

Tecnología para la mejora de la comunicación entre una persona sorda y una persona oyente

Realizado por:



Grupo de Tecnología del Habla
Universidad Politécnica de Madrid



Financiado por:



Plan Avanza del Ministerio de
Industria, Turismo y Comercio



Fundación ONCE

Rubén San Segundo Hernández

Índice

- Objetivo del proyecto
- Estado del arte
- Corpus paralelo
- Traducción de voz a LSE (Lengua de Signos Española)
- Editor de signos
- Generación de voz a partir de signos
- Evaluación final
- Resumen de resultados y líneas futuras
- Difusión del proyecto



Objetivo del proyecto

Diseño, desarrollo y evaluación de una **sistema avanzado de comunicación bidireccional** entre una persona oyente y una persona sorda. Aplicado a un servicio de atención personal de la administración.

Dos servicios elegidos:

- ☐ Renovación del Documento Nacional de Identidad (DNI).
- ☐ Renovación del permiso de conducir en la DGT (Dirección General de Tráfico).



Estado del arte (I)

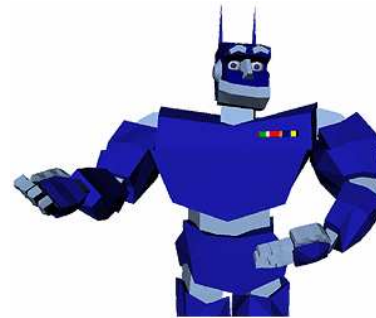
- Traducción automática:
 - Métodos basados en reglas
 - Transferencia: **Systran, METAL, Logos, WebSphere.**
 - Interlingua: **ROSETTA, KBMT, UNL, DLT, KANT-KANTOO-CATALYST.**
 - Métodos basados en corpus
 - Estadística: **CANDIDE, EGYPT, GIZA++, Pharaoh, Moses, Xerox, FSTs.**
 - Basada en ejemplos: **DÉJÀ-VU, OmegaT, Swordfish.**
- Proyectos de traducción voz a voz:
 - Internacionales: **TC-STAR, GALE.**
 - Nacionales: **AVIVAVOZ** (<http://www.avivavoz.es/>)



Estado del arte (II)

- Agentes animados virtuales
 - Gran variedad de servicios en distintos ámbitos.

Sistema August



Sistema MACK

- Sistemas de traducción de voz a Lengua de Signos

- SignTel
- Proyecto eSign
- Proyecto TESSA
- IBM - Sistema SiSi (“Say It, Sign It”)



Estado del arte (III)

- Corpus lingüísticos (Lengua de Signos):
 - Estudios sociolingüísticos de la **ASL** en 1995 (**Lucas**).
 - Proyecto **ECHO: BSL, NGT, SSL**.
 - Creación de corpus para **Auslan**.
 - Proyectos de creación de corpus parciales:
 - **ATIS, Phoenix, LSI, reconocimiento de signos** Bohemia.
 - Proyectos en desarrollo:
 - **BSL** en University College de Londres, California Berkeley, Orebro, Investigación Educativa en Noruega,
 - **LSE** en Universidad de Vigo e Instituto Cervantes (no corpus paralelo)



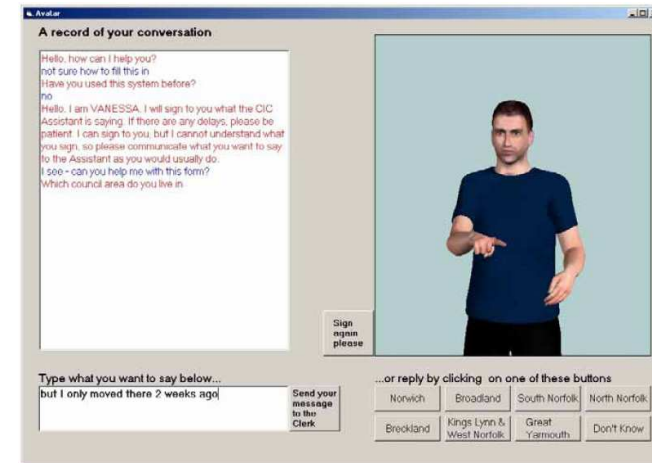
**Necesidad de
corpus en LSE**



Estado del arte (IV): LSE-voz

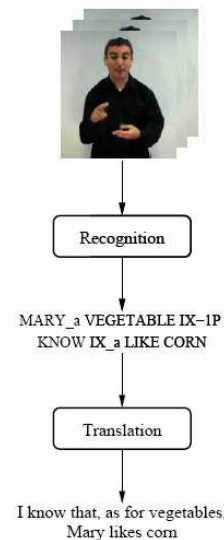
Proyecto **VANESSA**:

- Comunicación con funcionarios del ayuntamiento
- Interfaz simple y limitada:
 - Entrada de texto.
 - Botones con fórmulas básicas.

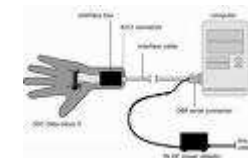


Proyecto **Hand in Hand**:

- Reconocimiento de signos:
 - Modelos Ocultos de Markov.
 - Trigramas.
- Métodos estadísticas de traducción.
- Experimentos:
 - Corpus 104 palabras.
 - Corpus 375 palabras.



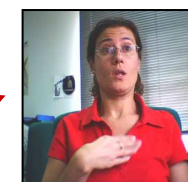
Guantes para el reconocimiento de lengua de signos



Corpus paralelo (I)



INDICE	DOMINIO	VENTANILLA	SERVICIO	TIPO	CASTELLANO	LSE	VIDEO
1	DL	CAJA	Decir la cantidad	Funcionario	ahora tiene que ir a la ventanilla de conductores	AHORA TU VENTANILLA ESPECÍFICO PERSONA CONDUCTOR IR-ALLÍ	1.wmv
2	DL	CAJA	Decir la cantidad	Funcionario	ahora vaya a la ventanilla de conductores	AHORA TU VENTANILLA ESPECÍFICO PERSONA CONDUCTOR IR-ALLÍ	2.wmv
3	DL	CAJA	Decir la cantidad	Funcionario	catorce con veinte	CATORCE COMA VEINTE EUROS	3.wmv
4	DL	CAJA	Decir la cantidad	Funcionario	catorce euros con veinte céntimos	CATORCE COMA VEINTE EUROS	4.wmv
5	DL	CAJA	Decir la cantidad	Funcionario	ocho euros	OCHO EUROS	5.wmv
6	DL	CAJA	Decir la cantidad	Funcionario	si usted tiene más de setenta años no tiene que	EJEMPLO TU EDAD SETENTA MÁS CARNET CONDUCIR RENOVAR PAGA	6.wmv
7	DL	CAJA	Decir la cantidad	Funcionario	son catorce con veinte	CATORCE COMA VEINTE EUROS	7.wmv
8	DL	CAJA	Decir la cantidad	Funcionario	son catorce euros con veinte céntimos	CATORCE COMA VEINTE EUROS	8.wmv
9	DL	CAJA	Decir la cantidad	Funcionario	son ocho euros	OCHO EUROS	9.wmv
10	DL	CAJA	Decir la cantidad	Funcionario	vaya directamente a conductores	TU VENTANILLA ESPECÍFICO PERSONA CONDUCTOR IR-ALLÍ DIRECTO	10.wmv
11	DL	CAJA	Decir la cantidad	Funcionario	vaya directamente a ventanilla de conductores	TU VENTANILLA ESPECÍFICO PERSONA CONDUCTOR IR-ALLÍ DIRECTO	11.wmv
12	DL	CAJA	Decir la cantidad	Funcionario	ve directamente a conductores	TU VENTANILLA ESPECÍFICO PERSONA CONDUCTOR IR-ALLÍ DIRECTO	12.wmv
13	DL	CAJA	Decir la cantidad	Funcionario	ve directamente a la ventanilla de conductores	TU VENTANILLA ESPECÍFICO PERSONA CONDUCTOR IR-ALLÍ DIRECTO	13.wmv
14	DL	CAJA	Horarios	Funcionario	el horario de la ventanilla de caja es de nueve a	VENTANILLA ESPECÍFICO DINERO PAGAR HORA HORARIO ABRIR NUEV	14.wmv
15	DL	CAJA	Horarios	Funcionario	el resto de las ventanillas cierran a las dos	VENTANILLA LOS-DEMÁS CERRAR HORA DOS	15.wmv
16	DL	CAJA	Ir después	Funcionario	por favor vaya a la ventanilla de conductores	TU VENTANILLA ESPECÍFICO PERSONA CONDUCTOR IR-ALLÍ	16.wmv
17	DL	CAJA	Ir después	Funcionario	por favor ve a la ventanilla de conductores	TU VENTANILLA ESPECÍFICO PERSONA CONDUCTOR IR-ALLÍ	17.wmv

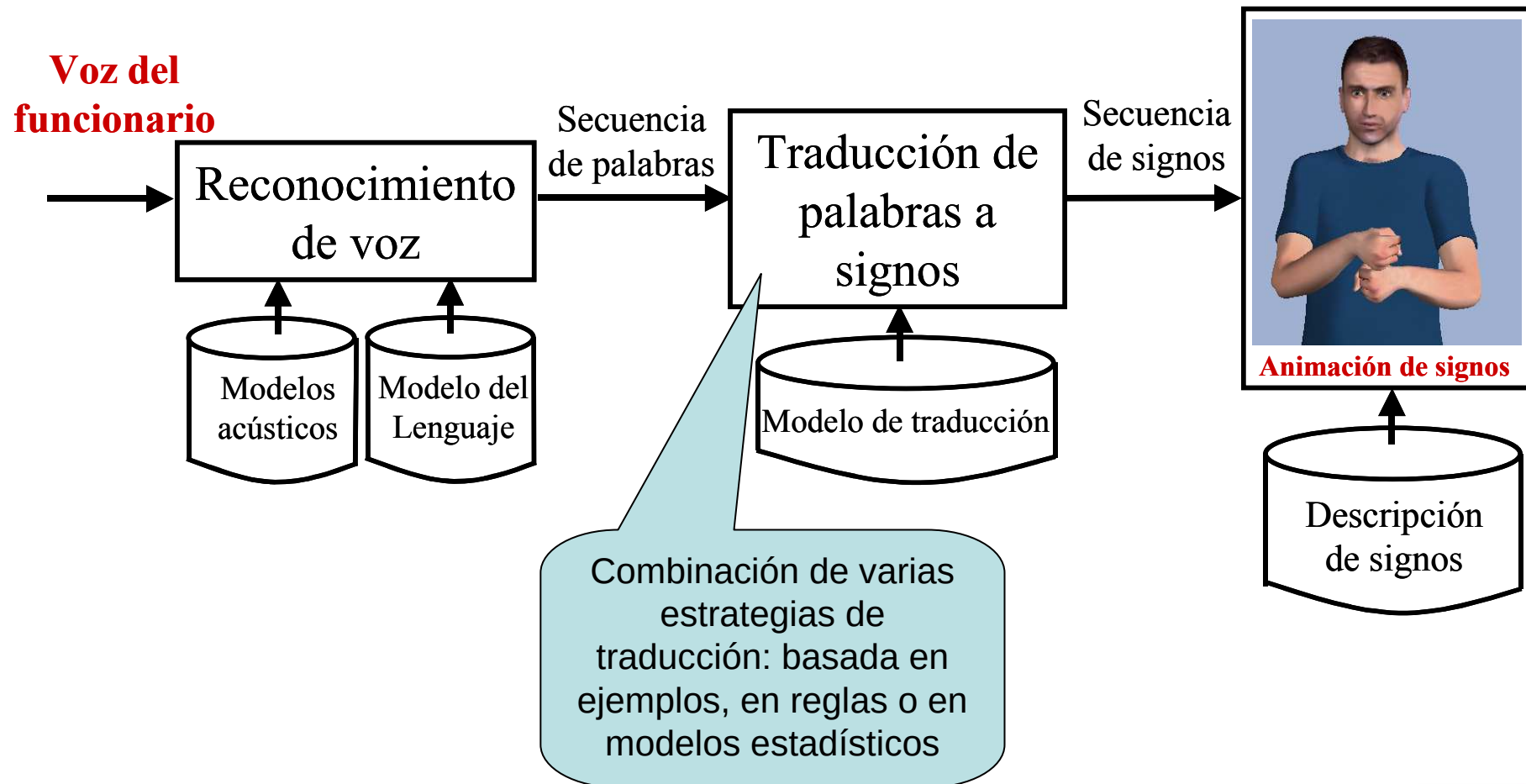


Corpus paralelo (II)

	DNI		DGT	
Funcionario	Español	LSE	Español	LSE
Pares de frases	1,425		1,641	
Frases diferentes	1,236	389	1,413	199
Palabras	8,490	6,282	17,113	12,741
Vocabulario	652	364	527	237
Usuario	Español	LSE	Español	LSE
Pares de frases	531		483	
Frases diferentes	458	139	389	93
Palabras	2,768	1,950	3,130	2,283
Vocabulario	422	165	294	133

Alrededor de 4000 pares de frases

Traducción de voz a LSE (I)



Traducción de voz a LSE (II)

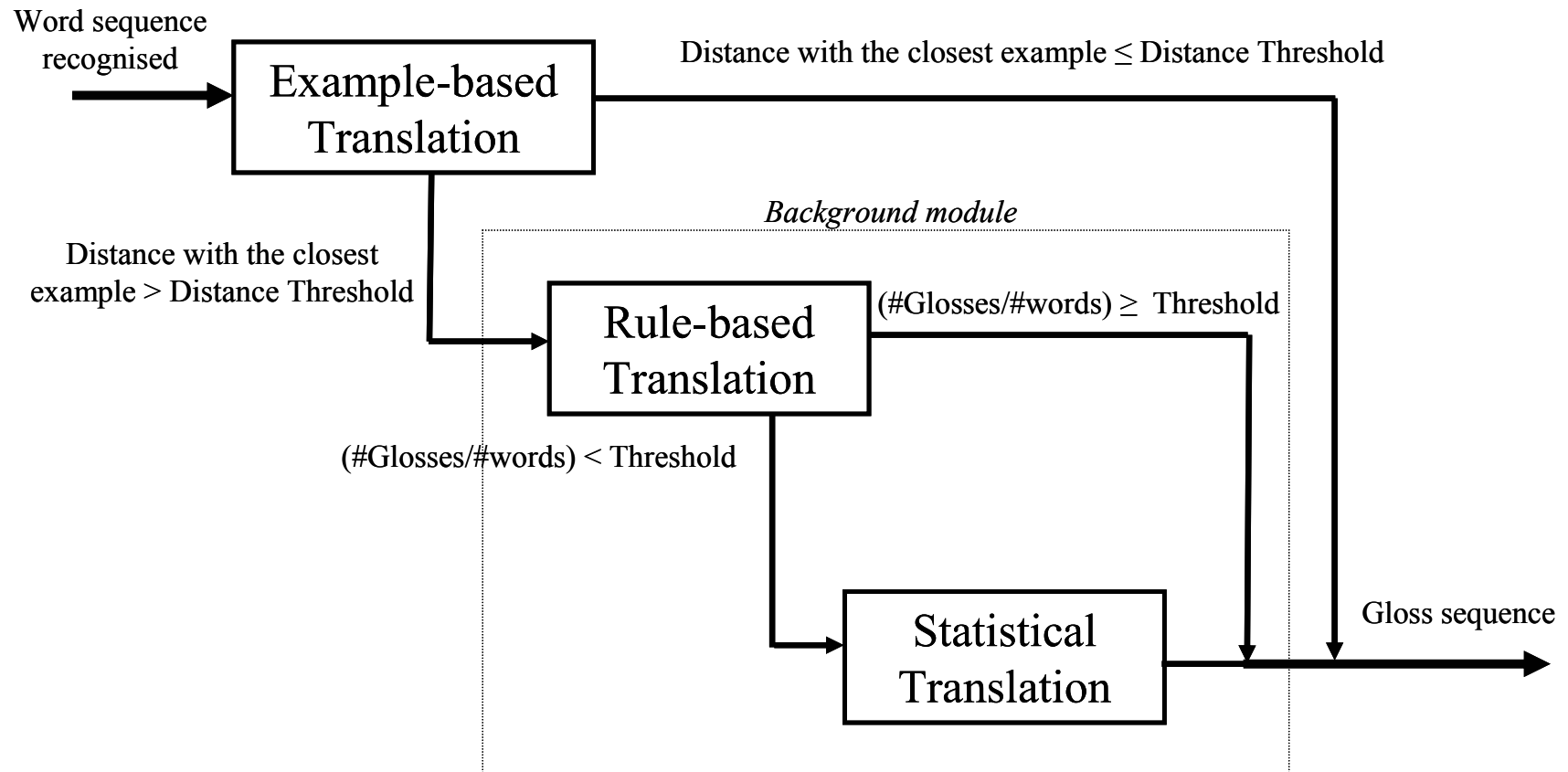


Reconocimiento de voz

- Reconocedor del Grupo de Tecnología del Habla (GTH-UPM)
 - Habla Continua.
 - Independiente de locutor.
 - Medidas de confianza.
 - Ofrece la posibilidad de adaptar los modelos acústicos a un entorno acústico o a un grupo de locutores.
 - 5760 modelos de trifenema.
 - Modelos de ruido.
 - Modelo de lenguaje N-gram adaptado a la tarea.



Traducción a LSE



Traducción basada en ejemplos

Ejemplos

Castellano	LSE
el denei es obligatorio desde los \$NUMERO años	DNI OBLIGATORIO DESDE \$NUMERO EDAD
puede venir a recoger el denei el \$DIA por la \$P_DIA	TÚ DNI RECOGER \$DIA \$P_DIA PODER
necesita \$NUMERO fotografías	TÚ \$NUMERO FOTOGRAFÍA NECESITAR
venga el \$NUMERO de \$MES	TÚ \$NUMERO \$MES VENIR
...	...

Distancia entre
ejemplos

“DNI OBLIGATORIO DESDE \$NUMERO EDAD”

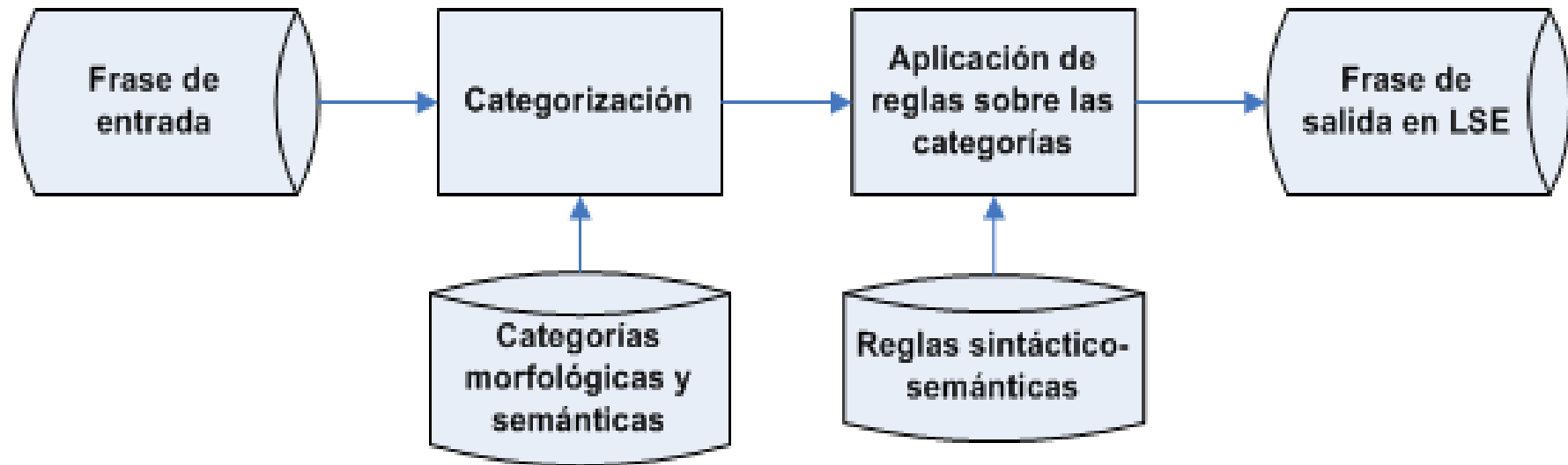
“DNI OBLIGATORIO DESDE CATORCE EDAD”

“denei obligatorio de los \$NUMERO años”

categorización

“denei obligatorio de los catorce años”

Traducción basada en reglas



Traducción por reglas: categorización

- Se asignan una o más categorías por palabra.
- Las reglas operan con bloques de categorías.
- La primera categoría de cada bloque son uno o más signos. Las siguientes contienen información sintáctico-semántica en base a la cual se aplicarán reglas.
- Categoría “basura”: bloques no relevantes para la traducción.
- Las palabras desconocidas no son basura. Podrían ser nombres propios a representar por deletreo.



Traducción por reglas: aplicación de reglas

Primitivas

- Reglas diseñadas en base a primitivas: operaciones sobre los bloques.
- Utilidad de las primitivas:
 - Reemplazar bloques con categorías específicas.
 - Cambiar el orden de los bloques.
 - Detectar, añadir o eliminar categorías.
 - Otros (ej: separar bloques que correspondan a varios signos para la salida).

Tipos de reglas utilizadas

- Cuatro fases:
 - Traducción de perífrasis y algunas fórmulas concretas, y manejo de basuras.
 - Agrupamiento de bloques y preparación para su ordenamiento, proceso de expresiones en las que habrían intervenido las basuras, y combinación de signos para obtener otros más complejos.
 - Ordenamiento de bloques según la gramática de la LSE.
 - Ajustes finales.



Traducción estadística

Base de Datos

La base de datos contiene 414 frases de un contexto restringido: las que diría un funcionario al dar información sobre renovación del DNI. El conjunto de frases se divide aleatoriamente en 3 grupos:

- Un primer grupo de entrenamiento, con el 75% de las frases
- Un segundo grupo de evaluación, con el 12,5% de las frases
- Un grupo de test, con el 12,5% de frases

		Castellano	LSE
Total	Pares de frases	414	
	Número de palabras o signos	4847	4564
Entrenamiento	Pares de frases	314	
	Número de palabras o signos	3071	2876
Validación	Pares de frases	50	
	Número de palabras o signos	582	574
Test	Pares de frases	50	
	Número de palabras o signos	570	505

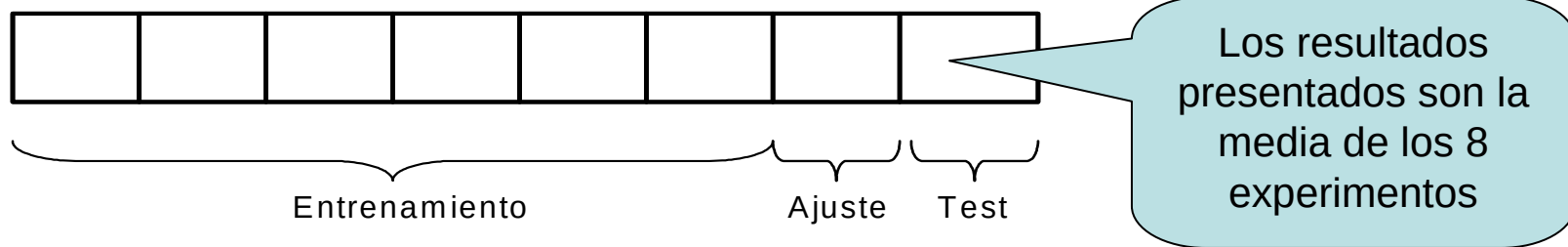
Estadísticas de ejemplo para una de las listas



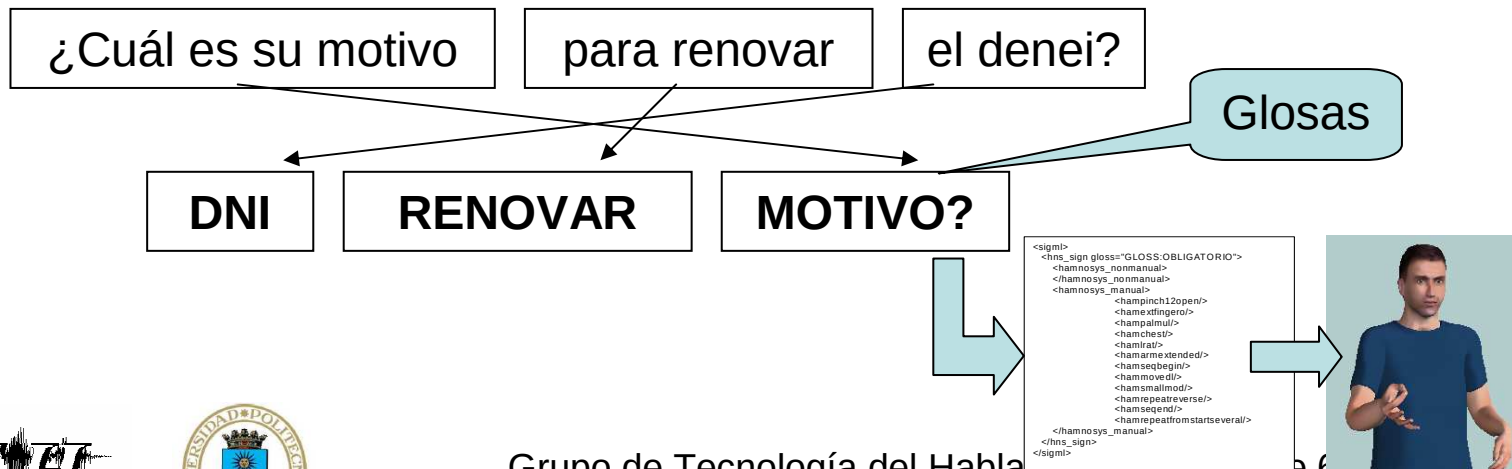
Traducción estadística

Base de Datos

- o Se divide el corpus paralelo en 8 listas de 50 frases cada una.
- o Con una estrategia Round-Robin, se elegirán cada vez 6 de ellas para Entrenamiento, 1 para Ajuste y 1 para Test



Ejemplo



Grupo de Tecnología del Habla

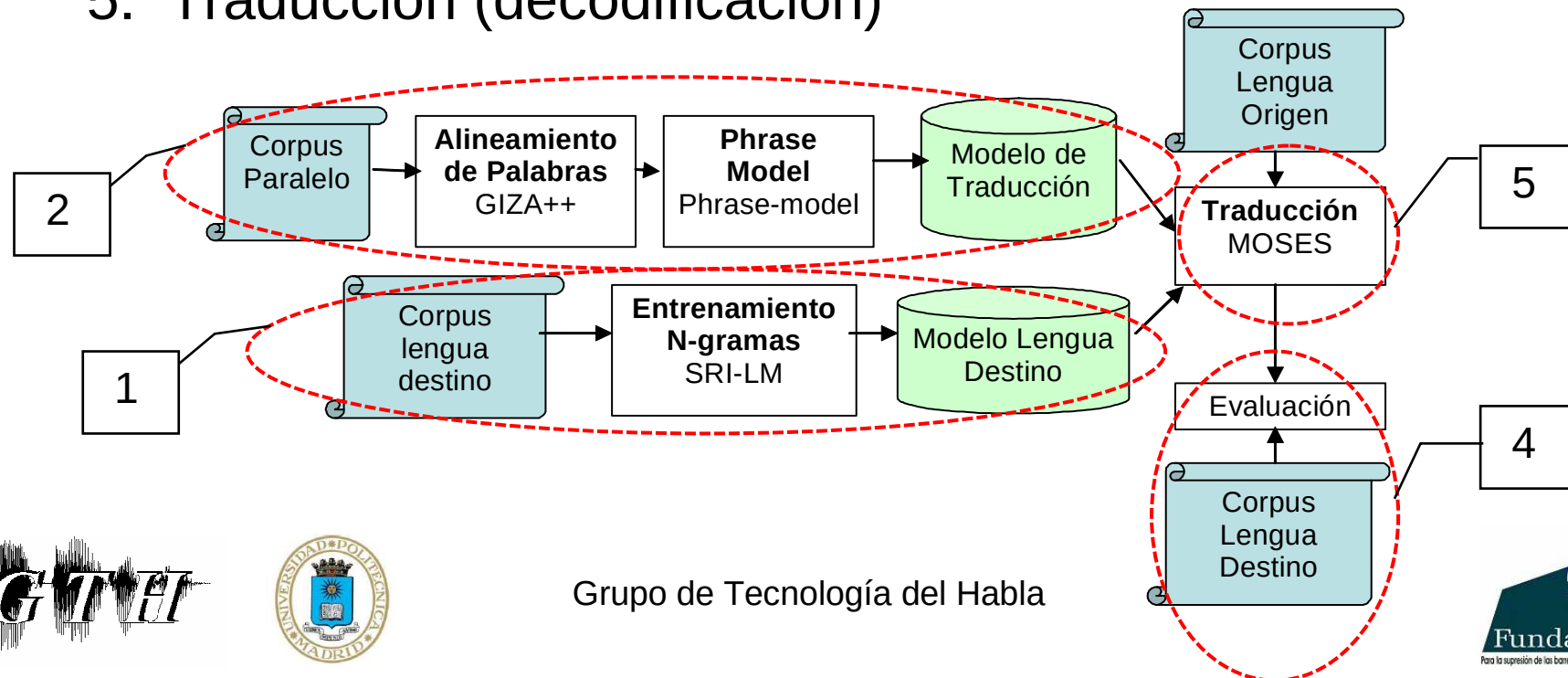


Traducción estadística

Subfrases

Visión General de la Arquitectura

1. Generación del Modelo de Lenguaje
2. Generación del Modelo de Traducción
3. Fase de Ajuste
4. Fase de Test
5. Traducción (decodificación)



Traducción estadística

Subfrases

Modelo de Lenguaje

Según el Teorema de Bayes, para encontrar la mejor traducción de una palabra:

$$\tilde{d} = \arg \max_{d \in d^*} p(d/o) = \arg \max_{d \in d^*} p(o/d) p(d)$$

o: idioma origen
d: idioma destino

Modelo de Traducción Modelo de Lenguaje

Se utiliza la herramienta SRI-LM para calcular el modelo de lenguaje mediante `ngram-count.exe`. Se calcula la probabilidad hasta trigramas.

Traducción estadística

Subfrases

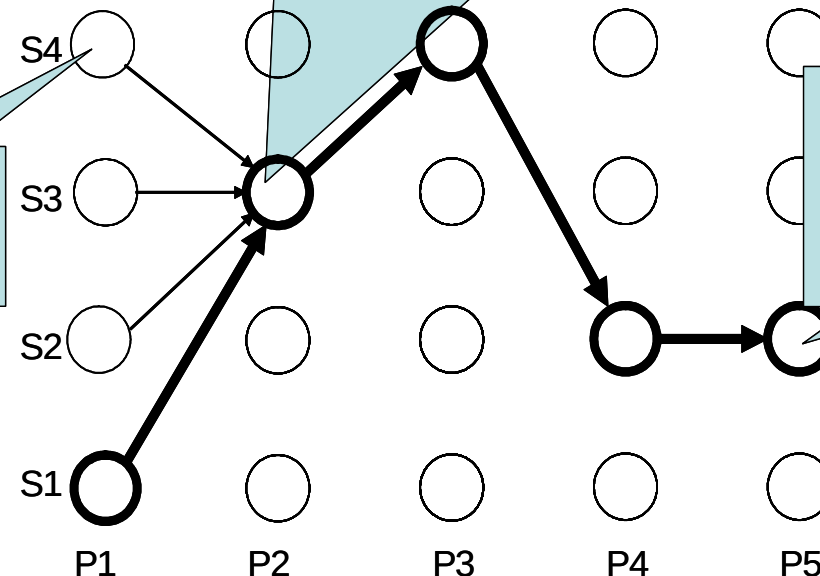
Modelo de Traducción

Se hace uso del script `train-factorized-phrase-model.perl`, que genera el modelo en los siguientes pasos:

1. Preparación del corpus y alineamiento **bidireccional** con el corpus de entrenamiento

Secuencia de Signos (i)

$$Pr_{acum(i,1)} = Pr_{LM}(S_i/INI) \times Pr_{TRAD}(S_i/W_j)$$



Volver hacia atrás empezando en el mejor nodo final (Pr_{acum} más alta)

Secuencia de Palabras (j)

Grupo de Tecnología del Habla

22 de 63

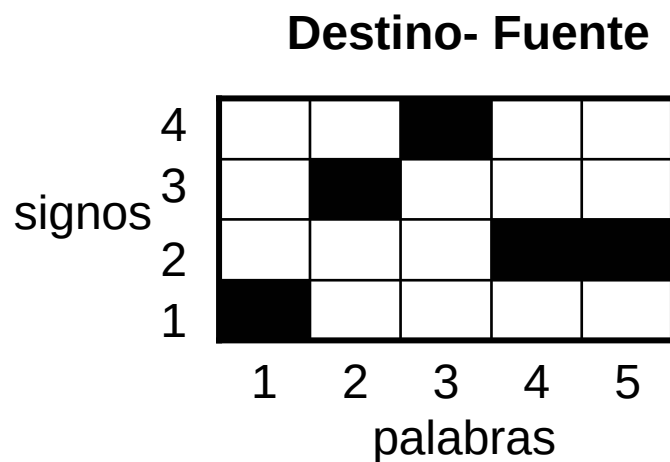


Traducción estadística

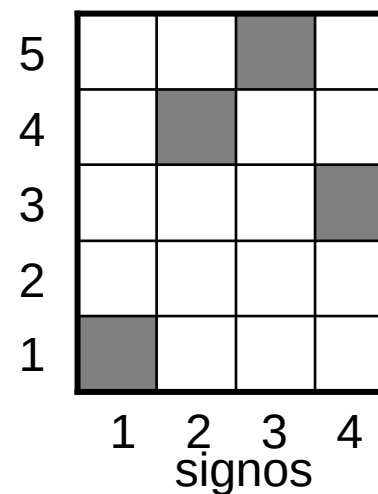
Subfrases

Modelo de Traducción

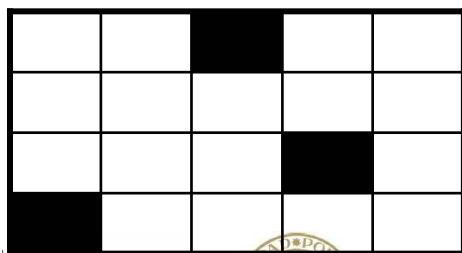
1. Preparación del corpus y alineamiento **bidireccional** con el corpus de entrenamiento



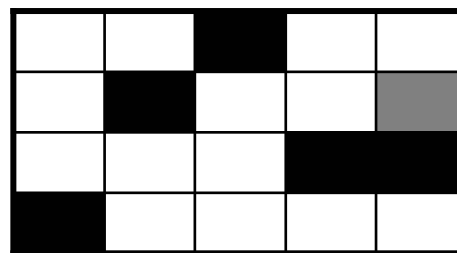
Fuente-Destino



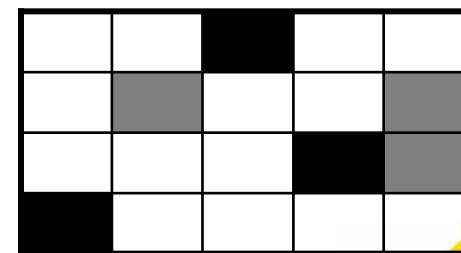
Intersection



Union



grow-diag-final



Traducción estadística

Modelo de Traducción

2. Extracción de subsecuencias de palabras

Subfrases

	ESTOS	PLURAL	PAIS
a			
los			
siguientes			
países			

consistente

	ESTOS	PLURAL	PAIS
a			
los			
siguientes			
países			

inconsistente

	ESTOS	PLURAL	PAIS
a			
los			
siguientes			
países			

inconsistente

3. Cálculo de las probabilidades de traducción de subfrases

algunas comunidades de españa puede ||| PLURAL ALGUNO PLURAL REGIÓN
 ESPAÑA TÚ PODER ||| (0,1) (2,3) () (4) (5,6) ||| (0) (0) (1) (1) (3)
 (4) (4) ||| 1 0.000906219 1 0.0099428 2.718

Traducción estadística

Subfrases

Traducción

Utilizando Moses se parte de las subsecuencias:

P1 P2 | P3 P4 P5
P1 | P2 P3 | P4 P5
P1 | P2 | P3 | P4 P5

Y se generan nodos con su posible traducción a LSE:

P1 P2 || S2

P3 P4 P5 || S6 S20

P2 P3 || S4

P4 P5 || S6 S20

P1 || S2

P3 || S6

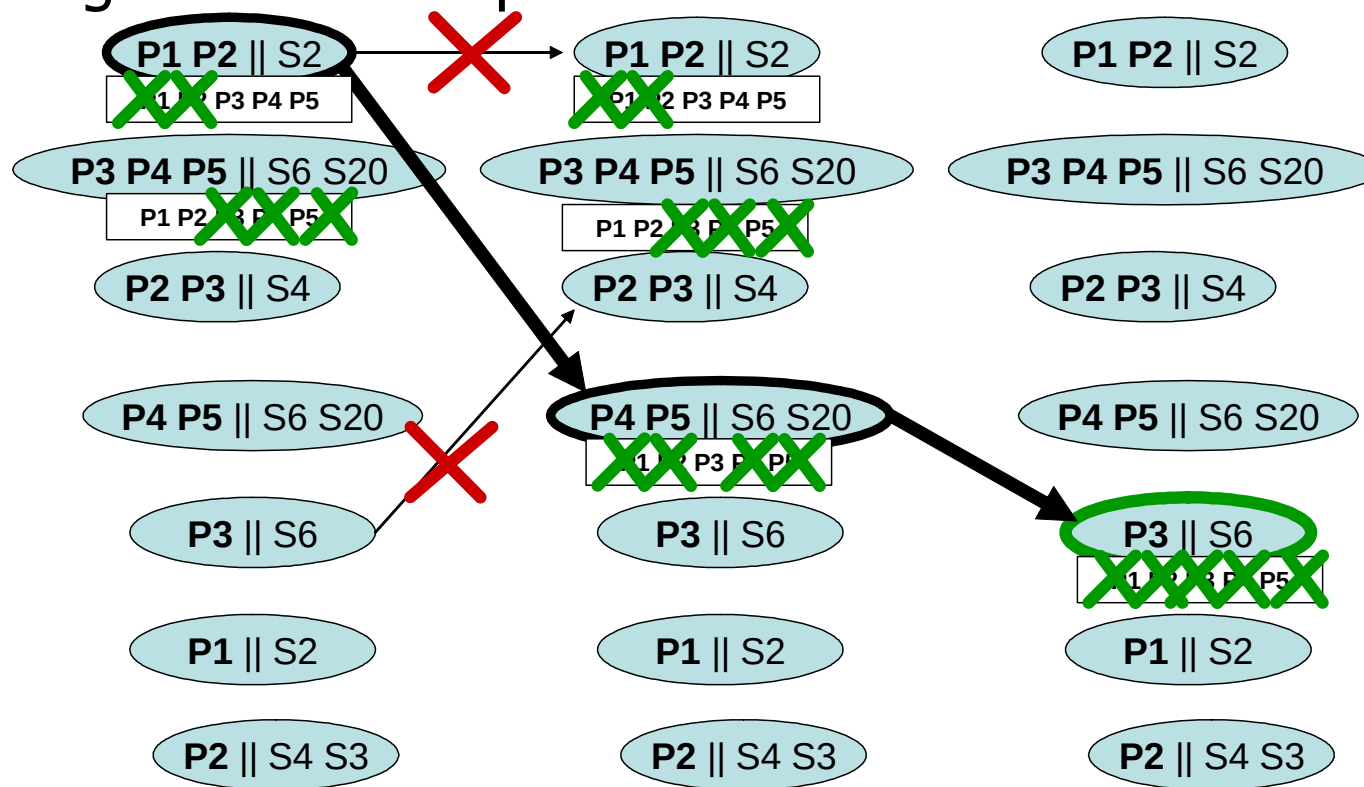
P2 || S4 S3

Traducción estadística

Subfrases

Traducción

Se elige el camino óptimo:



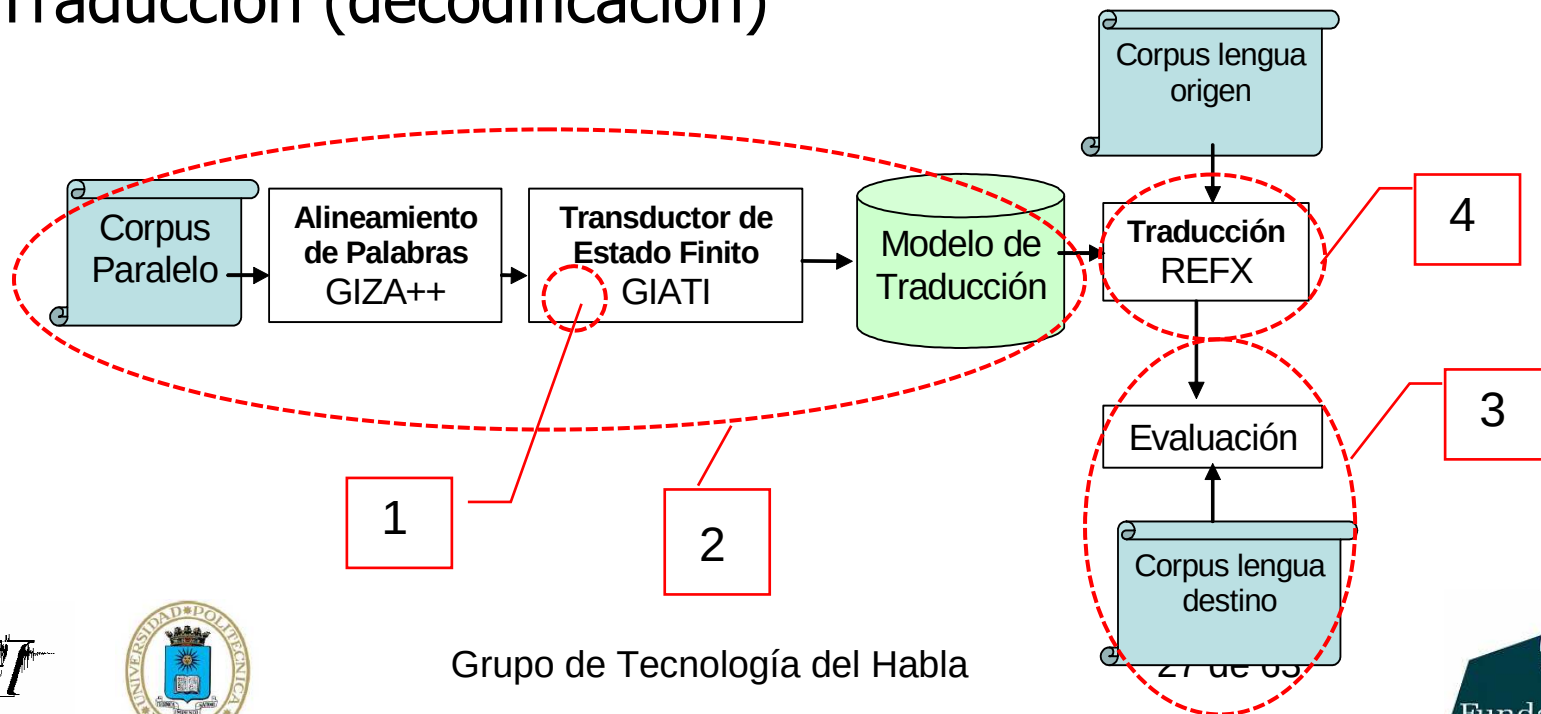
Finalmente se recorre el camino inverso y se obtiene la secuencia deseada de signos resultado de la traducción

Traducción estadística

FST

Visión General de la Arquitectura (Casacuberta & Vidal, 2004)

1. Generación del Modelo de Lenguaje
2. Generación del Modelo de Traducción
3. Fase de Test
4. Traducción (decodificación)



Traducción estadística

FST

Pasos para la traducción mediante FST

1. Alineamiento con GIZA++ a nivel de palabras

```
el denei es obligatorio desde los catorce años # DNI(2) SE-LLAMA(3)
OBLIGATORIO(4) DESDE(5) CATORCE(7) PLURAL(6) AÑO(8) EDAD(8)
el denei es obligatorio # DNI(2) SE-LLAMA(3) OBLIGATORIO(4)
el denei es el documento oficial # DNI(2) SE-LLAMA(3) DOCUMENTO(5)
OFICIAL(6)
el denei es oficial # DNI(2) SE-LLAMA(3) OFICIAL(4)
```

2. Combinación de ambas frases ("extended-words")

```
(el,  $\lambda$ ) (denei, DNI) (es, SE-LLAMA) (obligatorio, OBLIGATORIO) (desde,
DESDE) (los, PLURAL), (catorce, CATORCE) (años, AÑO EDAD)
(el,  $\lambda$ ) (denei, DNI) (es, SE-LLAMA) (obligatorio, OBLIGATORIO)
(el,  $\lambda$ ) (denei, DNI) (es, SE-LLAMA) (el,  $\lambda$ ) (documento, DOCUMENTO)
(oficial, OFICIAL)
(el,  $\lambda$ ) (denei, DNI) (es, SE-LLAMA) (oficial, OFICIAL)
```

Condición 1:

Cada palabra destino se debe unir con la correspondiente palabra origen según el alineamiento **siempre que no se modifique el orden de las palabras en destino**. Si eso ocurre la palabra destino se debe unir a la primera palabra fuente que no viole esta condición.



Traducción estadística

FST

Pasos para la traducción mediante FST

2. Combinación de ambas frases ("extended-words")

(el denei, DNI) (es, SE-LLAMA) (obligatorio, OBLIGATORIO) (desde, DESDE)
(los, PLURAL), (catorce, CATORCE) (años, AÑO EDAD)

(el denei, DNI) (es, SE-LLAMA) (obligatorio, OBLIGATORIO)

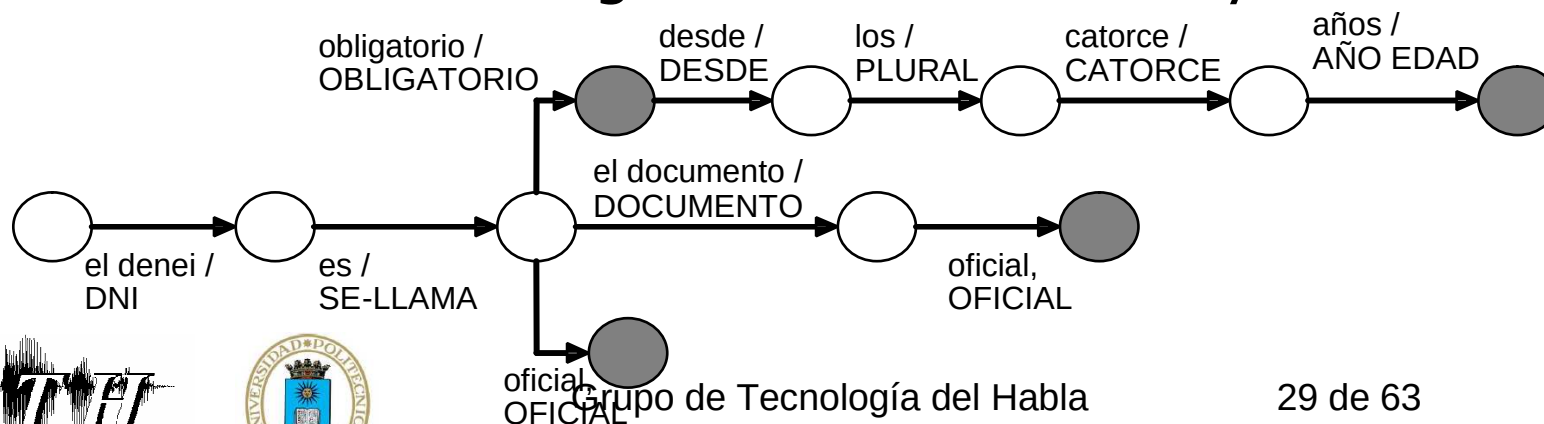
(el denei, DNI) (es, SE-LLAMA) (el documento, DOCUMENTO) (oficial, OFICIAL)

(el denei, DNI) (es, SE-LLAMA) (oficial, OFICIAL)

Condición 2:

Las **palabras fuente aisladas** (que no están alineadas a ninguna palabra destino) serán unidas a la primera palabra extendida que contenga palabras destino asignadas.

3. Inferencia de una gramática estocástica y un FST asociado



Traducción estadística **Medidas de Evaluación**

Para evaluar la calidad de una traducción se compara la salida del sistema automático con una referencia y se calculan unas medidas de evaluación:

✓ **WER:** "Word Error Rate", porción de palabras con error:

$$\text{WER}(\%) = \frac{S + B + I}{N} \cdot 100$$

✓ **BLEU:** precisión de n-gramas entre traducción del sistema y la referencia

✓ **NIST:** similar a BLEU, utiliza la media aritmética y considera la precisión de un solo n-grama

Experimentos de reconocimiento de voz

Tres experimentos diferentes:

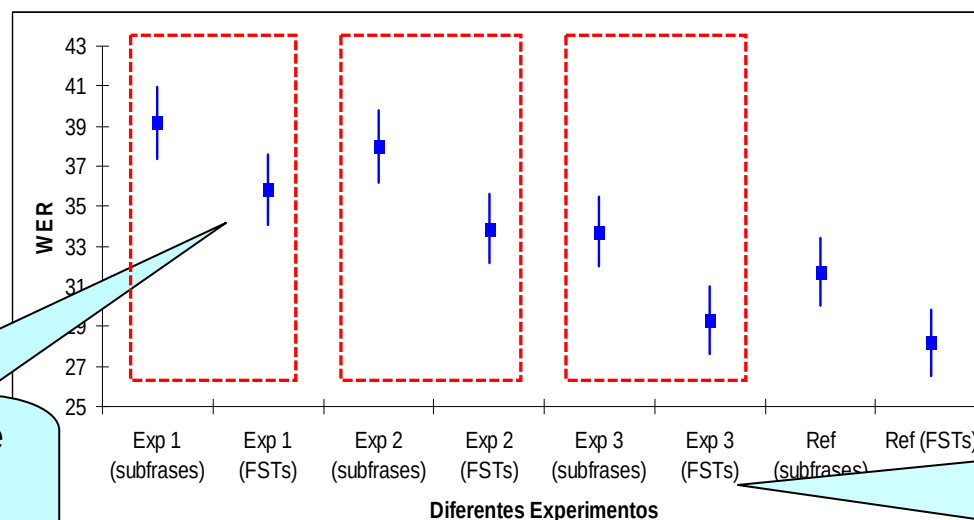
1. Tanto LM como vocabulario generados a partir del conjunto de entrenamiento
2. El LM usa el conjunto de entrenamiento y el vocabulario todas las frases
3. Se utilizan todas las frases tanto para el vocabulario como para el LM

	Word Error Rate	I (%)	B (%)	S (%)
Experimento 1	24,08	2,61	6,71	14,76
Experimento 2	15,84	1,19	5,93	8,72
Experimento 3	4,74	0,86	1,94	1,94

Traducción estadística

Resultados de referencia

Resultados



Al aumentar la tasa de reconocimiento, aumenta la de traducción (menor WER)

Traducción basada en FSTs da menor WER que la basada en subfrases para esta BD

Subfrases

	WER	$\pm \Delta$	BLEU	NIST
Experimento 1	39,17	1,78	0,4853	6,2806
Experimento 2	37,99	1,77	0,4900	6,4006
Experimento 3	33,72	1,73	0,5301	6,7238
Referencia	31,75	1,70	0,5469	6,8652

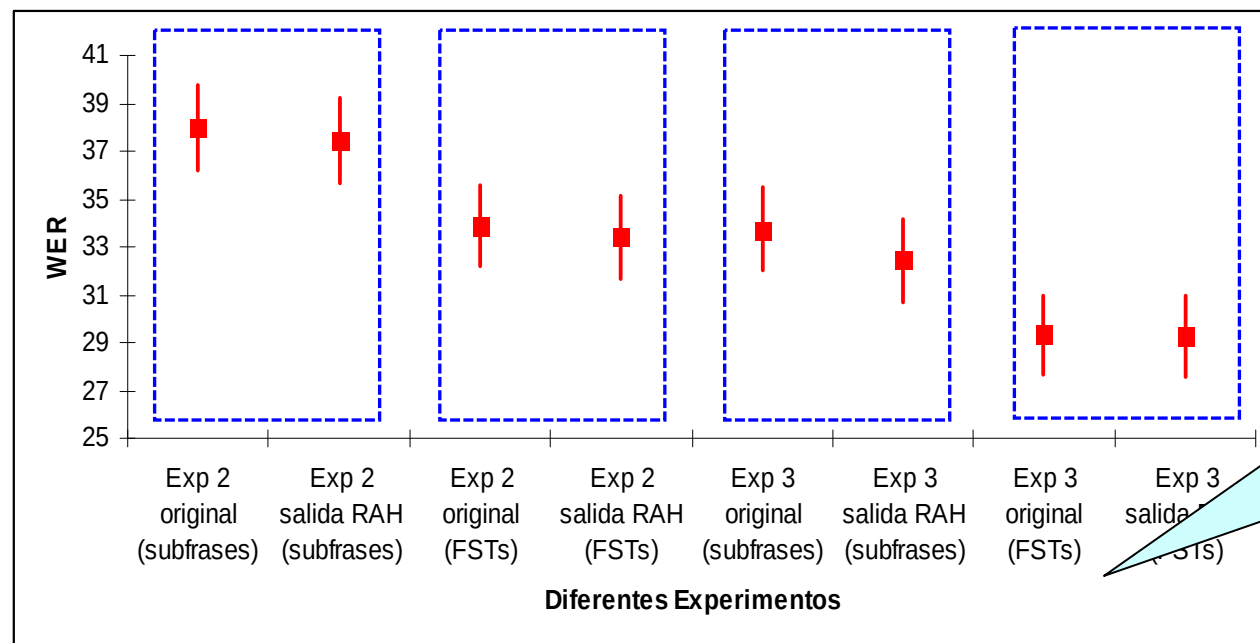
FSTs

	WER	$\pm \Delta$	BLEU	NIST
Experimento 1	35,85	1,75	0,509	6,6470
Experimento 2	33,89	1,73	0,5238	6,8610
Experimento 3	29,32	1,66	0,5804	7,3100
Referencia	28,21	1,64	0,5905	7,3500

Traducción estadística

Resultados

Incorporación de la salida del RAH en entrenamiento del modelo de traducción



Mejoras no significativas al añadir la salida del RAH en el entrenamiento del modelo de traducción

subfrases

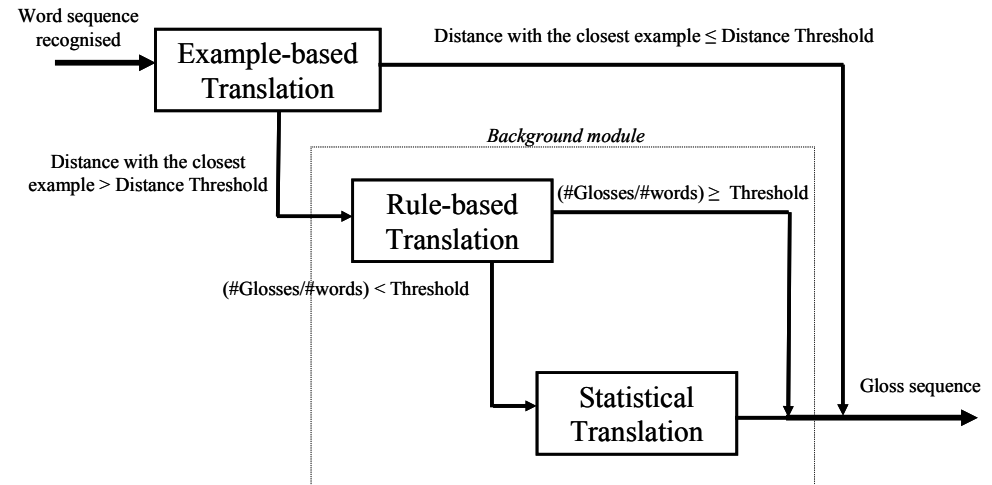
	WER	$\pm\Delta$	BLEU	NIST
Experimento 1	40,04	1,79	0,4775	6,2076
Experimento 2	37,46	1,77	0,4939	6,4738
Experimento 3	32,44	1,71	0,5449	6,8606

FSTs

	WER	$\pm\Delta$	BLEU	NIST
Experimento 1	36,33	1,76	0,5188	6,5273
Experimento 2	33,42	1,72	0,5235	6,8344
Experimento 3	29,27	1,66	0,5698	7,1953

Resultados en pruebas de laboratorio

- 1425 archivos, 3 locutores.
- Las métricas:
 - WA, B, S e I
 - PIWER.
 - BLEU y NIST.



	WA	B	S	I	PIWER
Reconocimiento	92.24	3.06	4.36	0.34	7.67
Traducción	84.89	5.79	7.77	1.55	14.24

N	1	2	3	4	5	6	7	8
NIST	6.5108	8.4427	9.047	9.2837	9.4165	9.4736	9.4886	9.4886
BLEU	0.8679	0.8392	0.8138	0.7919	0.7743	0.7596	0.7445	0.733

Representación de signos (I)

Proyecto eSign - Agente Virtual Guido

- Sistema de notación HamNoSys
- Herramienta eSign Editor
- Lenguaje SIGML
- Ventaja: Flexibilidad



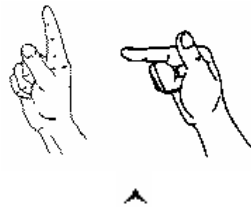
Representación de signos (II)

- **HAMNOSYS** (<http://www.sign-lang.uni-hamburg.de/Projekte/HamNoSys>)

- Descripción de los signos:

- Forma de la mano → 
- Orientación de la mano 

- Dirección de la base del dedo índice



- Orientación de la palma de la mano

Ejemplo

^○: Dirección del eje de la mano hacia arriba, palma de frente.

^0: Dirección del eje de la mano hacia arriba, palma hacia la izquierda.

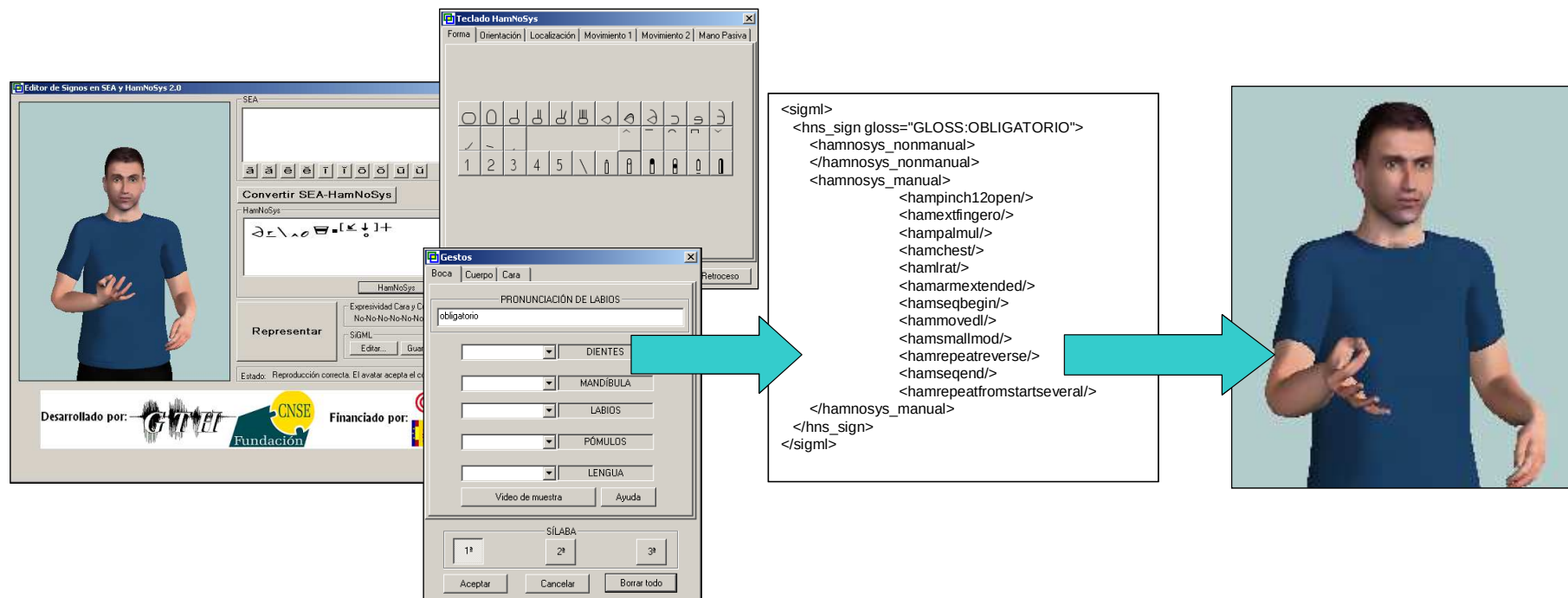
Representación de signos (III)

	A la izquierda	En la parte izquierda	Centro	En la parte derecha	A la derecha
Cabeza	◻○	▪○	○	○▪	○◻
Frente	◻∩	▪∩	∩	∩▪	∩◻
Nariz	◻∩	▪∩	∩	∩▪	∩◻
Boca	◻∞	▪∞	∞	∞▪	∞◻
Barbilla	◻∪	▪∪	∪	∪▪	∪◻
Cuello	◻∩∩	▪∩∩	∩∩	∩∩▪	∩∩◻
Línea de los hombros	◻∩	▪∩	∩	∩▪	∩◻
Línea del pecho	◻∩	▪∩	∩	∩▪	∩◻
Línea del estómago	◻∩	▪∩	∩	∩▪	∩◻

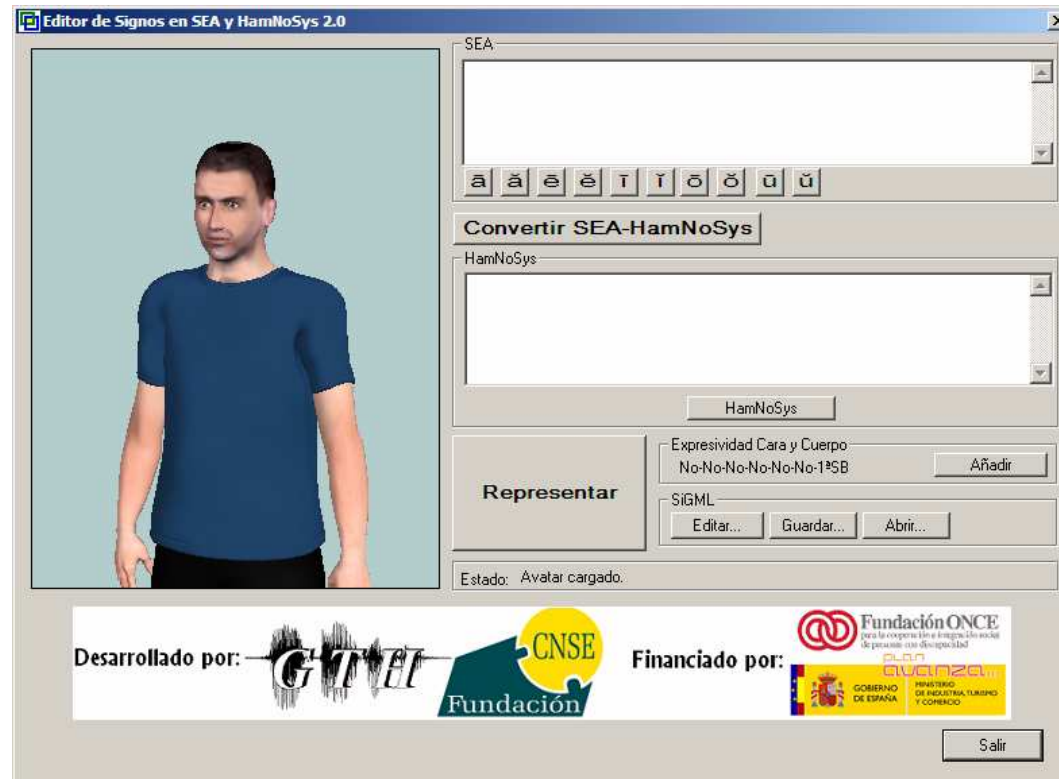
Editor de signos (I)



Editor de signos (II)



Editor de signos (III)



El objetivo de la herramienta es la creación del archivo en código SiGML en el que se especifican las características del signo

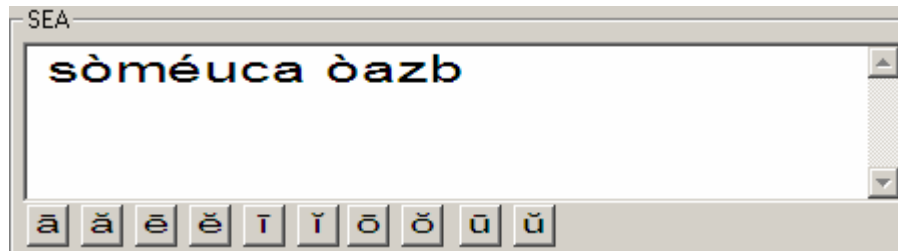
Existen **3 modos de edición**:

- En SEA.
- En HamNoSys.
- Directamente en SiGML.

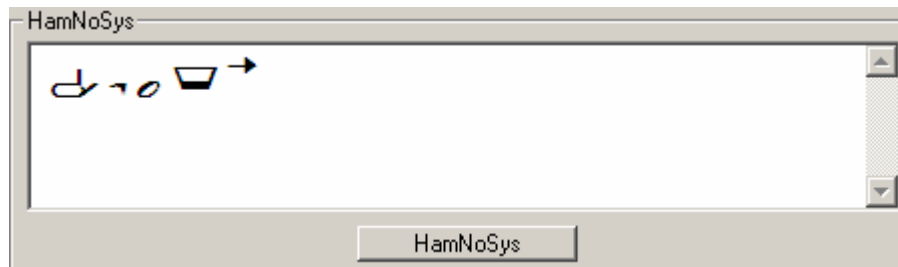
- La conversión entre HamNoSys y SiGML es directa, ya que son equivalentes.
- Para pasar de SEA a HamNoSys tenemos que realizar la conversión, obteniendo un resultado no siempre perfecto.

Editor de signos (IV)

Modo de edición en SEA



Modo de edición en HamNoSys



Representación animada

- Para que el agente animado represente el signo es necesaria la especificación en **HamNoSys** o **SiGML**.

- La especificación de SEA utiliza caracteres corrientes.
- En ocasiones utiliza algún carácter especial, que se insertan mediante botones.



- Para insertar los caracteres de HamNoSys utilizamos el **teclado de HamNoSys**, ya utilizado en el sistema de traducción.

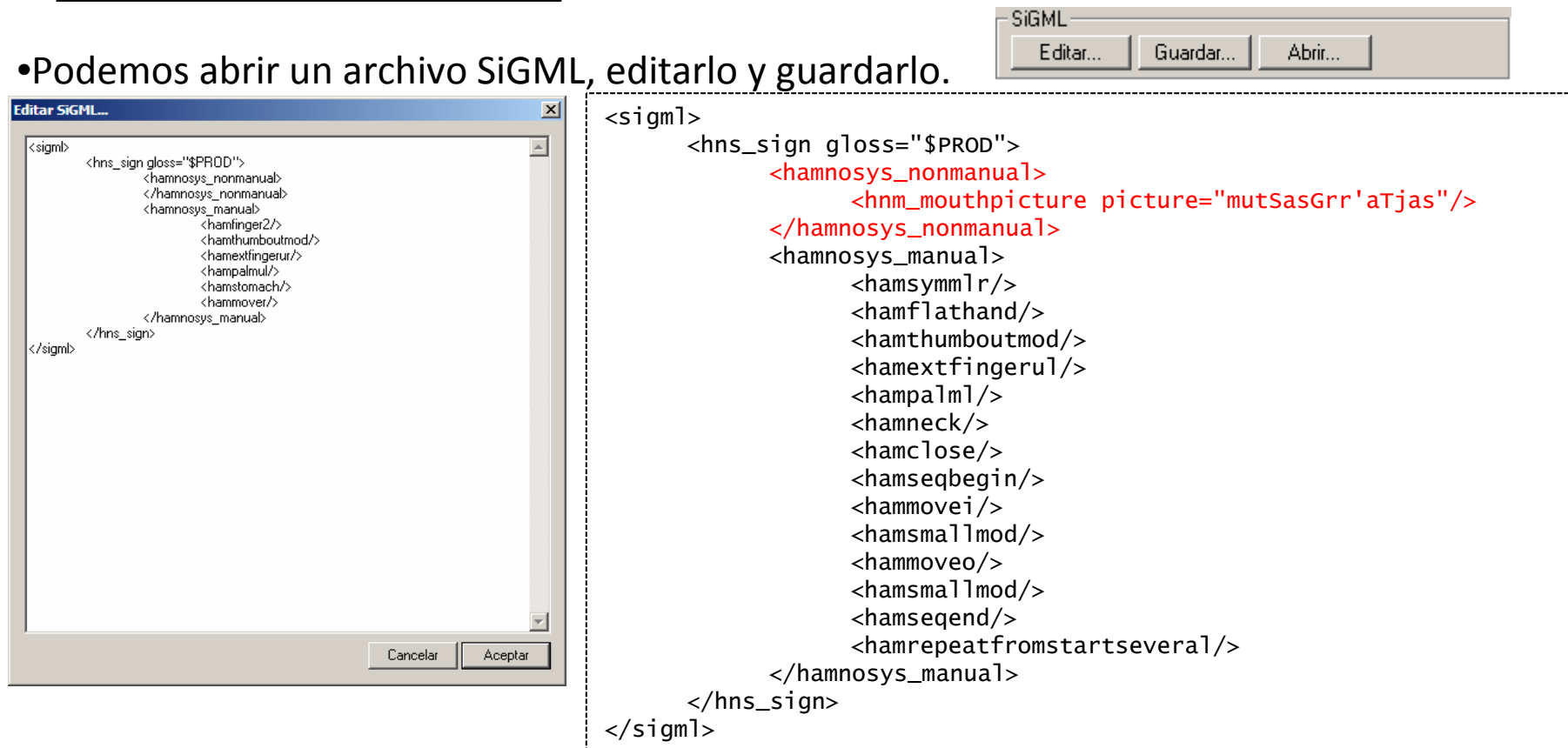


Representar

Editor de signos (V)

Modo de edición en SiGML

- Podemos abrir un archivo SiGML, editarlo y guardarlo.



The image shows the SiGML editor interface. On the left is a window titled 'Editar SiGML...' containing a text area with XML code. On the right is a control panel titled 'SiGML' with three buttons: 'Editar...', 'Guardar...', and 'Abrir...'. Below the control panel is a larger window showing the same XML code, with some elements highlighted in red.

```
<sigml>
  <hns_sign gloss="$PROD">
    <hamnosys_nonmanual>
    </hamnosys_nonmanual>
    <hamnosys_manual>
      <hamfinger2/>
      <hamthumboutmod/>
      <hamextfingerur/>
      <hampalmul/>
      <hamstomach/>
      <hammovei/>
    </hamnosys_manual>
  </hns_sign>
</sigml>
```

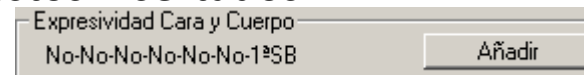
```
<sigml>
  <hns_sign gloss="$PROD">
    <hamnosys_nonmanual>
      <hnm_mouthpicture picture="mutSasGrr'aTjas"/>
    </hamnosys_nonmanual>
    <hamnosys_manual>
      <hamsymmlr/>
      <hamflathand/>
      <hamthumboutmod/>
      <hamextfingerul/>
      <hampalm1/>
      <hamneck/>
      <hamclose/>
      <hamseqbegin/>
      <hammovei/>
      <hamsmallmod/>
      <hammoveo/>
      <hamsmallmod/>
      <hamseqend/>
      <hamrepeatfromstartseveral/>
    </hamnosys_manual>
  </hns_sign>
</sigml>
```

- La parte en rojo del código SiGML corresponde a los gestos del cuerpo y de la cara

Editor de signos (VI)

Inserción de gestos de la cara y del cuerpo

- En el cuadro de estado aparecen los últimos gestos insertados.
- Podemos insertar en **3 sílabas** diferentes.



- Contiene 3 pestañas:
 - Gestos producidos por la boca, incluyendo pronunciación de sonidos mediante los labios.
 - Movimiento de hombros.
 - Otros gestos de la cara: mirada, cejas, párpados y nariz.

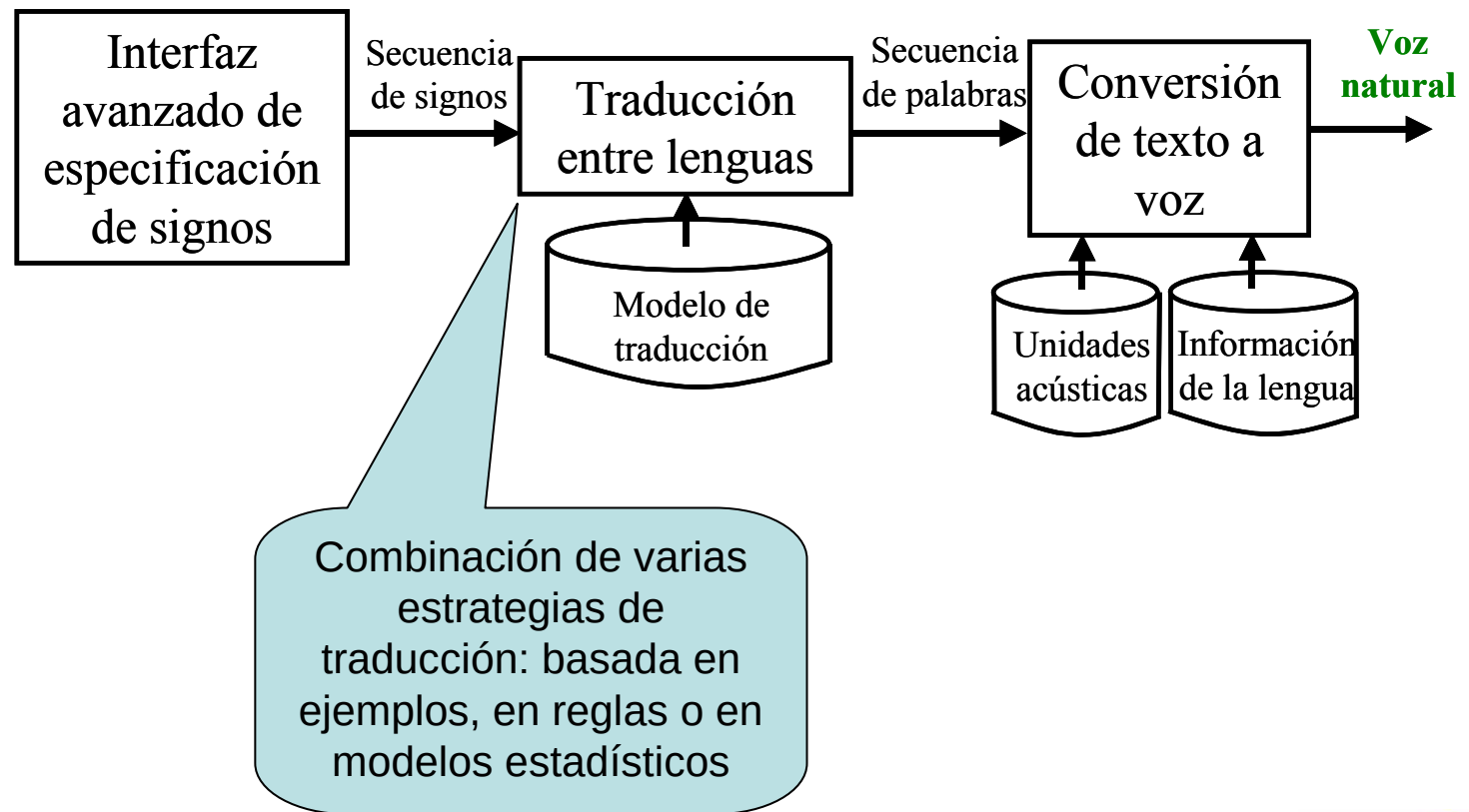
Editor de signos (VII)

[illegible]

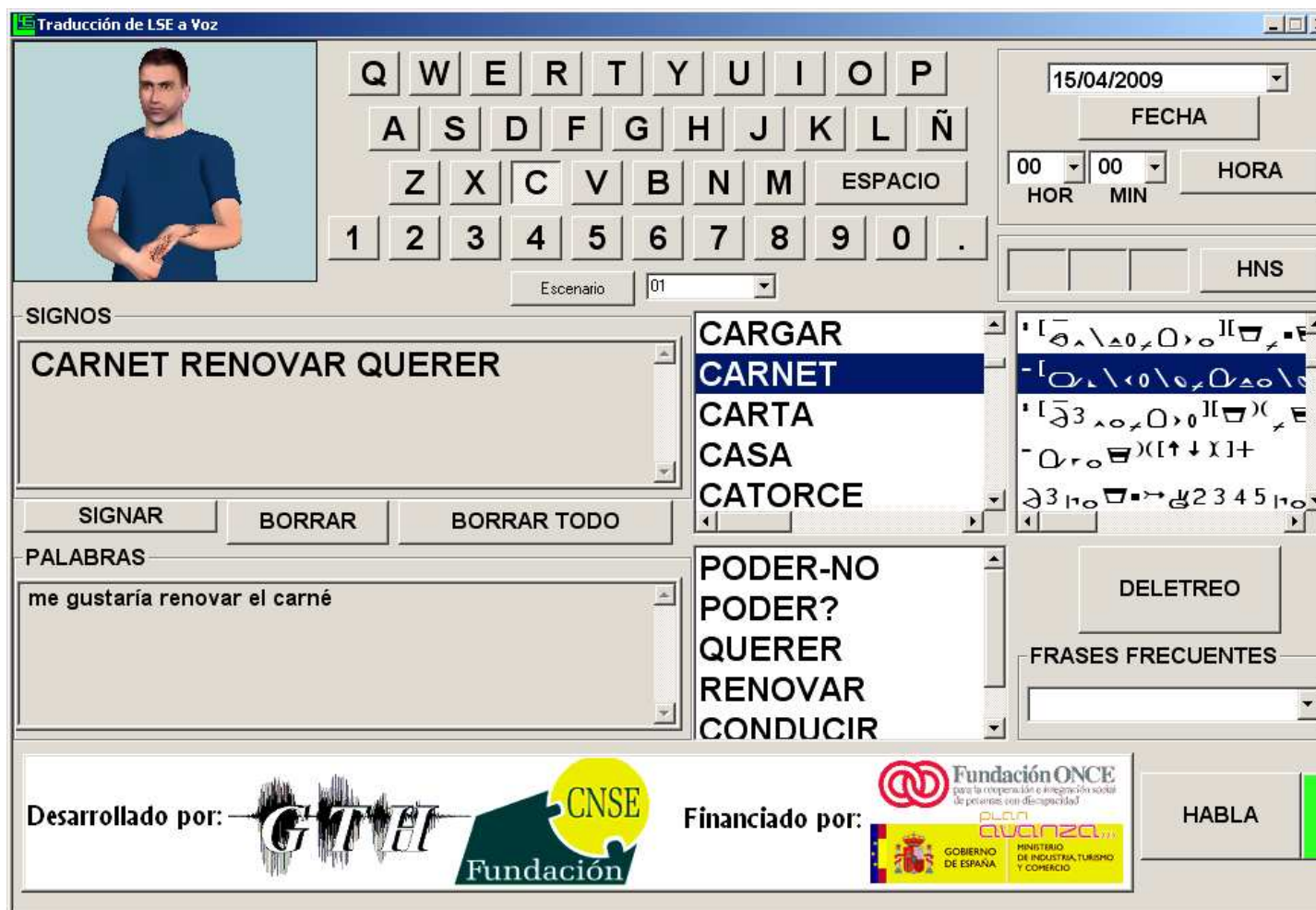
Alrededor de 800 signos descritos en
glosas, SEA, HamNoSys y SIGML



Generación de voz a partir de LSE (I)



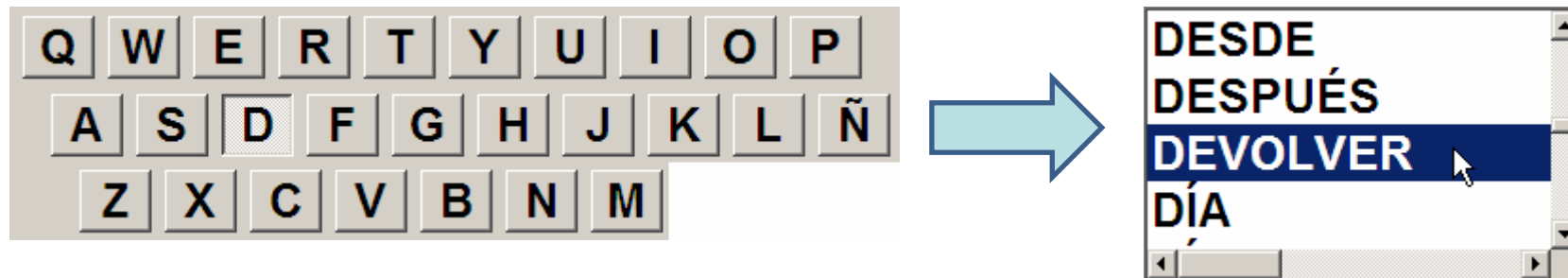
Generación de voz a partir de LSE (II)



Interfaz visual (I)

Búsqueda y selección de signos mediante glosas

- Mediante el teclado alfabético realizamos la búsqueda.
- Seleccionamos el signo deseado haciendo doble click.



Posibilidad de deletrear signos, letra a letra, para insertar nombres propios:

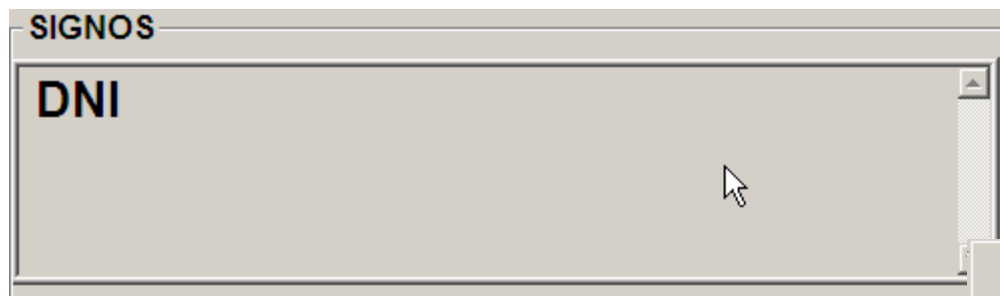
- El modo de deletreo se activa mediante el botón *DELETREO*.
- Cada letra se introduce mediante el teclado alfabético.



Interfaz visual (II)

Sugerencias de signos mediante glosas

- Se buscan los signos más probables que siguen al último insertado:
 - Las probabilidades se calculan sobre la **base de datos de ejemplos**.
- Al posar el puntero del ratón sobre el cuadro de texto de inserción de signos:
 - Aparece un menú flotante con las sugerencias más probables.
 - Pueden aparecer hasta **tres signos** como sugerencia.

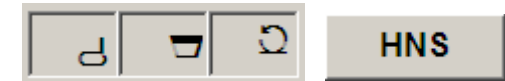


Para seleccionar un signo
debemos realizar un
único click

Interfaz visual (III)

Búsqueda y selección de signos mediante HamNoSys

- Como parámetros de búsqueda seleccionamos caracteres de HamNoSys.
- Se buscan aquellos signos que contengan todos los parámetros escogidos.
- Podemos seleccionar hasta **3 caracteres** de HamNoSys

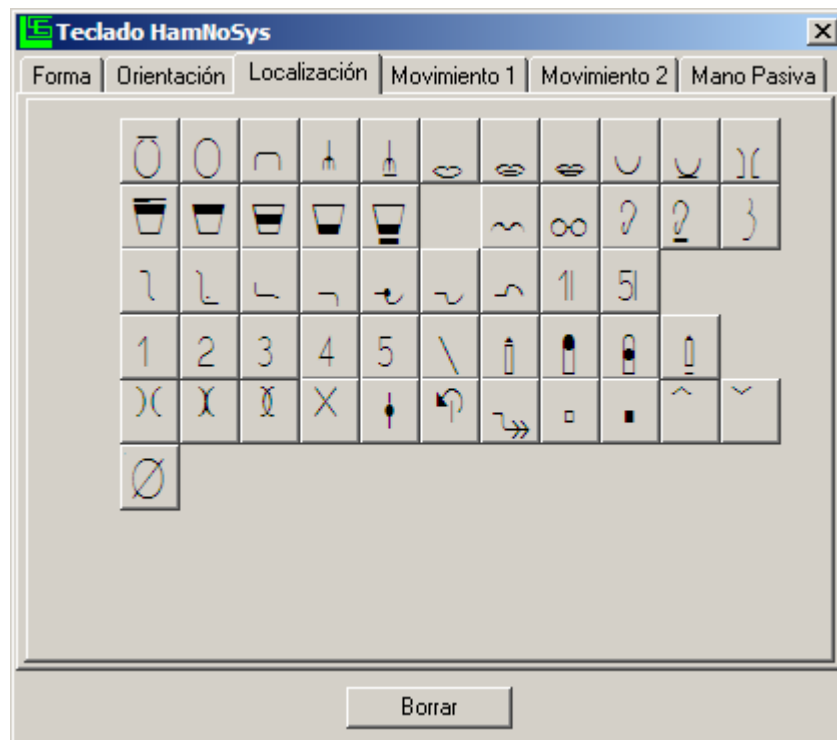


- El resultado de la búsqueda aparece en las listas de glosas y HamNoSys:



Interfaz visual (IV)

Teclado de HamNoSys

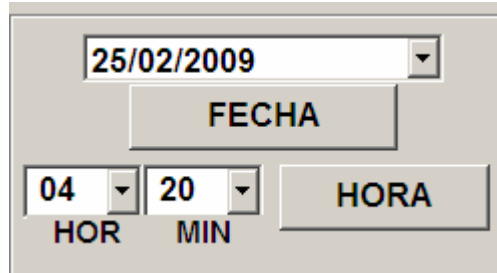


- Corresponde a un nuevo diálogo.
- El elemento contiene seis pestañas:
 - Forma.
 - Orientación.
 - Localización.
 - Movimiento 1.
 - Movimiento 2.
 - Mano Pasiva.
- Cada vez que pulsamos un carácter de HamNoSys, éste aparece como parámetro de búsqueda en el interfaz principal.

Interfaz visual (V)

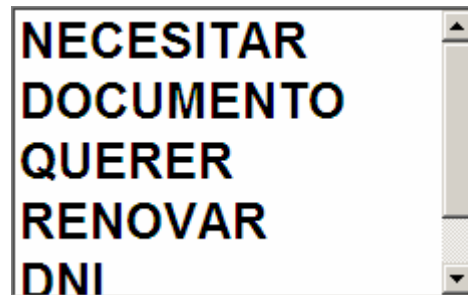
Ayudas generales

Podemos insertar fechas y horas



A screenshot of a date and time selection interface. At the top, there is a text box containing '25/02/2009' with a small downward arrow to its right. Below this is a button labeled 'FECHA'. Underneath the button are two more text boxes: the first contains '04' with a downward arrow, and the second contains '20' with a downward arrow. Below these two boxes are the labels 'HOR' and 'MIN' respectively. To the right of these is a button labeled 'HORA'.

Los últimos 5 signos insertados se almacenan



A screenshot of a list of inserted signs. The list is contained within a rectangular box with a vertical scrollbar on the right. The text in the list, from top to bottom, is: 'NECESITAR', 'DOCUMENTO', 'QUERER', 'RENOVAR', and 'DNI'.

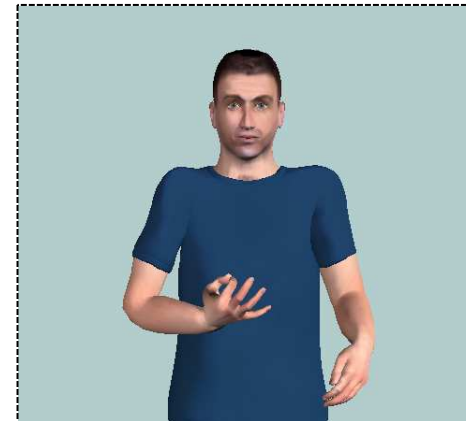
Podemos insertar fórmulas básicas

FRASES FRECUENTES

MUCHAS-GRACIAS

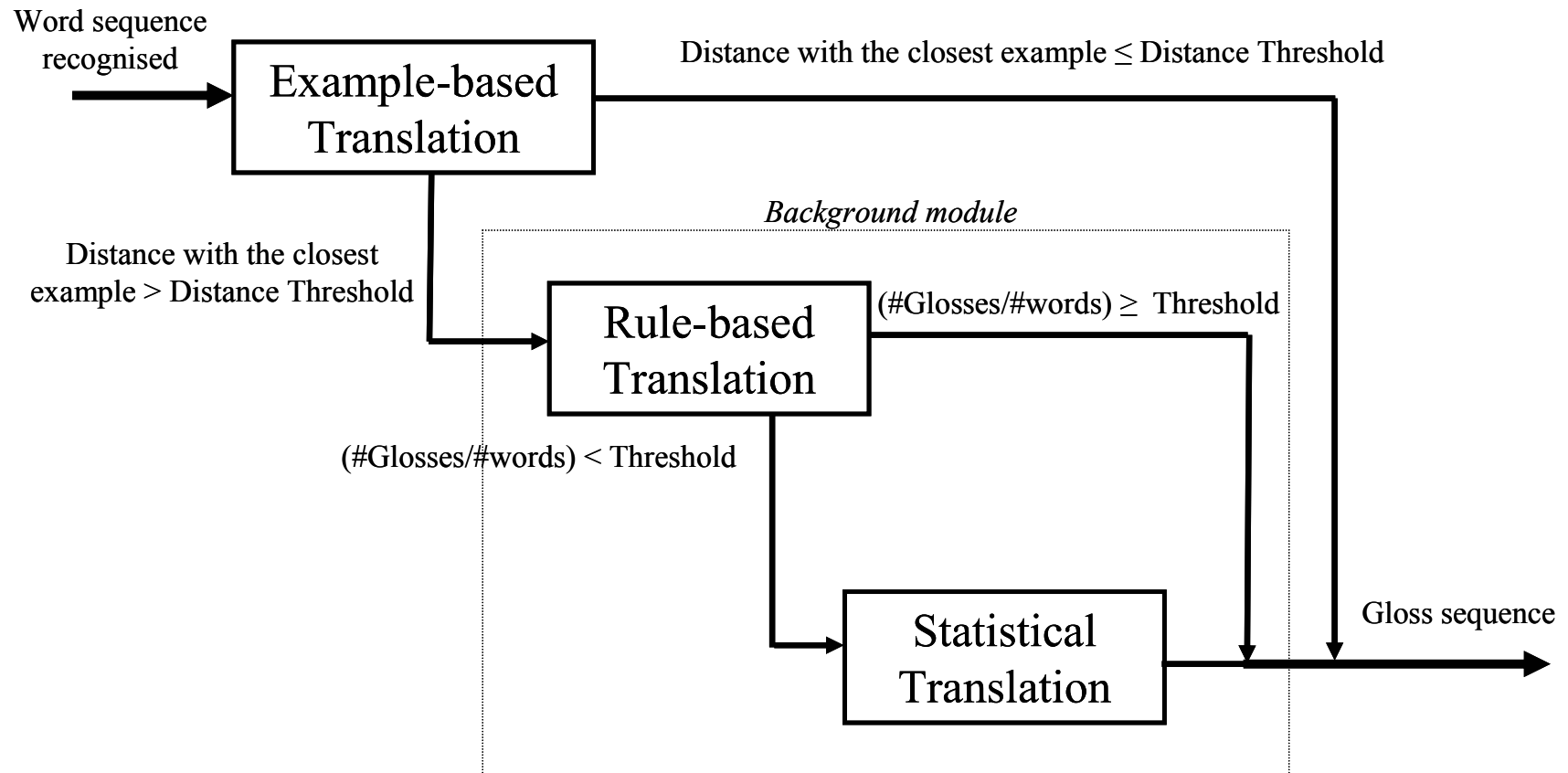
SIGNAR

- En vez de seleccionar un signo en las listas desplegables, si realizamos un único click el agente animado representa el signo.
- De esta manera podemos comprobar cualquier inserción.
- Es una gran ayuda para la elección del signo.



- Para signar toda la cadena insertada:

Traducción de LSE a texto



Conversión texto a voz



Voz Jorge



Grupo de Tecnología del Habla

53 de 63



Evaluación



- Jefatura Provincial de Tráfico de Toledo: renovación del permiso de conducir
 - 2 funcionarios
 - 10 personas sordas (5 de Madrid y 5 de Toledo)
 - 6 escenarios
 - 2 días de evaluación

Evaluación: resultados (I)

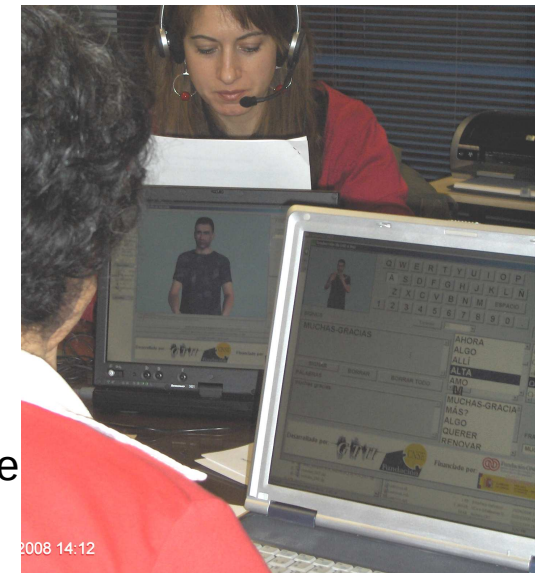
- Sistema de traducción de voz a LSE
 - Muy buena tasa de reconocimiento y de traducción



AGENTE	MEDIDA	VALOR
Sistema	Tasa de reconocimiento	95,2%
	Tasa de traducción	91,1%
	Tiempo de reconocimiento	3,2 seg.
	Tiempo de traducción	0,0013 seg.
	Tiempo de signado	4,7 seg.
	Traducción basada en ejemplos	94,9%
	Traducción basada en reglas	4,2%
	Traducción estadística	0,8%
	Número de turnos del funcionario por diálogo	8,4
	Número de diálogos	48

Evaluación: resultados (II)

- Sistema de traducción de voz a LSE
 - Muy buena tasa de reconocimiento y de traducción
 - Principales problemas
 - Agente animado poco natural
 - Bajo nivel de estandarización de la LSE: tanto en las estructuras de las frases como en la representación de los signos

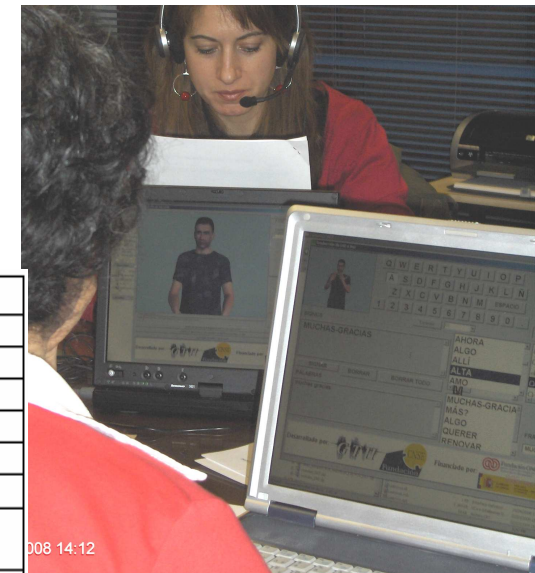


AGENTE	MEDIDA	VALOR (0-5)
Funcionario	Rapidez del sistema	4
	Tasa de reconocimiento	3,5
	Facilidad de uso	3,5
	Rapidez de aprendizaje para el manejo del sistema	3,5
	Si no hubiera un intérprete de LSE, ¿usaría este sistema?	3,5
	Valoración GLOBAL	3,5
Usuario	¿Los signos son correctos?	2,1
	Comprendo los signos	2,2
	El signado es natural	0,8
	Si no hubiera intérprete, ¿usaría el sistema?	2,0
	Valoración GLOBAL	2,2

Evaluación: resultados (III)

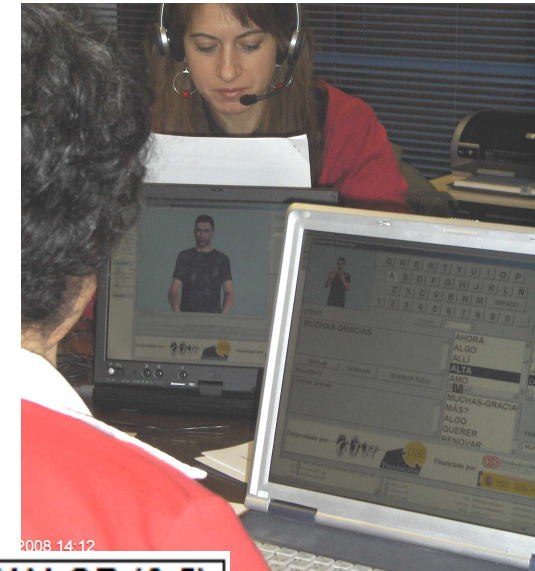
- Generación de voz
 - Muy buena tasa de traducción

AGENTE	MEDIDA	VALOR
Sistema	Tasa de traducción	98%
	Tiempo de traducción	0,001 seg.
	Tiempo de conversión de texto a voz	1,7 seg.
	Traducción basada en ejemplos	91,9%
	Traducción basada en reglas	8,1%
	Traducción estadística	0,0%
	Tiempo en segundos que se tarda en especificar una secuencia de glosas.	18,5
	Número de clic de ratón que se tarda en especificar una secuencia de glosas.	8,6 clics
	Tiempo en segundos que se tarda en especificar una secuencia de glosas normalizado al número de glosas.	7,4 seg.
	Número de clic de ratón que se tarda en especificar una secuencia de glosas normalizado al número de glosas.	3,3 clics.
	Número de glosas por turno	2,6
	Número de veces que se selecciona un signo según:	
	- Lista de glosas	50,4%
	- Lista de Hamnosys	0,0%
	- Lista de siguientes	48,1%
	- Lista de últimos	1,5%
	- Fecha	0,0%
	- Hora	0,0%
	- Deletreo	0,0%
	Número de turnos con frases frecuentes por diálogo	0,4
	Número de turnos del usuario (botón HABLAR) por diálogo	4,0
	Número de diálogos	48



Evaluación: resultados (IV)

- Generación de voz
 - Muy buena tasa de traducción
 - Principales problemas
 - No suficiente tiempo para entrenar con el sistema
 - Bajo nivel de estandarización para signo escritura: glosas, SEA o HamNoSys



AGENTE	MEDIDA	VALOR (0-5)
Funcionario	Inteligibilidad de las frases habladas	4
	Naturalidad de la voz generada	3
	Si no hubiera un intérprete de LSE, ¿usaría este sistema?	4
	Valoración GLOBAL	4
Usuario	Rapidez del sistema	3,2
	Facilidad de aprendizaje del programa	2,4
	Necesita más signos para expresarse	2,6
	Si no hubiera intérprete, ¿usaría el sistema?	2,8
	Valoración GLOBAL	3,00

Resumen de resultados

- Primer corpus paralelo Español y LSE para un dominio de aplicación concreto.
- 2 sistemas de traducción de voz a LSE para la renovación del DNI y del permiso de conducir.
- 2 sistemas de generación de voz para la renovación del DNI y del permiso de conducir.
- Base de datos con la descripción en glosas, SEA y HamNoSys de más de 800 signos (de estos dominios de aplicación)
- Evaluación en situaciones reales



Líneas futuras

- Mejora de los módulos de traducción
- Mejorar la flexibilidad de las herramientas para adaptarlas fácilmente a otros dominios
- Ampliación de la base de datos
- Nuevos avatares más flexibles y naturales



Difusión del proyecto (I)



Para la supresión de las barreras de comunicación

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



Grupo de Tecnología del Habla de la Universidad Politécnica de Madrid

Contenidos

Introducción
Objetivos
[Documentos Generados](#)
[Contenidos Digitales](#)
Publicaciones
Participantes
[Demostraciones](#)

07 / 01 / 2009

Novedades

Organización del primer curso de HamNoSys para los días 14 y 15 de julio

Reunión de seguimiento

Hemos estado en la JPT de Toledo para comenzar el estudio lingüístico en este entorno

Comenzamos el estudio lingüístico de las frases para la renovación del DNI

Primera Reunión de Coordinación

PROYECTO CONSEGUIDO!!!!

Redacción de la memoria final

Los principales documentos generados son:

- Memoria de técnica del proyecto
- Transparencias de la primera reunión de coordinación
- Documento de reglas de generación de glosas V1.0
- Documentos para la generación de signos en HamNoSys
 - Resumen en castellano
 - Ayuda en inglés del Editor eSign
 - Manual de consulta en inglés del Editor
 - HamNoSys v3.0
 - HamNoSys v4.0
- Fotos y Vídeos de la reunión en la Jefatura Provincial de Tráfico de Toledo
 - Vídeos de la JPT
 - Fotos de la JPT
- Transparencias de la reunión de seguimiento
- Documentación del Curso de Introducción a HamNoSys 4.0
 - Transparencias del curso
 - [Introducción a HamNoSys](#)

Proyecto financiado por:



GOBIERNO DE ESPAÑA
MINISTERIO DE INDUSTRIA, TURISMO Y COMERCIO

Plan Avanza del MITYC



Fundación ONCE
para la cooperación e integración social de personas con discapacidad

Fundación ONCE

Para contactar con el webmaster. Fecha de actualización: 21/11/2008

- Página web: www.traduccionvozlse.es
- 1 noticia en la revista de la UPM: El Mundo, ABC,...
- 2 artículos en revistas de investigación (en curso)
- 3 artículos en congresos nacionales (publicados)



Difusión del proyecto (II)



- Noticias,...



¿Preguntas?

