

# Proyecto ALIADO: Tecnologías del habla y el lenguaje para un asistente personal

**José B. Mariño**

Centro TALP-UPC

Campus Nord UPC. 08034-Barcelona  
[canton@talp.upc.es](mailto:canton@talp.upc.es) (telf: 93 401 6444)

**Horacio Rodríguez**

Centro TALP-UPC

Campus Nord UPC. 08034-Barcelona  
[horacio@talp.upc.es](mailto:horacio@talp.upc.es) (telf: 93 401 7024)

**Resumen:** ALIADO aborda el desarrollo de tecnologías del habla y el lenguaje de interés para el diseño de asistentes personales en un entorno plurilingüe. Se dedica especial atención al interfaz oral del usuario con el asistente. Las ayudas facilitadas por el asistente se centra en el uso del lenguaje: la búsqueda de respuesta y la traducción de texto y habla. Finalmente, se propone la construcción de dos demostradores.

**Palabras clave:** agente personal, reconocimiento de voz, robustez, conversión texto a voz, traducción estocástica, búsqueda de la respuesta.

**Abstract:** ALIADO undertakes the developing of spoken and written language technologies for the design of personal assistants in a multilingual environment. Main attention is paid to the design of the oral interface. We consider two examples of language centred help that can be provided by the assistant: “question answering” and text or speech machine translation. The technologies developed will be used to implement two showcases.

**Keywords:** agent technology, speech recognition, robustness, text-to-speech conversion, stochastic speech to speech translation, question answering.

## 1 Consorcio del proyecto

El consorcio está constituido por el Grupo de Tratamiento del Habla del Departamento de Teoría de la Señal y Comunicaciones de la Universidad Politécnica de Cataluña (UPC) y el Grupo de Tratamiento del Lenguaje Natural del Departamento de Lenguajes y Sistemas Informáticos de la UPC. En marzo de 1999 ambos grupos constituyeron el Centro de Investigación en Tecnologías y Aplicaciones del Lenguaje y el Habla (TALP).

## 2 Entidad financiadora

ALIADO está financiado por el Ministerio de Ciencia y Tecnología como proyecto coordinado (TIC2002-04447-C02).

## 3 Los agentes personales

### 3.1 El concepto

En los últimos años se ha dedicado un gran esfuerzo al desarrollo de la tecnología del habla y, lentamente, comienza a tener presencia en la vida cotidiana. Aparte de los sistemas de dictado, la tecnología del habla ha encontrado su mayor campo de aplicación en servicios de

comunicaciones (centralitas, telebanca, etc.), si bien es cierto que, en muchos casos, como tecnología de apoyo a operadores humanos. Frecuentemente, cuando se hacen previsiones sobre la futura sinergia entre la tecnología del lenguaje y la tecnología de las comunicaciones, se menciona a los “personalized intelligent agents”, que en adelante llamaremos asistentes personales (AP). Por un lado, el interfaz oral resulta una solución natural a la demanda contradictoria entre mayores funcionalidades y un tamaño más reducido de los terminales, y por otro el texto constituye un formato habitual para la representación de información. En general, este agente es imaginado como un sistema cuya interacción con el usuario se realiza mediante un dispositivo móvil y cuya inteligencia puede residir en un servidor al que se accede a través de red o en el propio dispositivo, y que puede realizar múltiples tareas, tales como:

- Gestión de comunicaciones y mensajes.
- Búsqueda y resumen de información.
- Gestión de agenda.
- Traducción de mensajes.
- Interpretación en conversaciones.
- Asistencia a usuarios con problemas de audición, vista, etc.

### 3.2 Proyectos seminales

En la actualidad se realizan esfuerzos dirigidos al desarrollo de las tecnologías precisas para el diseño de asistentes personales de estas características. En los Estados Unidos cabe citar los proyectos COMUNICATOR, financiado por DARPA, y el proyecto PERSONA de Microsoft. En la Unión Europea no existe ningún proyecto equivalente a los citados, aunque pueden señalarse varios que afrontan aspectos parciales (DUMAS, IMAGINE) o atienden a la obtención de recursos para su desarrollo (SPEECON, LC-Star).

### 4 El proyecto ALIADO

ALIADO se plantea la creación de conocimientos y el ensayo de metodologías de interés para el desarrollo de asistentes personales (AP) que se comuniquen con el usuario mediante el habla. No se trata del diseño de tal asistente personal, lo que por la diversidad de tecnologías que implica escapa a las posibilidades del consorcio, sino de atender a un conjunto de ellas relacionadas con el lenguaje (oral o escrito). Aunque estas tecnologías tienen interés en otros dominios de aplicación, en ALIADO se contempla al asistente personal como un marco de referencia para el desarrollo metodológico.

Es importante destacar también que en un mundo globalizado, donde deben relacionarse personas de lenguas diferentes, en el seno de la Unión Europea, cuya integración precisa romper las barreras lingüísticas, o en un estado como el español, que incluye regiones bilingües, se hace necesario afrontar el plurilingüismo en las tecnologías del lenguaje. Dado que los sistemas actuales que tratan el habla se basan en el diseño a partir de datos, para atender a estos problemas es preciso disponer de corpus plurilingües. El consorcio de ALIADO dispone de corpus del habla de todas las lenguas de la Unión Europea (fruto de su participación en el proyecto SPEECDAT y de las principales variedades dialectales del español (adquiridos en el seno del proyecto SALA). Además, como participe del proyecto LC-STAR, dispondrá a corto plazo de recursos para el desarrollo de la traducción automática del habla entre las principales lenguas europeas y el inglés americano.

En este entorno plurilingüe, el proyecto desarrolla un interfaz oral, que incluya las herramientas para el reconocimiento y síntesis

de voz necesarias para establecer una negociación entre el usuario y el asistente, y abordar la traducción del habla. En particular, se estudiará la parametrización del habla captada por un micrófono lejano. Se investigará el modelado acústico-fonético en un entorno plurilingüe, que incluye las lenguas romances, inglés y árabe mediterráneo. Se desarrollará una arquitectura para el reconocimiento de voz (en tiempo menor que el real) adecuada para el tratamiento de vocabularios de decenas de miles de palabras y que se integre eficazmente con las tareas de comprensión y traducción del habla espontánea. Se incluye el desarrollo de un controlador de diálogo de iniciativa mixta. Se aborda también la conversión de texto a voz, a fin de proporcionar sistemas totalmente orales; en particular se estudia la generación de voces de calidad a partir de bases de datos de tamaño reducido. Por último, se aborda la traducción del habla entre inglés y castellano. Merece ser destacado que aquí la tecnología del habla supera el tradicional papel de interfaz entre usuario y máquina, para intermediar entre dos personas,

En este contexto es asimismo importante señalar la importancia que el acceso selectivo y preciso a la información tiene para un AP. El gran reto de las tecnologías de acceso a la información no es hoy en día su adquisición sino el filtrado y resumen de la información adquirida de forma que lo que se proporcione al usuario sea la respuesta exacta o la información más relevante de acuerdo a su consulta. Las tecnologías de búsqueda de la respuesta y el resumen automático guiado por la consulta proporcionan herramientas potentes y precisas para hacer frente a este reto.

Se realizará un demostrador de las tecnologías del habla y el lenguaje desarrolladas para un asistente personal en un viaje turístico. El usuario, provisto de su asistente, visita un lugar cuya lengua no domina o no conoce, lo que ofrece una motivación para la traducción del habla y la búsqueda de respuesta en un entorno plurilingüe.

### 5 Página web

En página web del proyecto

<http://gps-tsc.upc.es/veu/aliado/>

puede consultarse la memoria del mismo. En dicha página se harán públicos también los principales resultados alcanzados con el progreso de ALIADO.