

# El plataforma de adquisición de diálogos en el proyecto DIHANA\*

M.I.Galiano<sup>†</sup> R.Granell<sup>†</sup> L.I.F.Hurtado<sup>†</sup> A.Miguel<sup>‡</sup> J.A.Sánchez<sup>†</sup> E.Sanchís<sup>†</sup>

<sup>†</sup>DSIC - Universidad Politécnica de Valencia  
Camino de Vera s/n  
Valencia 46022 (Spain)  
{mgaliano,rgranell,lhurtado,jandreu,esanchis}@dsic.upv.es

<sup>‡</sup> DIEC - Universidad de Zaragoza  
C/ María de Luna 1, Edificio A, L 3.05  
Zaragoza 50018 (Spain)  
amiguel@unizar.es

## 1. Introducción

Un sistema de diálogo es un sistema automático capaz de emular a un ser humano en un diálogo con otra persona, con el objetivo de que el sistema cumpla con una cierta tarea (dar una cierta información, o proporcionar ciertos servicios). El objetivo del proyecto “DIHANA: Sistema de Diálogo para el acceso a la información mediante habla espontánea en diferentes entornos”(TIC2002/04103-C03) es el estudio y desarrollo de uno de estos sistemas. El sistema de diálogo diseñado en DIHANA es un sistema modular y distribuido con las siguientes partes: módulo de *Reconocimiento automático del habla*; módulo de *Comprensión*; módulo de *Diálogo*; módulo de *Acceso a la base de datos de la aplicación*; y módulo de *Generación de respuestas de usuario*. En DIHANA se pretende obtener cada uno de estos módulos de manera *automática*, a partir de un conjunto de ejemplos de *consulta telefónica de horarios y precios de trenes de grandes líneas*. Esta forma de aprendizaje requiere una fase previa de adquisición de un corpus específico de diálogo usuario-sistema.

La adquisición de un corpus específico de diálogo usuario-sistema es una de las tareas más sensibles para el desarrollo de un sistema de diálogo: para que esta adquisición se realice de una manera *natural* se precisa un sistema de diálogo que funcione eficientemente; sin embargo, para desarrollar un sistema de diálogo eficiente es necesario una gran cantidad de datos (diálogos naturales) para el entrenamiento de sus modelos. La solución típica para obtener una adquisición inicial consiste en emplear la técnica basada en un *Mago de Oz* en la que una persona sustituye al sistema de diálogo. La labor del Mago consiste en simular el sistema automático, ayudando al usuario a obtener respuesta a sus consultas, por lo que, en sucesivos turnos de diálogo, interacciona con el usuario siguiendo una estrategia dada. Para realizar este trabajo, el Mago puede estar asistido por un sistema de adquisición que facilite su labor.

En esta demostración se pretende ilustrar el comportamiento de la plataforma de adquisición desarrollada en el marco de DIHANA para la adquisición de un corpus de diálogo de consulta telefónica de horarios y precios de trenes de grandes líneas.

## 2. El plataforma de adquisición

La plataforma desarrollada en el marco del proyecto está formada por diferentes componentes que son utilizadas en el proceso de adquisición: un gestor de comunicaciones (línea telefónica), un sistema de parametrización y reconocimiento automático del habla, un sistema de diálogo, un sistema de comprensión y un sistema de síntesis de voz. Todas estas componentes excepto el sistema de diálogo están automatizadas y, en parte, son el fruto del proyecto BASURDE (TIC98/0423-C06). El sistema de diálogo se pretende simular mediante la técnica del Mago de Oz.

La forma en la que un usuario interacciona con el Mago es como sigue:

- el usuario realiza una consulta telefónica al sistema de manera que el Mago escucha lo que realmente dice el usuario. Las consultas son sobre alguno o algunos de los siguientes conceptos o *frames*: horarios, precios, duración o tiempo de recorrido, tipos de tren y servicios. Cada uno de estos *frames* constituye el concepto principal de la consulta que se está realizando;
- la señal vocal pasa al reconocedor automático del habla, el cual proporciona la secuencia de palabras que con mayor probabilidad ha dicho el usuario, junto con un valor numérico que da cuenta de la fiabilidad con que se ha reconocido dicha secuencia;
- a continuación el módulo de comprensión toma de dicha secuencia la información semántica relevante. Dicha información es utilizada para completar la información asociada al *frame*. Si dicha información es insuficiente o no es muy precisa, el Mago solicita nueva información al usuario en un nuevo turno de diálogo, y así sucesivamente hasta que el Mago considera que la información disponible es suficiente para realizar una consulta a la base de datos.

---

\*Este trabajo ha sido subvencionado por el proyecto TIC2002/04103-C03 de la CICYT.

La labor del Mago consiste en ayudar al usuario a obtener respuesta a sus consultas, por lo que, en sucesivos turnos de diálogo, interacciona con el usuario siguiendo una estrategia dada. Para realizar este trabajo el Mago siempre dispone de la información sobre el diálogo que ha sido o está siendo procesado. Esta información está recogida en una estructura a la que llamamos *pizarra*. Inicialmente, la pizarra siempre tiene cierta información que obtiene del sistema: fecha actual, hora, etc.

Dado un estado del diálogo y la información proporcionada por el usuario, el Mago procesa esa información, la entiende, actualiza la información de la pizarra con ella y, finalmente, en función de dicha actualización, interacciona nuevamente con el usuario siguiendo una estrategia determinada. Para realizar este proceso de cambio de estado del diálogo, el Mago debe contar, además de con la pizarra, con otra información que le permitan entender lo dicho por el usuario, actualizar la pizarra, aplicar la estrategia marcada y si corresponde, consultar la base de datos y responder al usuario. Esta información es la salida del módulo de reconocimiento y la salida del módulo de comprensión. El Mago interacciona con el usuario haciendo uso de un módulo de generación de respuestas escritas que posteriormente son comunicadas a través de un sintetizador.

En la Figura 1 puede verse el aspecto de la plataforma de adquisición que ha sido desarrollada. Esta plataforma comprende las diferentes fuentes de conocimiento que han sido mencionadas anteriormente.

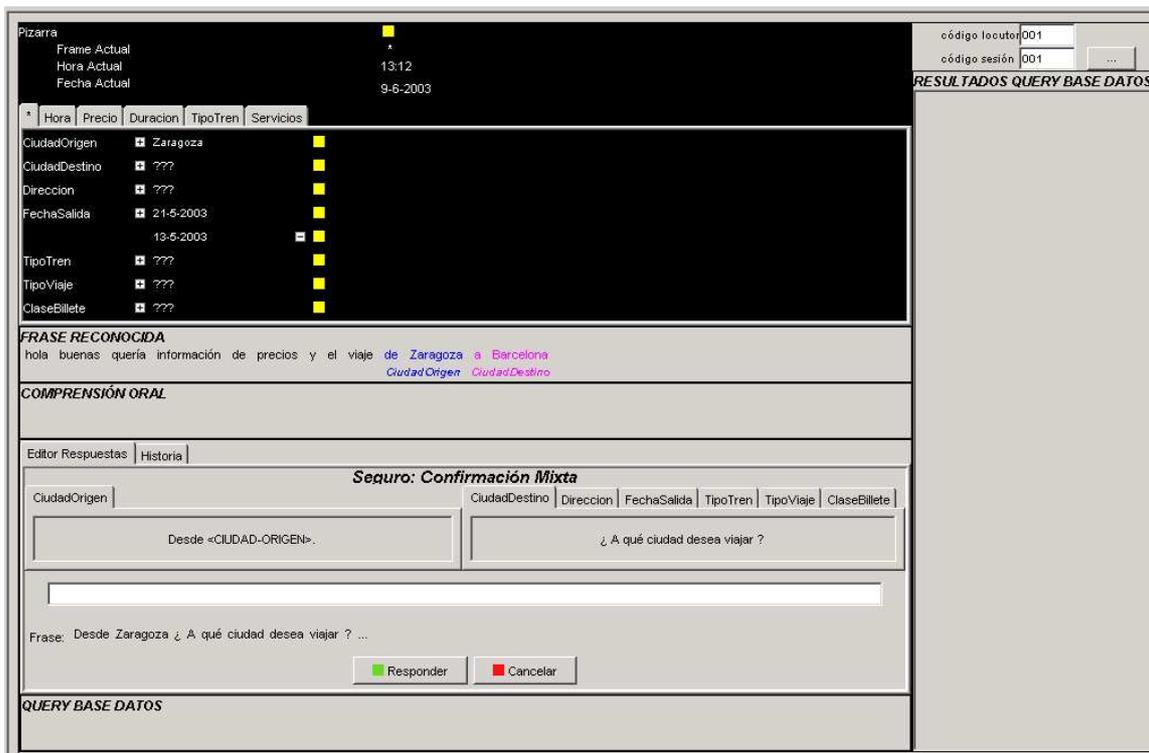


Figura 1: Aspecto de la plataforma de adquisición desarrollada en el proyecto DIHANA

El Mago supervisa la información que se genera durante el proceso de diálogo, modificándola si fuera necesario, e interacciona con el usuario siguiendo alguna de las siguientes modalidades:

- completar información,
- confirmar datos,
- indicar al usuario que no ha entendido lo dicho,
- o indicar al usuario que la información proporcionada no es válida.

La elección de una de las modalidades va ligada a la confianza que se tenga en lo que ha dicho el usuario. Dicha confianza está automatizada en algunos casos. Por ejemplo, el reconocedor proporciona un valor ligado a la verosimilitud que se tiene en lo dicho por el usuario. En otros casos el Mago debe decidir acerca de lo dicho por el usuario. Cuando el Mago considera que ya puede realizar una consulta a la base de datos del sistema, se lo indica al usuario y la efectúa.