

Área de
Programas y Proyectos



PROPUESTA DE PROYECTO PARA LA CONVOCATORIA SOBRE CONTENIDOS DIGITALES EN EL
MARCO DEL PLAN AVANZA (AVANZA CONTENIDOS)

EXPERIENCIAS PILOTO DEMOSTRADORAS

Desarrollo de experiencias piloto y demostradores que promuevan la utilización de la
información y contenidos del sector público.

MEMORIA TÉCNICA

Desarrollo de un sistema de traducción de voz
a Lengua de Signos Española para un servicio
público de atención personal

Proyecto Coordinado

Fundación de la Confederación Estatal de
Personas Sordas (Fundación CNSE)



ÍNDICE

Resumen.....	i
1 Introducción	2
2 Justificación.....	3
3 Objetivos del proyecto	4
4 Tareas a desarrollar en el proyecto.....	5
4.1 Tarea 1. Estudio lingüístico de los servicios de atención abordados	5
4.2 Tarea 2. Diseño de la arquitectura del sistema.....	6
4.3 Tarea 3: Adaptación del sistema de reconocimiento de voz.....	7
4.4 Tarea 4: Desarrollo del módulo de traducción basado en reglas de un experto	7
4.5 Tarea 5: Desarrollo del módulo de traducción estadística	8
4.6 Tarea 6: Desarrollo de la base de datos de signos con una avatar en 3D.....	9
4.7 Tarea 7: Evaluación final de los demostradores	10
4.8 Tarea 8: Gestión y documentación	10
5 Planificación de tareas.....	11
6 Resultados del proyecto previstos y contenidos digitales a generar	11
7 Plan de explotación de resultados.....	12
8 Difusión de resultados y conocimiento	13
9 Adecuación a los objetivos de la propuesta	14
10 Referencias.....	15

RESUMEN

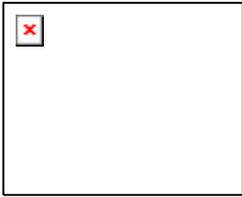
El objetivo principal del proyecto que se propone en esta solicitud es el diseño, desarrollo y evaluación de una arquitectura software que permita la traducción de voz a Lengua de Signos Española (LSE). En esta arquitectura, las traducciones a LSE se representan mediante un agente animado en 3D. Con el fin de evaluar y probar la arquitectura planteada se desarrollarán dos demostradores aplicados a la traducción de las frases que un funcionario público pronuncia cuando atiende personalmente a una persona sorda. Estos demostradores estarán orientados a algunos servicios públicos de atención concretos, como por ejemplo, el servicio de información y gestión para la renovación del Documento Nacional de Identidad (DNI), el servicio de información y gestión de los servicios que presta la DGT (Dirección General de Tráfico), o el servicio de información sobre los centros de atención e información sobre las prestaciones económicas del Instituto Nacional de la Seguridad Social (INSS).

Principales resultados esperados:

- Arquitectura software de traducción de voz a LSE.
- Dos demostradores que permiten la traducción de voz a LSE aplicados a traducir las frases de un funcionario en servicios públicos de atención personal.
- Herramientas de ayuda a la traducción entre el castellano y la LSE que faciliten una comunicación fluida entre personas sordas y oyentes en servicios públicos de atención.

Principales contenidos multimedia a generar:

- Un corpus paralelo con las frases en castellano y su traducción en LSE.
- Generación de una base de datos de contenidos multimedia con las descripciones, características, significado y animación de los signos correspondientes a los servicios públicos de atención abordados en este proyecto.



INTRODUCCIÓN

Durante los últimos 20 años se han producidos importantes avances tecnológicos en las tres áreas tecnológicas que están implicadas en el desarrollo de un sistema de traducción de voz a lengua de signos: estudios de la lengua de signos, algoritmos de traducción entre lenguas y sistemas de representación de movimientos mediante agentes virtuales en 3D.

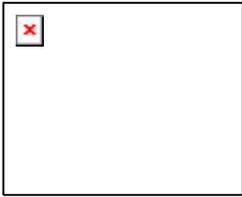
La lengua de signos presenta una gran variabilidad entre diferentes países e incluso en diferentes áreas de un mismo país. Por esta razón, desde 1960, la lengua de signos ha sido objeto de estudio en Estados Unidos (Stokoe, 1960; Christopoulos and Bonvillian, 1985; Pyers, 2006), en Europa (Engberg-Pedersen, 2003; Atherton, 1999; Meurant, 2004), Africa (Nyst, 2004) y Asia (Abdel-Fattah, 2005; Masataka, 2006). En España, durante los últimos 20 años han aparecido una gran cantidad de estudios de la Lengua de Signos Española (LSE). En 1991 M.A. Rodríguez (Rodríguez, 1991) llevó a cabo un análisis detallado de las diferencias entre la lengua de signos utilizada por las personas sordas y las propuestas de estandarización. Este trabajo ha sido ampliado significativamente con trabajos posteriores (Gallardo, 2002; Herrero-Blanco and Salazar-Garcia, 2005; Reyes, 2005).

Por otro lado, la traducción de habla natural ha sido y sigue siéndolo hoy en día una de las principales áreas de investigación en proyectos como C-Star, ATR. Vermobil, Eutrans, LC-Star, PF-Star y TC-Star. Todos estos proyectos, a excepción de TC-Star, son proyectos enfocados a la traducción de habla en dominios restringidos (como por ejemplo una agencia de viajes o turismo) con vocabularios de tamaño medio (10.000 palabras). Los mejores sistemas de traducción están basados en soluciones estadísticas (Och and Ney, 2002), incluyendo técnicas basadas en ejemplos (Sumita et al, 2003), traductores de estados finitos (Casacuberta and Vidal, 2004) y otras soluciones basadas en datos. Los importantes progresos conseguidos en traducción de habla se deben principalmente a factores como la aparición de medidas de error (Papineli et al, 2002), mejora de la eficiencia de los algoritmos de entrenamiento (Och and Ney, 2003), desarrollo de modelos dependientes del contexto (Zens et al, 2002) y algoritmos de generación eficientes (Koehn, 2003).

El proyecto europeo eSIGN (Essential Sign Language Information on Government Networks) (<http://www.sign-lang.uni-hamburg.de/eSIGN/>) constituye uno de los esfuerzos más importantes para el desarrollo de herramientas de apoyo a la generación de contenidos para personas sordas. En este proyecto, uno de los principales resultados ha sido la generación de un avatar en 3D (VGuido) capaz de representar cualquier signo de la lengua de signos.

En los últimos años ha aumentado el interés de varios grupos por la traducción automática a lengua de signos desarrollándose varios prototipos: basados en ejemplos (Morrissey and Way, 2005), basados en reglas (Huenerfauth, 2004), de frases completas (Cox et al, 2002) o soluciones estadísticas (Bungeroth et al, 2006).





En este proyecto se propone el diseño y desarrollo de uno de los primeros sistemas de traducción de voz a signos y el primero para Lengua de Signos Española. Para abordar este proyecto, el consorcio proponente tiene las mejores garantías para completar con éxito dicho proyecto. Por un lado se dispone de un socio tecnológico como el Grupo de Tecnología del Habla de la Universidad Politécnica de Madrid con más de 30 años de investigación en esta tecnología, y por otro lado, la Fundación CNSE que aporta el conocimiento experto imprescindible para que los sistemas desarrollados sean realmente útiles a la comunidad de personas sordas.

JUSTIFICACIÓN

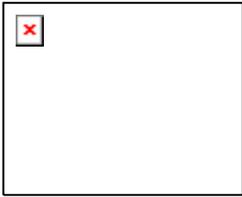
Basándonos en datos del INE y del MEC (CNSE, 2003), el 47% de las personas sordas o con discapacidad auditiva mayores de 10 años no tienen estudios o son analfabetas. La realidad nos muestra que de este colectivo, hasta un 92% padece analfabetismo funcional: tienen serias dificultades para comprender textos y expresarse correctamente por escrito. De esta manera, solo entre un 1% y un 3% de las personas sordas en España han superado estudios universitarios, unos niveles muy inferiores al del conjunto de la población española.

Esto deriva en diversas consecuencias en la calidad de vida de las personas sordas, aunque la que nos interesa destacar aquí está relacionada con el acceso de las personas sordas a la información. No cabe duda de que, según la definición propuesta por la OMS en la Clasificación Internacional de Deficiencias, Discapacidades y Minusvalías (habitualmente tomada como referencia), las personas sordas son personas con discapacidad. Sin embargo, el concepto de discapacidad de la OMS pretende evitar relegar a la condición de enfermos a las personas con discapacidad y desligarlas de cualquier estimagtización. En este sentido, la lengua de signos española, una lengua natural en las personas sordas que se estructura en los mismos niveles lingüísticos y cumple las mismas funciones lingüísticas que cualquier lengua oral, puede contribuir de forma efectiva al mejor acceso a la información por las personas sordas, la supresión de las barreras de comunicación y la igualdad de oportunidades.

El primer análisis lingüístico moderno de una lengua de signos fue llevada a cabo en los Países Bajos por Bernard Tervoort en 1956 (Leeson, 2000), aunque el primer estudio descriptivo y riguroso de la lengua de signos como lengua natural de las personas sordas fue realizado en EE.UU. por W. Stokoe en 1960 (Minguet, 2000: 208), a la que inmediatamente siguieron otras investigaciones en países americanos (por ejemplo, Friedman, 1975, 1976; Kegl y Wilbur, 1976), europeos (v.gr., Klima y Bellugi, 1979), y finalmente españoles (Rodríguez González, 1992; Muñoz Baell, 1998, 1999). Dichas investigaciones confirman que todas y cada una de las lengua de signos¹ estudiadas

¹ Cada país tiene su propia lengua de signos, e incluso dentro de un mismo país pueden existir varias lenguas de signos o variedades lingüísticas de la misma lengua de signos. Esto se debe básicamente a que el desarrollo de la lengua de signos como lengua se





cumplen las mismas funciones que cualquier otra lengua oral y se estructura a sus mismos niveles lingüísticos, con reglas complejas de gramática, aunque diferente a las lenguas orales en cuanto a su producción y comprensión.

Aunque desde la década de los años noventa se han logrado importantes avances en España para la supresión de las barreras de comunicación en las personas sordas con la progresiva incorporación de recursos en lengua de signos española, como la dotación de intérpretes, dicho servicio es realmente insuficiente en la medida en que existe una ratio de 1 intérprete por cada 105 personas sordas, mientras que en otros países europeos hay 1 intérprete por cada 10 personas.

El diseño y desarrollo de una tecnología que permita la traducción del castellano a la LSE supone una iniciativa novedosa en nuestro país la cual permita continuar avanzando en la mejora de las condiciones de accesibilidad para el colectivo de personas sordas usuarios de la LSE. La puesta en marcha de este tipo de herramientas contribuyen a la normalización de la vida de este colectivo facilitando su acceso a los mismos lugares, ámbitos, bienes y servicios que están a disposición de cualquier persona.

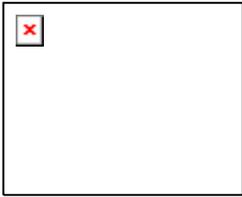
OBJETIVOS DEL PROYECTO

El objetivo principal del proyecto que se propone en esta solicitud es el diseño, desarrollo y evaluación de una **arquitectura software que permita la traducción de voz a Lengua de Signos Española (LSE)**. En esta arquitectura, las traducciones a LSE se representan mediante un agente animado en 3D. Con el fin de evaluar y probar la arquitectura planteada, se desarrollarán dos demostradores aplicados a la **traducción de las frases que un funcionario público pronuncia cuando atiende personalmente a una persona sorda**. Este objetivo principal se puede estructurar en varios subobjetivos que se detallan a continuación:

- Realización de un estudio lingüístico sobre las frases en castellano pronunciadas por el funcionario público y su traducción a la Lengua de Signos Española. Como resultado de este estudio se obtendrá un corpus paralelo con las frases en castellano y su traducción en LSE. Estos contenidos serán de gran importancia para realizar futuros estudios de traducción entre ambas lenguas y para desarrollar nuevas aplicaciones de aprendizaje de la LSE o de traducción en otros servicios.
- Diseño, implementación y evaluación de las diferentes tecnologías de reconocimiento de voz y de traducción entre lenguas, y su aplicación para el caso particular de traducción de castellano a LSE en servicios públicos de atención personal.

encuentra condicionada por el grado y cantidad de intercambios comunicativos entre los/as usuarios/as de la lengua, generalmente dentro de un mismo país o región, por lo que las lenguas de signos siguen una evolución lingüística muy diferente de las lenguas habladas. Skutnabb-Kangas (2002) cree que el número de lenguas de signos en el mundo puede ascender a 5.000, casi tantas como lenguas habladas.





- Desarrollo de varios sistemas para la generación de propuestas de traducción que posteriormente son corregidas por los especialistas en LSE, y un entorno para el diseño y generación de las animaciones de los signos mediante un agente animado virtual.
- Generación de una base de datos de contenidos multimedia con las descripciones, características, significado y animación de los signos correspondientes a los servicios públicos de atención abordados en este proyecto.

Los dos demostradores realizados estarán orientados a algunos servicios públicos de atención concretos, como por ejemplo:

- El servicio de información y gestión para la renovación del Documento Nacional de Identidad (DNI). En este caso, el sistema traducirá las informaciones y explicaciones que el funcionario público ofrece cuando atiende personalmente a una persona que quiere solicitar o renovar el DNI.
- El servicio de información sobre los centros de atención e información sobre las prestaciones económicas del Instituto Nacional de la Seguridad Social (INSS).
- Al igual que el servicio de información y gestión de los servicios que presta la DGT (Dirección General de Tráfico).

Los servicios finalmente considerados dependerán de las administraciones que los ofrecen de forma que nos permitan obtener la información necesaria para hacer un estudio lingüístico, y además, sea posible implantarlos en servicios reales para su evaluación.

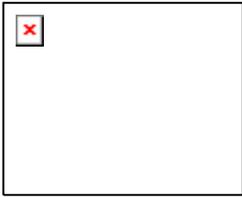
TAREAS A DESARROLLAR EN EL PROYECTO

Tarea 1. Estudio lingüístico de los servicios de atención abordados

En esta tarea se aborda el estudio lingüístico de las frases pronunciadas por el funcionario público en los dos servicios propuestos (o dominios de aplicación). Este estudio lingüístico supone la recopilación de las frases que con mayor frecuencia pronuncian los funcionarios en el servicio de atención personal. Consideraremos un número de frases adecuado de entre 500 y 1000 frases.

Una vez recopiladas estas frases, se deben generar las traducciones a LSE. Para ello, se estudiarán las diferentes posibilidades de representación de los signos de forma textual y se elegirá la que mayor flexibilidad ofrezca. Esta fase será la más costosa puesto que requiere la revisión manual de las frases en castellano, la búsqueda de los signos a los que se debe traducir y la posterior ordenación de los signos para generar la secuencia definitiva.





Participantes: Fundación CNSE (14 pm: personas x mes) y socio investigador En esta tarea la mayor carga de trabajo corresponderá a la Fundación CNSE en su labor de traducción de las frases recopiladas.

Resultados: el principal resultado de esta tarea es el corpus paralelo de frases en castellano y LSE.

Tarea 2. Diseño de la arquitectura del sistema

En esta tarea se propone el estudio y diseño de la arquitectura software que tendrá el sistema de traducción de voz a LSE.

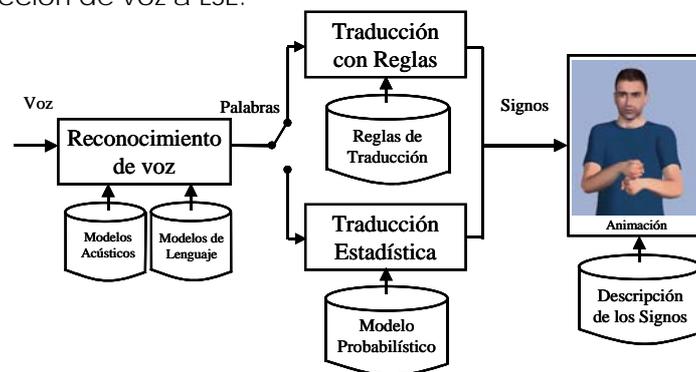


Figura 1. Etapas del sistema a desarrollar.

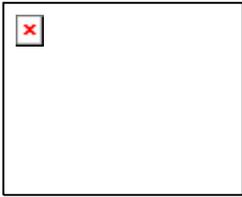
A grandes rasgos, el proceso de traducción se realizará en 3 pasos (Figura 1):

- En el primer paso, un reconocedor de voz que se encarga de traducir la voz en una secuencia de palabras. Este módulo permite reconocer habla en lenguaje natural (habla continua) e independiente del locutor.
- En segundo lugar, un módulo de traducción automática traduce la secuencia de palabras en una secuencia de signos. Para la implementación de este módulo se propondrán dos alternativas tecnológicas. La primera estrategia se basa en reglas de traducción desarrolladas por una persona experta. La segunda estrategia está basada en métodos estadísticos cuyos modelos se aprenden a partir de un corpus paralelo frases de texto-secuencias de signos.
- Finalmente, el módulo de animación de los signos basado en un agente animado en 3D (incorporado en el sistema como un control ActiveX). Este agente animado ha sido desarrollado en el proyecto eSIGN (<http://www.sign-lang.uni-hamburg.de/eSIGN/>) y está disponible para investigación.

Participante: Socio investigador y Fundación CNSE (1,5 pm).

Resultados: arquitectura software de traducción de voz a LSE.





Tarea 3: Adaptación del sistema de reconocimiento de voz

Esta tarea consiste en la adaptación del sistema de reconocimiento de voz disponible a los dominios de aplicación definidos en este proyecto. El sistema de reconocimiento de voz permite reconocer las frases pronunciadas por cualquier locutor/a hablando en castellano de forma continua (sin hacer pausas entre palabras), y genera a su salida la transcripción en texto de las palabras pronunciadas. Aunque el sistema de reconocimiento es muy flexible, se requiere de una adaptación a los dominios concretos de aplicación con el fin de conseguir las mejores tasas de reconocimiento. Esta adaptación consiste en:

- Utilizando el corpus de frases desarrollado en la Tarea 1, se genera el vocabulario de reconocimiento (conjunto de palabras), así como las transcripciones de cada una de las palabras consideradas. Con estas transcripciones se realiza un estudio sobre la confusión entre las diferentes palabras y se incorporan modificaciones específicas para reducir esta confusión.
- Con estas mismas frases, se generarán los modelos de lenguaje que se utilizarán en el proceso de reconocimiento. Estos modelos de lenguaje serán probabilísticos porque ofrecen un mejor compromiso entre flexibilidad y calidad de reconocimiento. Estos modelos recogen las probabilidades de las diferentes secuencias de palabras.

Una vez realizadas las labores de adaptación, se evaluarán las prestaciones finales del sistema de reconocimiento para los dos servicios planteados.

Participantes: Socio investigador

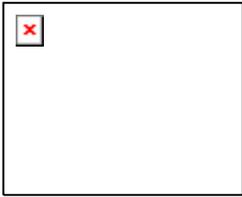
Resultados: sistema de reconocimiento de voz adaptado a dos servicios públicos de atención.

Tarea 4: Desarrollo del módulo de traducción basado en reglas de un experto

La primera propuesta de traducción considerada se basa en reglas de traducción desarrolladas por una persona experta. En este caso, las relaciones entre las palabras y los signos de LSE son definidas manualmente por una persona experta en un entorno de programación específico. Esta propuesta tecnológica está basada en una estrategia "bottom-up" a través de la cual, partiendo de las palabras reconocidas y tras un proceso de etiquetado se van generando conceptos semánticos. Estos conceptos se van combinando y generando signos, que a su vez se pueden recombinar con otros signos o conceptos para dar lugar a la secuencia de signos final. Este proceso conlleva dos fases:

- En una primera fase se realiza un proceso de categorización en el que cada una de las palabras se etiqueta con una o varias categorías sintáctico-semánticas.
- En una segunda fase se aplican un conjunto de reglas de traducción que permiten ir combinando las diferentes categorías para ir generando conceptos semánticos y signos. Estos conceptos podrán ser objeto de otras reglas que





realizarán cambios de orden en la secuencia y nuevas combinaciones hasta generar la secuencia de signos definitiva.

En esta tarea hay que realizar dos labores principales. La primera de ellas es desarrollar las herramientas software necesarias que permitan a los especialistas en LSE y el lingüista, definir las reglas necesarias para la traducción (aportación principal de Socio investigador). En segundo lugar, teniendo en cuenta el estudio lingüístico realizado en la tarea 1, hay que definir las categorías y las reglas de experto que permiten traducir una frase en forma de texto en una secuencia ordenada de signos (aportación combinada de Socio investigador y la Fundación CNSE). Una vez desarrollado el sistema de traducción se realizará una evaluación para medir sus prestaciones. El corpus de frases obtenido en la tarea 1 será un material muy importante para desarrollar y evaluar el módulo de traducción basado en reglas.

Participantes: Fundación CNSE (14 pm) y Socio investigador

Resultados: sistema de traducción (basado en reglas) de texto a LSE adaptado a los servicios.

Tarea 5: Desarrollo del módulo de traducción estadística

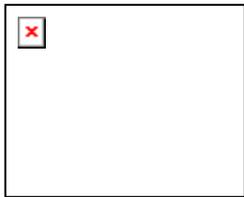
El sistema de traducción basado en reglas (descrito en la tarea anterior) es la mejor solución tecnológica para dominios restringidos como son los que se proponen en estos dos servicios de atención. En este proyecto se plantea también una investigación de nuevas técnicas de traducción entre lenguas basadas en modelos estadísticos. El motivo de abordar esta propuesta tecnológica es porque esta solución es muy interesante cuando los dominios de aplicación aumentan tanto su variedad lingüística que las propuestas basadas en reglas de experto suponen un coste muy elevado.

La traducción estadística consiste en un algoritmo de búsqueda dinámica que utiliza un modelo estadístico para obtener la mejor secuencia de signos resultado de la traducción de una secuencia de palabras obtenidas del reconocedor de voz. Este modelo integra principalmente información de dos tipos de probabilidades:

- Probabilidad de traducción: recoge información sobre qué palabras se traducen por qué signos.
- Probabilidad de la secuencia de signos: aporta información sobre qué secuencias de signos son más probables en la LSE.

En este tipo de sistemas de traducción, además del algoritmo de búsqueda, se puede realizar una etapa de preprocesado que mejora las prestaciones del sistema. En esta etapa de preprocesado se aplican, sobre las palabras de entrada, varias reglas de experto muy generales que facilitan enormemente la labor de traducción realizada por el algoritmo de búsqueda con lo que se aprovechará parte del trabajo y las reglas obtenidas en la tarea 4. Aunque este sistema de traducción será realizado principalmente por Socio investigador, se solicitará asesoría por parte de la Fundación CNSE para el diseño de esta etapa de preprocesado. Una vez desarrollado el sistema de traducción se realizará una evaluación de sus prestaciones finales. Al igual que ocurría en la tarea anterior el corpus de frases obtenido en la tarea 1 será un material





muy importante tanto para el entrenamiento de los modelos estadísticos como la posterior evaluación de la traducción estadística.

Participantes: Socio investigador y Fundación CNSE (8 pm)

Resultados: sistema de traducción estadística de texto a LSE adaptado a los servicios propuestos.

Tarea 6: Desarrollo de la base de datos de signos con una avatar en 3D

Para la representación visual de los signos se utilizará un agente animado en 3D que permita representar los signos con una gran calidad. Este agente animado ha sido desarrollado en el proyecto eSIGN (<http://www.sign-lang.uni-hamburg.de/eSIGN/>) y está disponible libremente para investigación.

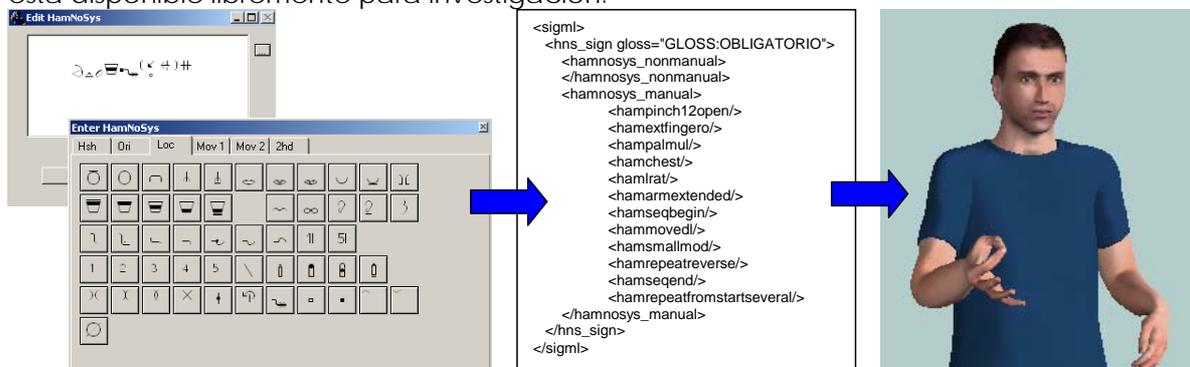


Figura 2. Proceso de generación de un signo a partir de su descripción

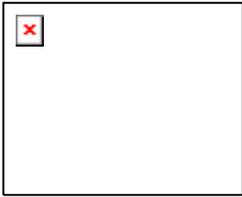
El proceso de generación de los signos se puede ver en la Figura 2. En primer lugar es necesario hacer un estudio de la producción de cada signo y generar en notación HamNoSys, la descripción detallada del signo. A partir de esta descripción se debe generar un script en lenguaje SIGML (Signing Gesture Markup Language). SIGML es una aplicación del lenguaje XML. Mediante este lenguaje se definen aspectos relacionados con la posición de las manos, la cabeza, la velocidad de signado, el tamaño del gesto, etc. Este script es interpretado por el agente animado que produce los signos con las características indicadas.

En esta tarea la labor del Socio investigador consistirá en la mejora del entorno para la caracterización y especificación de cada signo, mientras que la Fundación CNSE se encargará de utilizar este entorno de desarrollo para generar la base de datos con todos los signos necesarios para los demostradores a desarrollar en este proyecto. Esta labor de generación de los signos es la que supone el mayor esfuerzo dentro de esta tarea.

Participantes: Fundación CNSE (18 pm) y Socio investigador

Resultados: Entorno visual para el diseño y caracterización de signos, y base de datos de contenidos multimedia con las descripciones, características, significado y animación de los signos correspondientes a los servicios públicos de atención abordados en este proyecto.





Tarea 7: Evaluación final de los demostradores

Además de las evaluaciones parciales de cada uno de los módulos de los sistemas, en el proyecto se propone una evaluación externa final con potenciales usuarios que permita medir la utilidad de los demostradores desarrollados en los servicios concretos. En esta etapa final será responsabilidad del Socio investigador integrar los diferentes módulos desarrollados en las tareas anteriores y generar las versiones finales de los demostradores. En relación con la interfaz visual de los demostradores se contará con el asesoramiento de la Fundación CNSE para ofrecer la mayor ergonomía posible a las personas sordas que puedan utilizar el sistema. Por otro lado la Fundación CNSE será la encargada de contactar con grupos de personas sordas que evalúen el sistema desde el punto de vista del usuario final.

Para la evaluación de los demostradores se solicitará ayuda a las administraciones que están ofreciendo actualmente estos servicios de atención para que nos permitan montar los demostradores en sus instalaciones (durante algunos días) y así realizar la evaluación de campo. Durante esta evaluación se recopilarán medidas automáticas de aciertos del sistema recogidas por el propio demostrador y medidas subjetivas obtenidas de un cuestionario que rellenarán las personas que utilicen el sistema: tanto los funcionarios públicos que hablen al sistema, como las personas sordas usuarios del servicio.

La presentación y análisis de las medidas de evaluación se recogerán en el informe final del proyecto, incluyendo una reflexión sobre qué partes de los sistemas necesitan un esfuerzo de mejora mayor.

Participantes: Socio investigador y Fundación CNSE (9 pm)

Resultados: los demostradores de traducción de voz a LSE para los servicios planteados, así como su evaluación por funcionarios públicos y por usuarios finales.

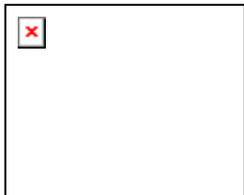
Tarea 8: Gestión y documentación

En esta tarea se incluyen los esfuerzos de gestión del proyecto y de documentación. Aunque el representante del proyecto sea Socio investigador la Fundación CNSE colaborará en la gestión del proyecto principalmente en la partida presupuestaria asignada a dicha fundación.

En relación con el proceso de documentación se prevé la generación de los siguientes documentos:

- Estudio lingüístico para los dos servicios analizados en este proyecto. Este estudio recogerá también una descripción detallada del corpus paralelo de frases en castellano y secuencias de signos generado en la tarea 1.
- Manual de usuario del entorno para la definición de las reglas de experto en el módulo de traducción basado en reglas. Esta descripción será muy útil para que expertos de la Fundación CNSE puedan especificar las reglas de traducción.
- Informe de resultados de traducción conseguidos con las dos estrategias de traducción evaluadas: traducción por reglas y traducción estadística.
- Manual de usuario del entorno para el diseño y especificación de los signos.





- Informe con la descripción detallada de la base de datos multimedia de signos generada: características, diseño y animaciones de los signos.
- Generación y documentación de los contenidos digitales en los formatos adecuados para su fácil distribución.
- Memoria de resultados y estado del proyecto del primer año.
- Memoria final de resultados del proyecto.
- Participantes: Socio investigador y Fundación CNSE (2 pm)
- Resultados: los documentos comentados anteriormente.

PLANIFICACIÓN DE TAREAS

La duración del proyecto planteado es de 15 meses y la planificación de las tareas se refleja en la Tabla 1. El proyecto iría desde el **1 de octubre de 2007** hasta el **31 de diciembre de 2008**.

Tareas	T0	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14
T1: Estudio Lingüístico	■	■	■	■											
T2: Diseño Arquitectura				■	■	■	■	■	■	■	■	■	■	■	■
T3: Reconocimiento de voz			■	■	■	■	■	■	■	■	■	■	■	■	■
T4: Traducción por reglas				■	■	■	■	■	■	■	■	■	■	■	■
T5: Traducción estadística					■	■	■	■	■	■	■	■	■	■	■
T6: Generación de signos						■	■	■	■	■	■	■	■	■	■
T7: Evaluación prototipos												■	■	■	■
T8: Gestión/documentación	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■

Tabla 1. Planificación de tareas del proyecto

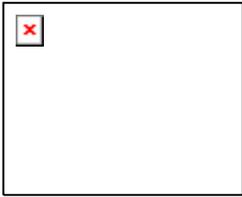
RESULTADOS DEL PROYECTO PREVISTOS Y CONTENIDOS DIGITALES A GENERAR

En relación con los principales resultados científicos y tecnológicos esperados cabe destacar:

- Una arquitectura software para la traducción automática de voz a LSE.



www.fundacioncnse.org
 C/ Islas Aleutianas, 28. 28035 Madrid
 Tel: 91 376 85 60. Fax: 91 376 85 64
 e-mail: proyectos@fundacioncnse.org



- Dos demostradores para la traducción de voz a LSE que permitan traducir las frases de un funcionario en servicios público de atención personal: el servicio de información y gestión para la renovación del Documento Nacional de Identidad (DNI) y el servicio de información sobre y el servicio de información y gestión de los servicios que presta la DGT (Dirección General de Tráfico), el servicio de información sobre los centros de atención e información sobre las prestaciones económicas del Instituto Nacional de la Seguridad Social (INSS), junto con el servicio de información y gestión de los servicios que presta la DGT (Dirección General de Tráfico).
- Herramientas de ayuda a la traducción entre el castellano y la LSE con el fin de facilitar la labor de los intérpretes en la generación de contenidos visuales para personas sordas. En concreto se obtendrán dos sistemas para la generación de propuestas de traducción que posteriormente son corregidas por los intérpretes, y un entorno para el diseño, caracterización y generación de las animaciones de los signos mediante un agente animado virtual.
- En el último año, los grupos implicados en este proyecto coordinado han invertido un esfuerzo muy importante en realizar publicaciones en revistas con factores de impacto importantes, habiendo conseguido hasta 5 nuevos artículos en revista. En este nuevo proyecto y dado que todas las tecnologías presentadas son innovadoras y muy prometedoras, se va a fomentar la propuesta de nuevas patentes o modelos de utilidad así como la publicación de los resultados en las principales revistas con factores de impacto, y en los congresos más reconocidos.

En relación con los contenidos multimedia a generar cabe destacar los siguientes:

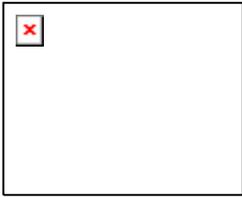
- Un corpus paralelo con las frases en castellano y su traducción en LSE. Estos contenidos serán de gran importancia para realizar futuros estudios de traducción entre ambas lenguas.
- Generación de una base de datos de contenidos multimedia con las descripciones, características, significado y animación de los signos correspondientes a los servicios públicos de atención abordados en este proyecto.

PLAN DE EXPLOTACIÓN DE RESULTADOS

En relación con los demostradores implementados se realizará una labor importante de evaluación con el fin de analizar los aspectos que hay que mejorar de cara a instalar de forma definitiva los sistemas de traducción desarrollados como apoyo a los servicios públicos considerados.

Este proceso de implantación de los demostradores en servicios reales conlleva las siguientes fases:





1. En primer lugar, se deben analizar y solventar los posibles problemas de los demostradores detectados en la evaluación final del proyecto.
2. Una vez resueltos los problemas se generarán nuevas versiones de los demostradores.
3. Seguidamente, se establecerán los contactos necesarios para incorporar los demostradores en algunos servicios puntuales bajo una gran supervisión.
4. Finalmente se propondrá una evaluación más extensa en duración y número de personas que lo utilicen con el fin de generar una versión final de los sistemas que se pueda implantar en una gran cantidad de administraciones.

En relación con las herramientas y entornos de desarrollo estarán a disposición de la Fundación CNSE para generar nuevos contenidos en LSE y estarán disponibles para el desarrollo de nuevos traductores adaptados a otros servicios de atención personal de la administración

El corpus paralelo con las frases en castellano y su traducción en LSE se podrán utilizar para realizar futuros estudios de traducción entre ambas lenguas.

Finalmente, en relación con la base de datos de contenidos multimedia con las descripciones, características, significado y animación de los signos estará disponible para el desarrollo de nuevas aplicaciones de aprendizaje de la LSE o el desarrollo de nuevos traductores adaptados a otros servicios de atención personal de la administración.

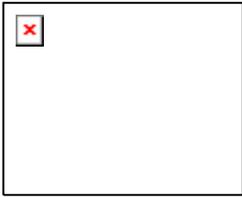
DIFUSIÓN DE RESULTADOS Y CONOCIMIENTO

Los demostradores que se obtendrán como resultado del trabajo en esta investigación serán propuestos para la presentación pública en todos aquellos eventos que la Universidad Politécnica de Madrid o la Fundación CNSE así lo requieran.

Adicionalmente, se realizarán gestiones con los medios de comunicación para que se difunda a la sociedad los resultados obtenidos, con especial énfasis en la exposición pública del funcionamiento de los demostradores realizados. Además de los mecanismos de comunicación de la UPM y de la Fundación CNSE a través de sus revistas de divulgación, se contactará con medios de comunicación de ámbito nacional que difundan los logros del proyecto. También para la difusión del proyecto al movimiento asociativo se ha previsto el envío puntual de circulares informativas y folletos divulgativos a las distintas federaciones y asociaciones de personas sordas del país, así como a diversos centros educativos relacionados con la LSE y/o la educación de personas sordas. Por otro lado, también se dará cuenta de este proyecto en la Memoria de Gestión de Fundación CNSE del año 2007, que se distribuirá entre entidades públicas y privadas con las que mantiene contacto la entidad.

Como ya se ha venido haciendo hasta ahora, el proyecto estará presente en todos aquellos foros relacionados con la materia en los que la entidad participe.





El objetivo es que este proyecto sea una herramienta de referencia y utilidad entre los usuarios de la LSE contribuyendo tanto a su normalización como a la estandarización del léxico. Por todo ello habrá que realizar una difusión que llegue a los máximos medios hacer mella en la población en general, no solo en el entorno de las personas sordas. Al igual que en todos los proyectos anteriores se generarán páginas web con toda la información del proyecto: documentación, publicaciones, bases de datos, herramientas software,...

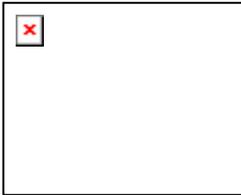
Finalmente, una labor importante de divulgación del conocimiento será la publicación de artículos científicos en revistas de investigación con factor de impacto elevado y en congresos de reconocido prestigio. Esta labor constituye una de las actividades más importantes de las universidades y en concreto de la UPM miembro del proyecto coordinado que se presenta.

ADECUACIÓN A LOS OBJETIVOS DE LA PROPUESTA

En la Tabla 2 se resume la adecuación del proyecto a los objetivos de la propuesta

Criterio	Justificación
Identificación de los objetivos del proyecto y su adecuación hacia la consecución de resultados explícitamente identificados	En la presente propuesta se ha incluido un apartado específico en el que se detalla el objetivo principal y los subobjetivos que se proponen. Estos subobjetivos están directamente relacionados con las tareas que se plantean para la realización del proyecto.
Calidad y viabilidad técnica, económica y financiera del proyecto	En esta propuesta se ha incluido una memoria técnica ampliamente documentada, junto con una memoria económica y financiera muy completa en la que se han justificado detalladamente todos los gastos necesarios para la realización del proyecto.
Plan de explotación de resultados del proyecto	Se han incluido un apartado específico de resultados previstos y otro apartado de explicación del plan de explotación donde se detallan las acciones a realizar en relación con cada uno de los resultados previstos del proyecto.
Grado de innovación del proyecto	El proyecto propone la realización dos demostradores que permitan la traducción de voz a LSE. A nivel internacional no existe ningún sistema ni demostrador que aborde esta tarea de traducción entre estas dos lenguas. Otros sistemas similares se han desarrollado para traducir el inglés en la Lengua de Signos Inglesa pero muy orientados a servicios muchos más restringidos que los que se proponen abordar en este proyecto. Además, en este caso se propone su aplicación sobre un servicio real de la administración.
Capacidad para difundir los resultados y conocimientos	El consorcio formado por la Fundación CNSE y la Universidad Politécnica de Madrid es una garantía muy importante de que los resultados y conocimientos generados con este proyecto tendrán la mayor difusión posible y podrán ser utilizados en nuevos proyectos fruto de la colaboración entre estas dos





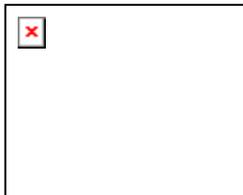
	entidades. La labor de difusión es una de las principales actividades de los miembros de este consorcio.
Impacto socio-económico	El impacto socio-económico es muy importante puesto que se trata del desarrollo de sistemas que permitan a la administración pública ofrecer servicios personales de mayor calidad a las personas sordas. Permitiendo que este colectivo pueda acceder con una mayor facilidad a dichos servicios. Por otro lado, el desarrollo de sistemas de traducción automática permitiría reducir sensiblemente la contratación de intérpretes con la reducción en coste que eso supone.
Impacto territorial y demográfico	En relación con el impacto territorial, los servicios propuestos son comunes a todas las comunidades autónomas (gestión del DNI y y el servicio de información sobre los centros de atención e información sobre las prestaciones económicas del Instituto Nacional de la Seguridad Social (INSS). El hecho de considerar el castellano como lengua origen para la traducción define el ámbito de aplicación a las comunidades autónomas que utilicen esta lengua en los servicios oficiales de la administración. El colectivo de usuarios de LSE en España ronda las 400.000 personas, de las cuales 100.000 son personas sordas usuarias de LSE
Integración de medidas específicas que fomenten la igualdad de género	Aunque en este proyecto no se han incluido medidas específicas para fomentar la igualdad de género, dicha igualdad ha sido uno de los principales criterios a la hora de planificar el actual proyecto considerando dicha igualdad en las posibilidades de uso de los demostradores desarrollados como en la conformación del equipo investigador del proyecto.

Tabla 2. Adecuación a los objetivos de la propuesta.

Referencias

- Abdel-Fattah, M. A. 2005 "Arabic Sign Language: A perspective". In: Journal of Deaf Studies and Deaf Education 10: 2 (2005) - pp. 212-221.
- Atherton, M. 1999 "Welsh today BSL tomorrow". In: Deaf Worlds 15: 1 (1999) - pp. 11-15
- Bertenstam, J., et al. 1995. "The Waxholm system-A progress report". In Proc. Spoken Dialogue Systems, Vigso, Denmark, 1995.
- Bungeroth, J. et. al. 2006. "A German Sign Language Corpus of the Domain Weather Report". 5th Int. Conf. on Language Resources and Evaluation, Genoa, Italy, 2006.
- Casacuberta F., E. Vidal. 2004. "Machine Translation with Inferred Stochastic Finite-State Transducers". Computational Linguistics, Vol. 30, No. 2, pp. 205-225, June 2004.
- Cassell, J., Stocky, T., Bickmore, T., Gao, Y., Nakano, Y., Ryokai, K., Tversky, D., Vaucelle, C., Vilhjálmsson, 2002. "MACK: Media lab Autonomous Conversational Kiosk" in Proc. Of Imagina: Intelligent Autonomous Agents, Monte Carlo, Monaco, 2002.
- Christopoulos, C. Bonvillian, J., 1985. "Sign Language". Journal of Communication Disorders, 18 1-20. Amsterdam. 1985.
- Cole, R., et al. 1999. "New Tools for interactive speech and language training: Using animated conversational agents in the classrooms of profoundly deaf children" in Proc. ESCA/SOCRATES Workshop on Method and Tool Innovations for Speech Science Education, London, 1999, pp. 45-52.
- Cole, R., Van Vuuren, S., Pellom, B., Hacıoglu, K., Ma, J., Movellan, J., Schwartz, S., Wade-Stein, D., Ward, W., Yan, J., (2003). "Perceptive Animated Interfaces: First Steps Toward a New Paradigm for Human





Área de Programas y Proyectos



- Computer Interaction", in IEEE Transactions on Multimedia: Special Issue on Human Computer Interaction, vol. 91, no. 9, pp. 1391-1405. 2003.
- Cox, S.J., M. Lincoln, J Tryggvason, M Nakisa, M. Wells, Mand Tutt, and S Abbott. TESSA, a system to aid communication with deaf people. In ASSETS 2002, pages 205-212, Edinburgh, Scotland, 2002.
- Engberg-Pedersen, E. 2003 "From pointing to reference and predication: pointing signs, eyegaze, and head and body orientation in Danish Sign Language". In: Kita, Sotaro (ed): Pointing : where language, culture, and cognition meet. Mahwah, NJ : Erlbaum (2003) - pp. 269-292
- Gallardo, B., Montserrat, 2002. V., "Estudios Lingüísticos sobre la lengua de signos española" Universidad de Valencia. Ed. AGAPEA ISBN: 8437055261. ISBN-13: 9788437055268
- Granström, B., House, D., Beskow, J., 2002. "Speech and Signs for Talking Faces in Conversational Dialogue Systems" Multimodality in Language and Speech Systems. Kluwer Academic Publishers. pp 209-241. 2002.
- Gustafson, J., 2002. "Developing multimodal spoken dialogue systems- Empirical studies of spoken human-computer interactions". Ph D. Dissertation. Dept. Speech, Music and Hearing, Royal Inst. of Technology, Stockholm, Sweden, 2002.
- Gustafson, J., and Bell, L., 2003. "Speech technology on trial: Experiences from the August system". Journal of Natural Language Engineering: Special Issue on Best Practice in Spoken Dialogue Systems. pp. 273-286. 2003.
- Herrero-Blanco, A., Salazar-Garcia, V. 2005 "Non-verbal predicability and copula support rule in Spanish Sign Language". In: de Groot, Casper / Hengeveld, Kees (eds): Morphosyntactic expression in functional grammar. (Functional Grammar Series; 27) Berlin [u.a.] : de Gruyter (2005) - pp. 281-315
- Huenerfauth. M. 2004. A multi-path architecture for machine translation of English text into American Sign language animation. HLT-NAACL, Boston, MA, USA.
- Koehn P., F.J. Och D. Marcu. 2003 "Statistical Phrase-based translation". Human Language Technology Conference 2003 (HLT-NAACL 2003), Edmonton, Canada, pp. 127-133, May 2003.
- Lundeberg, M., and Beskow, J., 1999. "Developing a 3D-agent for the August dialogue system" in Proc. Audio-Visual Speech Processing, Santa Cruz, CA, 1999.
- Masataka, N et al 2006. "Neural correlates for numerical processing in the manual mode". In: Journal of Deaf Studies and Deaf Education 11: 2 (2006) - pp. 144-152
- Meurant, L. 2004 "Anaphora, role shift and polyphony in Belgian sign language". Poster. In: TISLR 8 Barcelona, September 30 - October 2. Programme and Abstracts. (International Conference on Theoretical Issues in Sign Language Research; 8) (2004) - pp. 113-115
- Morrissey. S and A. Way. 2005. An example-based approach to translating sign language. In Workshop Example-Based Machine Translation (MT X-05), pages109-116, Phuket, Thailand, September.
- Nyst, V. 2004 "Verbs of motion in Adamorobe Sign Language". Poster. In: TISLR 8 Barcelona, September 30 - October 2. Programme and Abstracts. (International Conference on Theoretical Issues in Sign Language Research; 8) (2004) - pp. 127-129.
- Och J., Ney. H., 2002. "Discriminative Training and Maximum Entropy Models for Statistical Machine Translation". Annual Meeting of the Ass. For Computational Linguistics (ACL), Philadelphia, PA, pp. 295-302. 2002.
- Och J., Ney. H., 2003. "A systematic comparison of various alignment models". Computational Linguistics, Vol. 29, No. 1 pp. 19-51, 2003.
- Papineni K., S. Roukos, T. Ward, W.J. Zhu. 2002 "BLEU: a method for automatic evaluation of machine translation". 40th Annual Meeting of the Association for Computational Linguistics (ACL), Philadelphia, PA, pp. 311-318. 2002.
- Pyers J.E., 2006. "Indicating the body: Expression of body part terminology in American Sign Language". Language Sciences, In Press, Corrected Proof, Available online 4 January 2006.
- Reyes, I. 2005 "Comunicar a través del silencio :las posibilidades de la lengua de signos española". Sevilla : Universidad de Sevilla 2005 - 310 p.
- Rodríguez, M.A. 1991 "Lenguaje de signos" Phd Dissertation. Confederación Estatal de Personas Sordas (CNSE) and Fundación ONCE. Madrid. Spain. 1991
- Stokoe, W., 1960. "Sign Language structure: an outline of the visual communication systems of the American deaf". Studies in Linguistics. Buffalo, Univ. Paper 8. 1960.
- Sutton, S., Cole, R., 1998. "Universal speech tools: the CSLU toolkit" in Proc of Intl. Conf. On Spoken Language Processing, pp. 3221-3224, Sydney, Australia. 1998.

