludo*
 - *Afirmación*
 - *Pregunta por sí/no*
 - *Pregunta “Wh” (cómo, quién, cuándo, cuál, qué)*
 - *Solicitud*
 - *Agradecimiento*
 - ...
- Problema: clasificación automática de DAs.

NLU: Frame Semantics

- Ejemplo de Frame:
 - *Quiero ver los vuelos de Boston a San Francisco para el martes a la mañana.*

MOSTRAR:
 VUELOS:
 ORIGEN:
 CIUDAD:
 FECHA:
 HORA:
 DESTINO:
 CIUDAD:

NLU: Frame Semantics

- Ejemplo de Frame:
 - *Quiero ver los vuelos de Boston a San Francisco para el martes a la mañana.*

MOSTRAR:	
VUELOS:	
ORIGEN:	
CIUDAD:	Boston
FECHA:	Martes
HORA:	mañana
DESTINO:	
CIUDAD:	San Francisco

- Estos frames se completan mediante **gramáticas semánticas**, o enfoques basados en **HMM**.

NLU: Frame Semantics

- Enfoque 1: Gramáticas semánticas.

- SHOW → show me | I want | can I see | ...
- DEPART_TIME → (after | around | before) HOUR
| morning | afternoon | evening
- HOUR → one | two | three... | twelve (am | pm)
- FLIGHTS → (a) flight | flights
- ORIGIN → from CITY
- DESTINATION → to CITY
- CITY → Boston | San Francisco | Denver | Washington

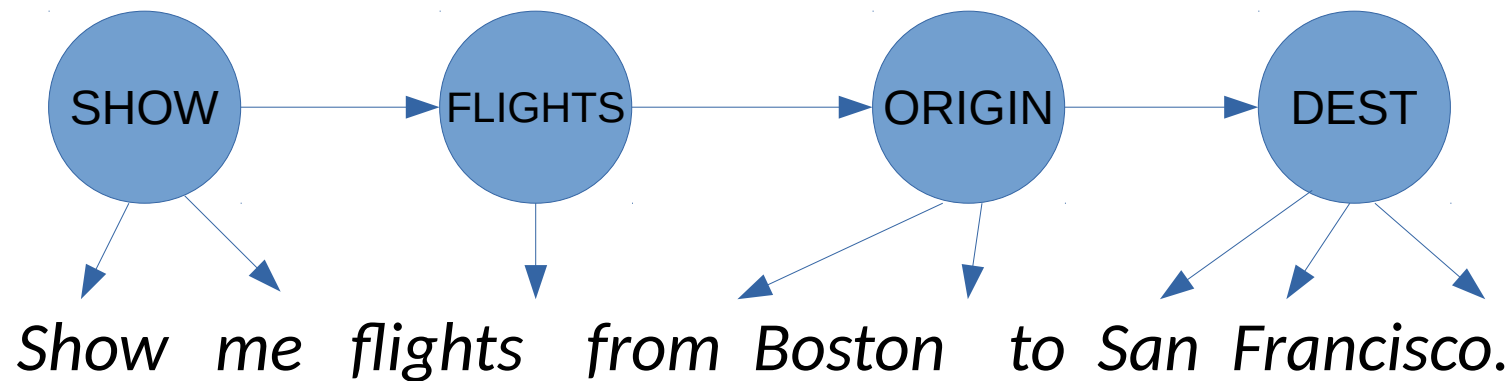
- Ejemplo:

Show me flights from Boston to San Francisco on Tuesday morning.

SHOW	FLIGHTS	ORIGIN	DESTINATION	DEPART_DATE	DEP_TIME
------	---------	--------	-------------	-------------	----------

NLU: Frame Semantics

- Enfoque 2: Hidden Markov Models (HMM)
- Estados ocultos: rótulos semánticos de los slots.
- Observaciones: palabras.



Problemas...

- **Sistema:** *¿Y en qué día de Mayo quiere viajar?*
- **Cliente:** *Bueno, eh, voy a un congreso del 12 al 15.*
- Notar que el cliente no respondió la pregunta.
- Significado de la oración del cliente:
 - Evento=congreso
 - Comienzo-del-evento: 12
 - Fin-del-evento: 15
 - ¡No dice nada sobre el vuelo!
- ¿Cómo puede el sistema inferir que el cliente está informándole las fechas del viaje?

Grice: Implicatura Conversacional

- **Implicatura**: Inferencias que el hablante espera que el oyente realice.
- Se basan en un **principio de cooperación**, acuerdo tácito entre hablante y oyente para cooperar en la comunicación.
- **Máximas** conversacionales de Grice:
 - **Relevancia**: Sé relevante.
 - **Cantidad**: No hagas una contribución más ni menos informativa que lo requerido.
 - **Calidad**: Tratá de hacer contribuciones ciertas (no digas cosas que son falsas, o para las que carecés de evidencia adecuada).
 - **Manera**: Evitá ser ambiguo o poco claro; sé breve y ordenado.

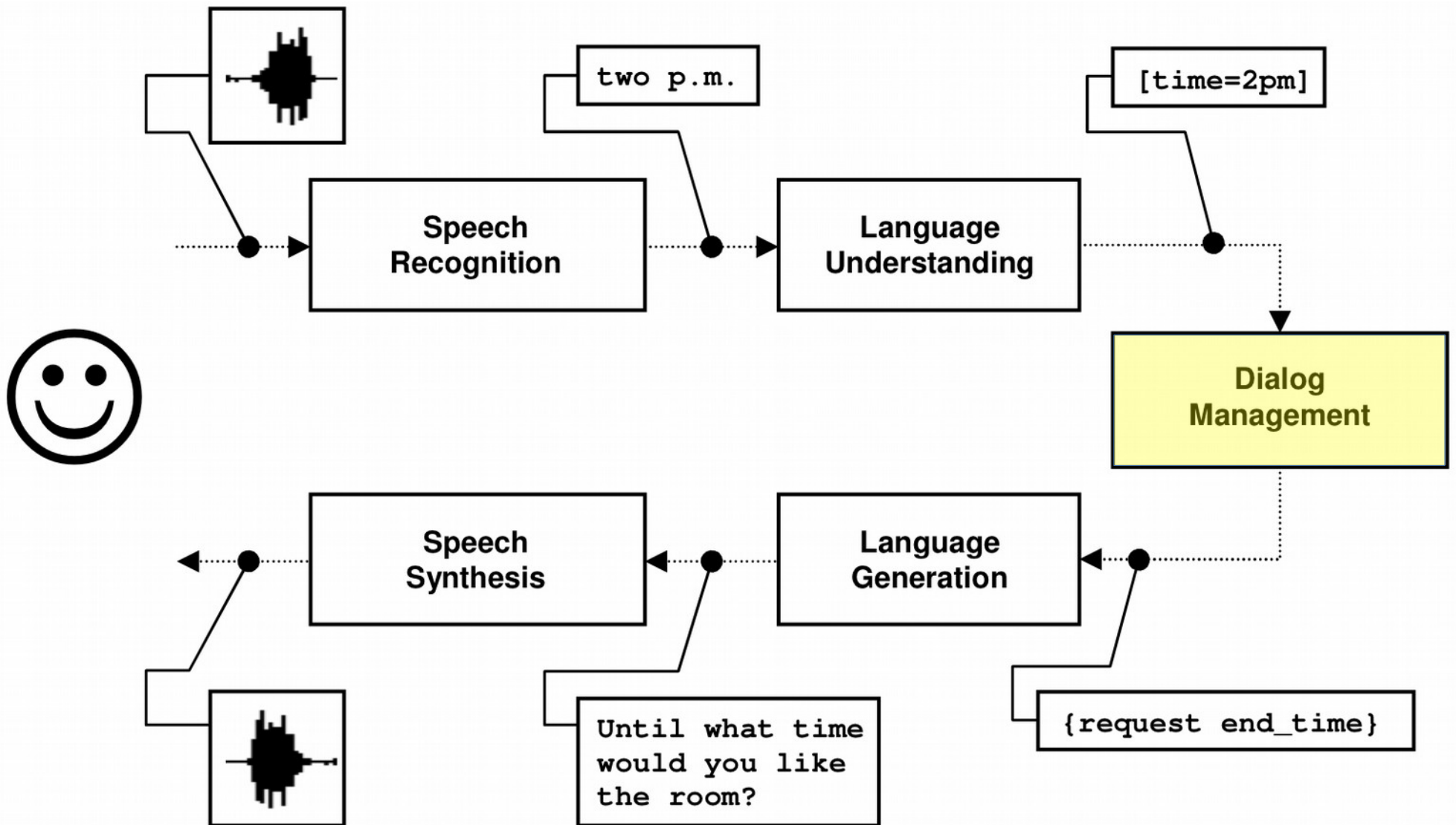
Máximas conversacionales de Grice

- **Máxima de Relevancia:** Sé relevante.
- A: *¿Está Federica?*
B: *Su auto está afuera.*
- **Implicatura:** Sí.
 - El oyente piensa: *¿Por qué menciona B el auto?*
Debe ser relevante. ¿Cómo? Debe ser porque si su auto está afuera, entonces ella probablemente esté acá.

Máximas conversacionales de Grice

- **Máxima de Cantidad:** No hagas una contribución más ni menos informativa que lo requerido.
- **A:** *¿Cuánta plata tenés?*
B: *Tengo 50 pesos.*
- **Implicatura:** B tiene *exactamente* 50 pesos.
 - El oyente piensa: *Si B tuviera más de 50 pesos, me lo diría.*

Arquitectura de un SDH

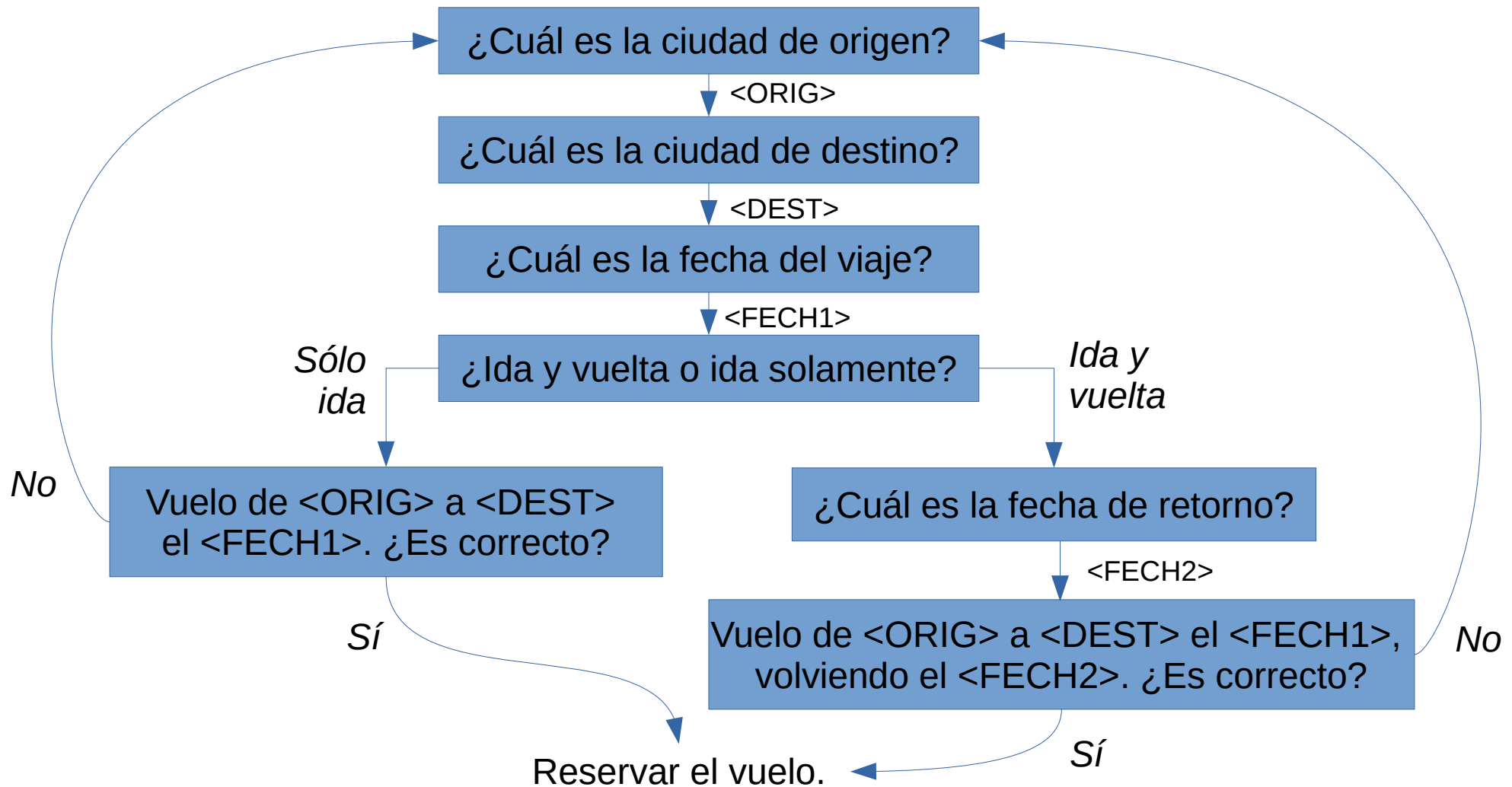


Administrador del diálogo

- ELIZA (Joseph Weizenbaum, MIT, 1966)
 - Chatbot psicoterapista
 - Pattern matching, reglas simples construidas a mano.
 - Demos:
 - <http://www.cyberpsych.org/eliza/>
 - <http://deixilabs.com/eliza.html>
 - ¡Pasa el test de Turing con alguna gente!
- Evolucionó en malware...
 - CyberLover, chatbot que roba info confidencial en chats.
- ...y en herramientas educativas!!
 - Chatbot de Dale Aceptar:
<http://www.daleacceptar.gob.ar/cms/>

Administrador del diálogo

- Autómata de estados finitos (FSA)



Administrador del diálogo

- Autómata de estados finitos (FSA)
 - Funcionan con buenas tasas de error.
 - Modelo ASR específico para cada respuesta del usuario: ciudades, fechas, horas, etc.
 - Pero... muy pobre experiencia de usuario.
 - Mejoras: comandos universales (*help, main menu*).
 - *System-initiative*.

Administrador del diálogo

- Sistemas *mixed-initiative* y *user-initiative*.
 - Ejemplo: Sistemas **open prompt**
 - El usuario puede formular su mensaje libremente.
 - Sistema: *¿En qué le puedo ayudar?*
- Flexibilidad ↔ Incertidumbre

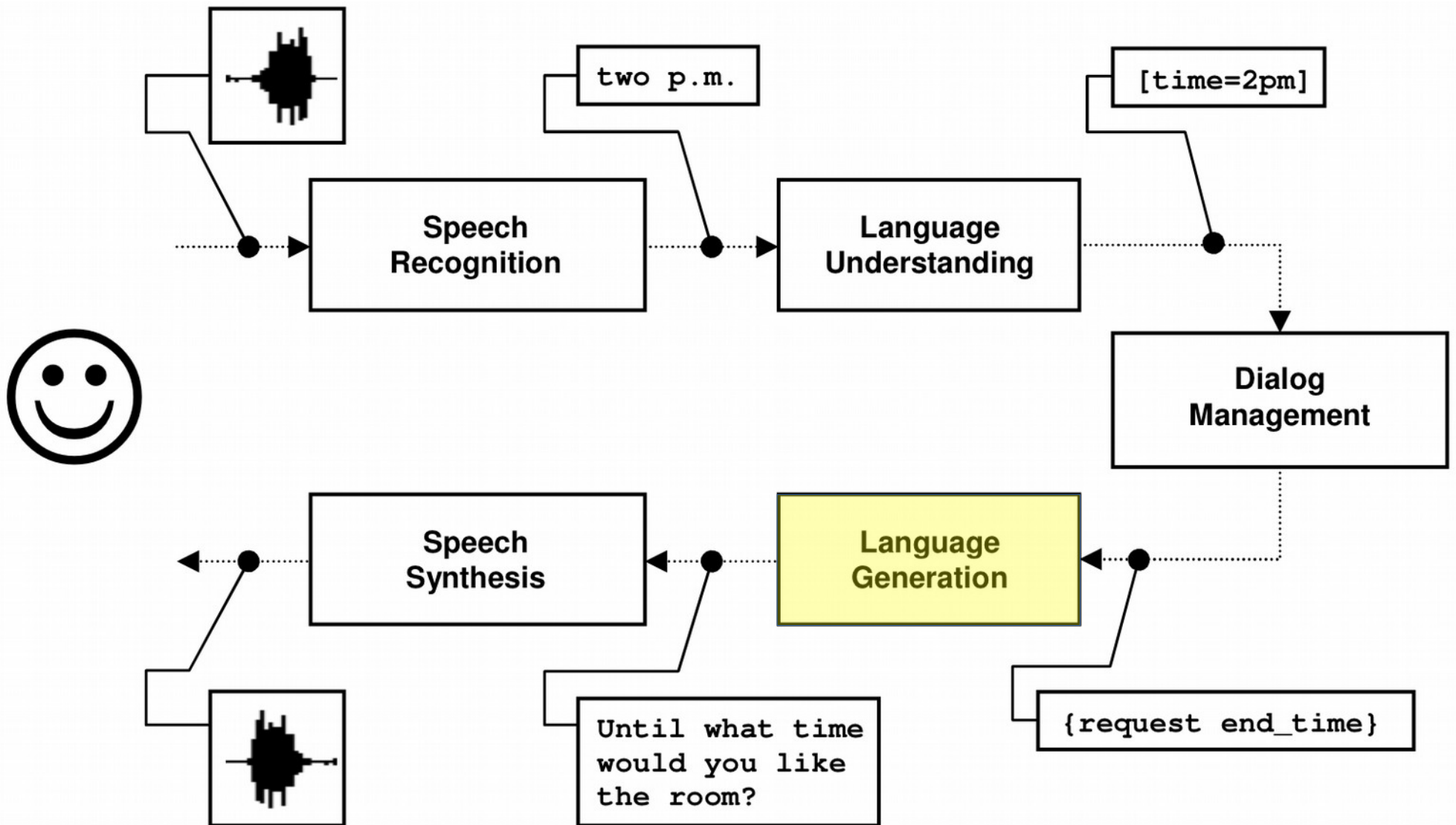
Administrador del diálogo

- Lidiando con la incertidumbre...
 - **Confirmación explícita vs. implícita**
 - Sistema: *¿Cuál es la ciudad de origen?*
Usuario: *Córdoba*
Sistema: *¿La ciudad de origen es Córdoba?*
Usuario: *Sí*
 - Sistema: *¿Cuál es la ciudad de origen?*
Usuario: *Córdoba*
Sistema: *¿Y quiere volar de Córdoba a qué ciudad?*
Usuario: *A Salta.*

Administrador del diálogo

- Lidiando con la incertidumbre...
 - **Progressive prompting**
 - Sistema: *¿Cuándo desea llegar a Salta?*
 - Usuario: *Quiero llegar a tiempo para ver la semifinal de la Liga Nacional de Basket.*
 - Sistema: *Lo siento, no he comprendido su respuesta. Por favor, diga el día y el mes en que le gustaría llegar a Salta.*
 - Usuario: *Ah, el 10 de diciembre.*
 - **Rapid reprompting**
 - Sistema: *¿Cómo dijo?*
 - La 2da vez que no entiende hace la pregunta completa.

Arquitectura de un SDH



Generación del Lenguaje Natural

- Construcción de oraciones a partir de conceptos.
 - Elección de estructura sintáctica y palabras.
- *Natural Language Generation* (NLG)
- Output de NLG → Input de TTS.
- Sistemas de diálogo hablado:
 - Generación basada en templates:
 - ¿A qué hora quiere partir de **CIUDAD-ORIG**?
 - ¿Quiere volver a **CIUDAD-ORIG** desde **CIUDAD-DEST**?

NLG: Restricciones de HCI

Por favor diga la fecha....

Por favor diga la hora de comienzo....

Por favor diga la duración...

Por favor diga a nombre de quién es la reserva...



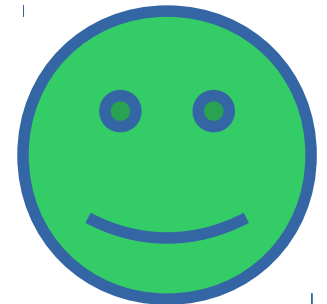
Marcadores de discurso y pronombres (“coherencia”):

Primero, dígame la fecha....

Lo siguiente que necesitaría saber es la hora de comienzo.

Gracias. <pausa> Ahora, ¿cuánto va a durar la reunión?

Por último, ¿a nombre de quién hago la reserva?



NLG: Restricciones de HCI

- Prompts cada vez más cortos:

Sistema: *¿Cuál es la primera compañía que desea agregar a su lista?*

Usuario: Cisco

Sistema: *¿Cuál es el nombre de la siguiente compañía? (O bien, puede decir “Listo”).)*

Usuario: IBM

Sistema: *¿La siguiente compañía? O diga, “Listo”.*

Usuario: Intel

Sistema: *¿La siguiente?*

Usuario: America Online.

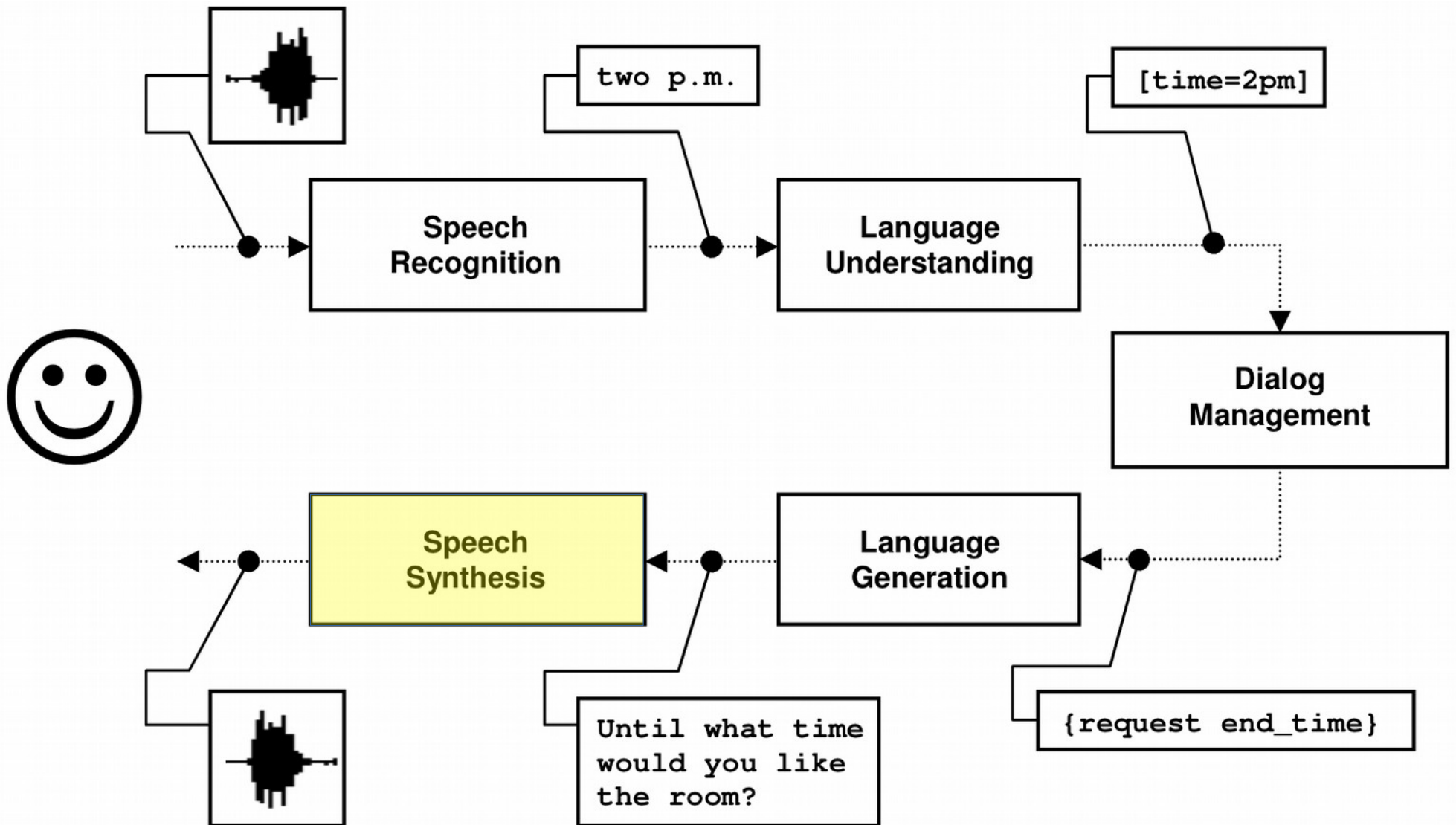
Sistema: *¿La siguiente?*

Usuario: ...

Más problemas...

- **Sistema:** ... *hay 3 vuelos sin escalas para ese día.*
- El sistema dirá esto luego de **encontrar** 3 vuelos en su base de datos. Podría haber más...
- Pero el cliente va a interpretar que hay **solamente** 3 vuelos.
- Tener en cuenta **Implicaturas Conversacionales** al diseñar el módulo de generación de un sistema.

Arquitectura de un SDH



TTS para SDH

- SDH de **dominio abierto** → TTS de **dominio abierto**.
- En **dominios limitados**:
 - Selección de unidades largas (palabras o incluso frases).
 - Mucho mejor calidad. Menos costoso.
- Problema frecuente: **nombres propios** (apellidos, ciudades, calles).
 - Concatenación de difonos; selección de unidades cortas (difonos); síntesis con HMMs, etc.

Arquitectura de un SDH

