Procesamiento del Lenguaje Natural, Revista nº 45, septiembre 2010, pp 207-214 recibido 27-04-10 revisado 17-05-10 aceptado 20-05-10

Estudio del tipo de alineamiento en un sistema de traducción estadística de castellano a Lengua de Signos Española (LSE)

Analysis of the alignment configuration in a statistical translation system of Spanish into Spanish Sign Language (LSE)

V. López, R. San-Segundo, R. Martín, J.M. Lucas, R. Barra-Chicote
Grupo de Tecnología del Habla
Universidad Politécnica de Madrid
veronicalopez@die.upm.es

Resumen: La principal aportación de este artículo es el estudio del efecto que tiene el tipo de alineamiento en un sistema de traducción estadística de castellano a Lengua de Signos Española (LSE). El sistema de traducción utiliza un modelo de traducción basado en subfrases o secuencias de palabras. El artículo describe el ajuste de los parámetros de configuración de este sistema para el problema de traducción concreto (castellano-LSE), siendo la selección del tipo de alineamiento un aspecto crítico en los resultados de traducción obtenidos. La selección del tipo de alineamiento se define en el proceso de generación del modelo de traducción basado en palabras como paso previo a la generación del modelo de secuencias de palabras. La evaluación de la arquitectura se realiza con varias métricas: WER (tasa de error de palabras), BLEU ("BiLingual Evaluation Understudy") y NIST. Finalmente, los resultados que se obtienen dan una tasa de error de 28,29%, consiguiendo una reducción relativa de más de un 35% en dicha tasa de error.
Palabras clave: Traducción Automática Estadística, Lengua de Signos Española, Traductor basado en Subfrases o Secuencias de Palabras, Tipo de Alineamiento de Palabras.

Abstract: The main aspect of this paper is the effect analysis of the alignment configuration in a statistical Spanish into Spanish Sign Language translation system. The translation system uses a phrase-based translation model. This paper describes the system configuration adapted for the specific translation problem (Spanish-LSE). In this configuration, the type of alignment is a critical aspect for the system performance. This alignment is used during the process of word-based translation model generation, preliminary step for generating the phrase-based translation model (finally used in translation). The translation system has been evaluated with several metrics: WER (Word Error Rate), BLEU ("BiLingual Evaluation Understudy") and NIST. Finally, the results show a word error rate relative reduction of more than 35% obtaining a final 28.29% WER.
Keywords: Automatic Statistic Translation, Sign Language Translation, Phrase-based Translator, Type of Word Alignment.

1 Introducción

Los sistemas de traducción estadística están adquiriendo cada vez mayor importancia debido a su buen uso de los recursos y a la calidad de las traducciones resultantes.

Este artículo describe los ajustes llevados a cabo en el desarrollo de un sistema de traducción automática estadística que ayude a las personas sordas a poder realizar tareas importantes de la vida cotidiana, como ir a renovar el DNI o el permiso de conducir, sin necesidad de intérpretes signantes que suponen un coste elevado.

La capacidad de una persona sorda (especialmente las personas sordas prelocutivas) de utilizar una lengua escrita es muy inferior a la de un oyente, presentando una capacidad lectora y de escritura muy inferiores en castellano que en Lengua de Signos Española (LSE), pues no pueden extraer información semántica de las palabras ni formar una imagen mental de lo que se les está comunicando: tienen muchos problemas con las conjugaciones de los verbos, concordancias de

ISSN 1135-5948

género y número, y la interpretación de conceptos abstractos. Por ello, actualmente las personas sordas se consideran un colectivo cultural y social diferenciado con cierto aislamiento con respecto a los oyentes.

Es ahí donde radica la importancia del sistema que se describe en este artículo, que trata de eliminar las barreras de comunicación existentes entre personas sordas y oyentes, mejorando un sistema de traducción de castellano a LSE, y que ha sido desarrollado por el Grupo de Tecnología del Habla en colaboración con la Fundación de la Confederación Estatal de Personas Sordas (Fundación CNSE).

2 Estado de la cuestión

En los últimos años, ha habido numerosos proyectos de investigación sobre traducción de habla natural. En Europa: C-Star, ATR, Vermobil, Eutrans, LC-Star, PF-Star y, el más ambicioso, TC_STAR. En Estados Unidos está el programa GALE, cuyo objetivo es desarrollar y aplicar tecnologías de procesado de habla y lenguaje natural para analizar e interpretar enormes volúmenes de voz y texto en varios idiomas.

También, en cuanto a sistemas de traducción a lengua de signos se refiere, varios grupos de investigación han mostrado su interés desarrollando varios prototipos: basados en ejemplos (Morrissey y Way, 2005), en reglas escritas por un experto (San-Segundo, 2008), frases completas (Cox et al, 2002) o métodos estadísticos (Bungeroth y Ney, 2004; Morrissey et al, 2007; sistema SiSi de IBM).

La investigación en lengua de signos ha sido posible gracias a los corpus generados por varios grupos. Algunos ejemplos son los siguientes. En primer lugar destacar el corpus compuesto por más de 300 horas de grabaciones obtenidas a partir de 100 signantes en Lengua de Signos Australiana (T. Johnston, 2008). La base de datos RWTH-BOSTON-400, que es una base de datos que contiene 843 frases, con alrededor de 400 signos diferentes de 5 signantes en Lengua de Signos Americana, con anotaciones en inglés (Dreuw et al., 2008). Otro ejemplo es el proyecto de generación del British Sign Language Corpus que trata de crear un corpus digital de lectura mecánica y espontánea en Lengua de Signos Británica (BSL), proveniente de personas sordas signantes nativas y alumnos, en todo el Reino Unido (Schembri, 2008). Finalmente, comentar el corpus desarrollado en el Institute for Language and Speech Processing (ILSP) y que contiene partes signadas de narración, así como una considerable cantidad de frases y expresiones agrupadas signadas (Efthimiou E., y Fotinea, E., 2008).

Este artículo describe los ajustes de los parámetros de configuración de una arquitectura de traducción estadística basada en subfrases o secuencias de palabras para la traducción de castellano a Lengua de Signos Española (LSE) en dos dominios de aplicación: la solicitud y renovación del DNI y la renovación del permiso de conducir. En el dominio de aplicación mencionado, el sistema de traducción castellano-LSE descrito se complementa con otro sistema de traducción LSE-castellano, de manera que se facilite el diálogo entre una persona sorda y un funcionario. La aportación más importante es el estudio realizado sobre el tipo de alineamiento considerado para la generación de los pares de frases alineadas, paso anterior a la generación del modelo de subfrases.

3 Base de Datos

Para el desarrollo de este sistema de traducción se cuenta con una base de datos formada por frases generadas en dos dominios de aplicación: la solicitud o renovación del DNI, y la renovación del permiso de conducir.

Las frases en castellano fueron obtenidas mediante entrevistas con funcionarios en su atención personal, añadiendo así frases típicas del día a día en este tipo de procesos burocráticos; y frases correspondientes a usuarios que son, por lo general, interrogativas, solicitando una cierta información. Posteriormente, miembros del Grupo de Tecnología del Habla de la Universidad Politécnica de Madrid ampliaron el número de frases añadiendo variantes en castellano diferentes a las de la base de datos inicial, pero con el mismo significado y traducción a LSE.

La traducción de cada una de estas frases a LSE fue realizada por personas sordas conocedoras de la lengua castellana y, posteriormente, expertos en LSE representaron y grabaron estas frases en videos.

De esta manera, la base de datos cuenta con 4080 frases en castellano con sus respectivas traducciones a LSE. Las frases en LSE son secuencias de glosas (palabras en mayúsculas)

que representan los signos. Las características principales de la base de datos pueden verse en la Tabla 1.

	DNI		Carné de Conducir	
Funcionario	Castellano	LSE	Castellano	LSE
Nº de frases	1425		1641	
Frases diferentes	1236	389	1413	199
Nº palabras	8490	6282	17113	12741
Vocabulario	652	364	527	237
Usuario	Castellano	LSE	Castellano	LSE
Nº de frases	531		483	
Frases diferentes	458	139	389	93
Nº palabras	2768	1950	3130	2283
Vocabulario	422	165	294	133

Tabla 1: Principales características del corpus paralelo con frases de diálogos para renovar el DNI y el carné de conducir.

La LSE, como otras lenguas signadas, comparte el canal gestual-visual, pero también tiene características gramaticales de las lenguas orales que conviven con características exclusivas de las lenguas de signos.

En castellano, el orden de los argumentos dentro de la predicación es sujeto-verbo-objeto (SVO) mientras que en LSE es sujeto-objeto-verbo (SOV). Además, en castellano hay concordancia del verbo con el sujeto, y en LSE hay concordancia de algunos verbos con el sujeto, el objeto e incluso el destinatario. A continuación se muestra un ejemplo:
CASTELLANO: Juan ha comprado las entradas
LSE: JUAN ENTRADAS COMPRAR

En LSE, el orden OV puede cambiar en algunas oraciones transitivas, mientras que en castellano siempre debe ser VO. Sin embargo, el orden SV sí debe respetarse en ambas lenguas.

Otras diferencias de la LSE con el castellano son: el género que no se marca en LSE, no hay artículos, existen clasificadores, el plural es descriptivo y existe la doble referencia.

Una propiedad importante de la LSE es la iconicidad, que es la propiedad que tienen los signos de asemejarse a aquello que sustituyen. En el caso de las frases escritas en LSE, las glosas tienen información principalmente semántica: en LSE el número de signos por frase (4,4) en muy inferior al número de palabras en castellano (5,9). Además, existe un vocabulario más reducido en LSE que en castellano: sinónimos en castellano se representan con el mismo signo en LSE.

Para el desarrollo de la arquitectura de traducción se generan dos tipos de documentos: de texto y de signos. Los documentos de texto contienen una serie de frases en castellano, mientras que los de signos contienen las frases en LSE (formadas por glosas) que son traducción de las anteriores en castellano.

Con estas frases se realizan los experimentos con la arquitectura de traducción estadística. Para ello, en primer lugar se generan de forma aleatoria 8 ficheros (listas) de texto con una lista de frases en castellano y lo mismo con las frases en LSE y, posteriormente, se utiliza el 75% de estas listas para el entrenamiento, el 12,5% para el ajuste y el otro 12,5% para el test, tal como muestra la Figura 1. De esta manera, 6 listas se agrupan en un único fichero de entrenamiento (*train.texto* para las frases en castellano y *train.signos* para las frases en LSE), una lista para el ajuste del sistema (*validacion.texto* y *validacion.signos*) y otra lista para el test final (*test.texto* y *test.signos*).

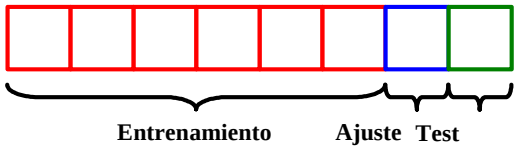

Figura 1: Esquema de la división del conjunto de frases en entrenamiento, ajuste y evaluación

En todos los experimentos, estas 8 listas van rotando con una estrategia de Round-Robin, de manera que se hacen 8 entrenamientos, 8 ajustes y 8 evaluaciones con distintas agrupaciones de listas y, posteriormente, se combinan los resultados y se calcula la media. Con esta estrategia se pretende aumentar la fiabilidad de los resultados obtenidos.

4 Arquitectura basada en Subfrases

La traducción estadística basada en modelos de subfrases (Figura 2) hace uso de un modelo de secuencias de palabras (extracción de subfrases) obtenido a partir de los pares de frases del conjunto de entrenamiento debidamente alineados. Para ello se utiliza el programa *GIZA++*.

Junto al modelo de subfrases se utiliza un modelo de lenguaje de la lengua destino para obtener una secuencia de signos dada una frase de entrada utilizando el módulo de traducción (programa *MOSES*). Finalmente, la secuencia

de signos se evalúa para calcular los errores en el proceso de traducción.

La traducción automática estadística se basa en la teoría de la información, aplicando el Teorema de Bayes para plantear el problema de traducción como un problema de maximización de la probabilidad de que la cadena del idioma destino (d) haya sido generada por la cadena origen (o). De tal manera que para conseguir *p(d|o)* se calcula *p(o|d)·p(d)*, donde *p(o|d)* es la probabilidad de que la cadena origen sea la traducción de la cadena destino (modelo de traducción), y *p(d)* es la probabilidad de ver aquella cadena destino (modelo de lenguaje de la lengua destino).

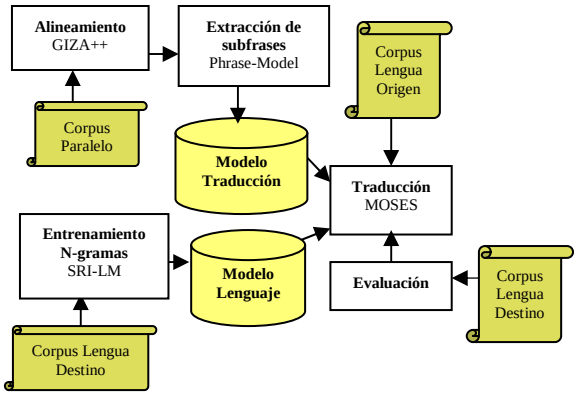

Figura 2: Arquitectura completa del sistema de traducción basado en subfrases

4.1 Modelo de Lenguaje

Para generar el modelo de lenguaje se emplean las herramientas *SRI-LM* (Stolcke, 2002), que permiten estimar modelos de lenguaje tipo n-grama. Un modelo de n-grama determina la probabilidad de una palabra dadas las n-1 palabras previas. Para ello, *SRI-LM* utiliza la herramienta *n-gram_count* para generar y manipular cuentas de n-gramas y estimar modelos de lenguaje a partir de ellas con un formato *.arpa*.

4.2 Modelo de Traducción

Para generar el modelo de traducción se emplean las siguientes herramientas: *Phrase_Extract*, *Phrase_Score* y *GIZA++*. Además, se necesita una colección de textos en lengua origen traducidos a lengua destino (corpus paralelo), para lo que se emplean las frases de la base de datos en castellano y su traducción a LSE (ficheros de entrenamiento).

GIZA++ (Och y Ney, 2000) es una herramienta software que permite alinear textos y aprender modelos de traducción basada en palabras a partir ellos, empleando en este sistema 5 iteraciones. De tal manera que, a partir del corpus bilingüe, se obtiene el alineamiento en los dos sentidos: origen-destino y destino-origen, para posteriormente combinar dichos alineamientos. Esta combinación puede realizarse de varias formas que se comentan a continuación. Notar que entre paréntesis se ha puesto unas siglas para poder identificar el tipo de alineamiento en la presentación de los resultados):

- **Destino-Origen (DO)**: Sólo se tiene en cuenta el alineamiento en un sentido: destino-origen (LSE-castellano en este caso), ignorando el alineamiento en sentido contrario. En este tipo de alineamiento, el proceso de alineamiento viene guiado por los signos: cada signo de LSE se alinea con una palabra en castellano, pudiendo quedar palabras sin alinear con ningún signo.
- **Origen-Destino (OD)**: En este caso el alineamiento que se tiene en cuenta es el contrario: origen-destino (castellano-LSE). En este caso, al contrario del anterior, el proceso de alineamiento viene guiado por las palabras que se alinean con un signo, pudiendo quedar signos sin alinear.
- **Intersección (I):** Partiendo de los dos alineamientos anteriores (destino-origen y origen-destino), se obtiene como alineamiento final los alineamientos intersección de ambos. Este tipo de alineamiento es el más exigente: permite tener alineamientos más fiables pero a costa de tener menos. Esta característica influye en la calidad del modelo de traducción cuando no se dispone de una gran cantidad de frases para su entrenamiento.
- **Unión (U):** En este caso, se toma la unión de los puntos de alineamiento en los dos sentidos (destino-origen y origen-destino). De esta manera, se obtienen puntos adicionales de alineamiento, consiguiendo más ejemplos para entrenar el modelo de traducción de palabras, pero también la calidad de los alineamientos considerados es menor que en el caso de la intersección.
- **Crecimiento (C):** En este tipo de alineamiento se toman los puntos de alineamiento de la intersección y, a continuación, se añaden los puntos de la unión que estén contiguos a los puntos de la intersección. Con esta posibilidad, se intenta buscar una situación intermedia entre la unión y la intersección consideradas

anteriormente. El objetivo es buscar un punto de equilibrio entre cantidad de puntos de alineamiento y su calidad.

- **Crecimiento diagonal (CD):** Igual que el alineamiento anterior, pero añadiendo sólo los puntos de la unión contiguos a la intersección y situados en la diagonal.
- **Crecimiento diagonal evitando casos de no alineamiento (CDP):** Este tipo de combinación es similar a la anterior con un postproceso adicional que tiene como objetivo evitar que ninguna palabra o signo se quede sin ningún punto de alineamiento. En el caso de que eso ocurra se añaden los alineamientos en uno u otro sentido necesarios.

En la siguiente figura (Figura 3) se muestran algunos ejemplos de alineamiento representados en forma de tablas donde en filas y columnas se ponen palabras o signos dependiendo del tipo de alineamiento. Los alineamientos se reflejan mediante el sombreado de las casillas correspondientes

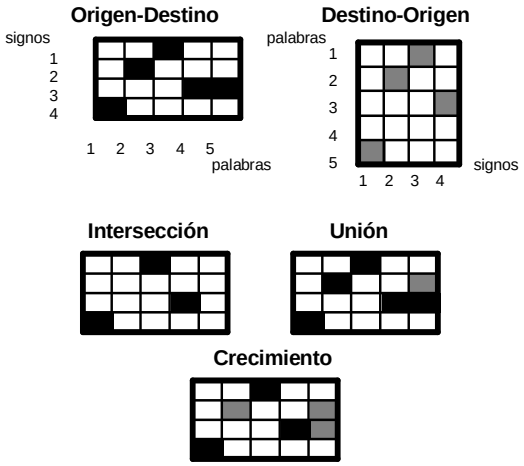

Figura 3: Distintos tipos de alineamiento entre frases

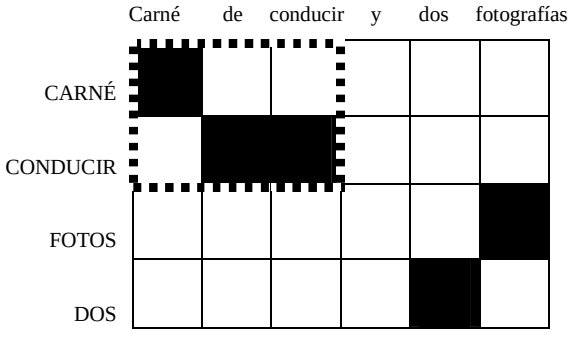

Figura 4: Ejemplo de alineamiento entre una frase en castellano y otra en LSE con una selección de secuencias: carné de conducir-CARNÉ-CONDUCIR

A partir del alineamiento, se obtienen las probabilidades de traducción para todos los pares de palabra (*w(d|o)* y *w(o|d)*) realizándose así una estimación de la tabla de traducción léxica más probable. La extracción de las secuencias de palabras y signos se realiza mediante el programa *phrase_extract.*

Esta extracción se realiza mediante el criterio de consistencia siguiente: un par de secuencias es consistente cuando todos los alineamientos que impliquen palabras o signos de esas secuencias caen dentro de esas mismas secuencias y no de otras palabras o signos. En la Figura 4 puede verse en un recuadro un ejemplo de secuencias de palabras y signos extraídas.

Finalmente, con el programa *phrase_score*, se calculan las probabilidades de traducción para todos los pares de subfrases en los dos sentidos: castellano-LSE y LSE–castellano.

4.3 Ajuste de los pesos de los modelos

En el archivo de configuración *moses.ini*, los distintos modelos implicados en el proceso de traducción (modelo de traducción y de lenguaje de la lengua destino) tienen unos pesos por defecto que no son los óptimos. En este paso, se ajustan estos pesos realizando la traducción de una lista de frases diferente a la utilizada en entrenamiento y a la que se utilizará en test.

Para realizar la traducción, se combinan linealmente las probabilidades de los modelos generados anteriormente cuyos pesos se quiere ajustar. Para ello, con el traductor *MOSES* se traduce y evalúa una lista de frases cuya traducción correcta se conoce, probando con distintos pesos y escogiendo los que dan los mejores resultados.

4.4 Traducción

Para la traducción se emplea el decodificador *MOSES*, que es un sistema de traducción automática estadística basado en subsecuencias de palabras que implementa un algoritmo de búsqueda que obtiene, a partir de una frase de entrada, la secuencia de signos que con mayor probabilidad corresponde a su traducción.

De tal manera que, utilizando los modelos de traducción y lenguaje obtenidos anteriormente, y con los pesos de las probabilidades ajustados, traduce una lista de frases de prueba, evaluando así el sistema.

5 Medidas de evaluación

Para evaluar cada uno de los experimentos que se realizan se emplean varias métricas que comparan la traducción que realiza el sistema con una traducción de referencia. La primera métrica utilizada es WER ("Word Error Rate", tasa de palabras con error), que calcula el número de inserciones, borrados y sustituciones de palabras cuando se compara la frase traducida por el sistema con la frase de referencia. Esta medida se basa en la distancia de edición o de Levenshtein. WER calcula la tasa de error de palabras teniendo en cuenta la posición de cada una de ellas dentro de la frase. Otra medida análoga es PER ("Position Independent Word Error Rate"), que no tiene en cuenta las posiciones de cada una de las palabras dentro de una frase.

Otra métrica utilizada es BLEU ("*BiLingual Evaluation Understudy*"), que es un método de evaluación de la calidad de las traducciones realizadas por sistemas de traducción automática. BLEU (calculado con la herramienta de NIST *mteval.pl*) compara los n-gramas de la frase generada por el sistema de traducción con los n-gramas de la frase de referencia, contando el número de n-gramas que coinciden independientemente de la posición.

La última medida empleada es NIST, basada en la BLEU con algunas modificaciones. En primer lugar, BLEU utiliza la media geométrica de la precisión de los n-gramas, mientras que NIST utiliza una media aritmética para reducir el impacto de bajas concurrencias para órdenes altos de n-gramas. Además, BLEU calcula la precisión de n-gramas utilizando pesos iguales para cada n-grama, mientras que NIST considera la calidad de la información que proporciona un n-grama particular en sí mismo (por ejemplo, cuanto menos frecuente sea un n-grama más peso se le asigna).

6 Experimentos iniciales

Para ajustar la arquitectura de traducción basada en subfrases, primero se realizan unos experimentos iniciales con los valores por defecto de los parámetros de configuración de los programas de la arquitectura, para ver los resultados de los que se parte.

Por otro lado, los ficheros necesarios para el entrenamiento, ajuste y evaluación en cada experimento se obtienen de la base de datos generada, tal como se explica en el apartado 3. De tal manera que se obtienen los resultados que se observan en la Tabla 2: la tasa de error es muy elevada superando el 40%.

WER	PER	BLEU	NIST
43,44	36,01	0,5168	6,9860

Tabla 2: Resultados iniciales con el sistema de traducción sin ajustar

7 Ajustes en la configuración de la arquitectura de traducción

Como se explicó anteriormente, la arquitectura de traducción basada en subfrases se compone de un módulo de entrenamiento del modelo de lenguaje, un módulo de entrenamiento del modelo de traducción y dos módulos de traducción: uno de ajuste y otro de test.

7.1 Modelo de lenguaje

Se comenzó ajustando los parámetros de las herramientas que generan el modelo de lenguaje (SRI-LM) sin conseguir mejoras estadísticamente significativas.

7.2 Modelo de traducción

En el proceso de entrenamiento del modelo de traducción, se analizaron dos parámetros principalmente: la longitud máxima de subfrase y el alineamiento.

En relación con la longitud máxima de subfrase (parámetro *-max-phrase-length*), se realizó un barrido con valores entre 7 y 40, obteniendo variaciones no significativas en la calidad de las traducciones conseguidas. Finalmente se fijó este parámetro a 20.

Donde los resultados sí ofrecieron variaciones significativas en la traducción fue a la hora de analizar la influencia del tipo de alineamiento utilizado.

El proceso de obtención del modelo de traducción comienza con la preparación del corpus paralelo de entrenamiento (en castellano y su traducción a LSE). Este corpus se codifica y se pasa el programa *GIZA++* que permite obtener los alineamientos en los dos sentidos: castellano-LSE y LSE-castellano para cada frase del corpus. A continuación, se combinan dichos alineamientos para obtener el alineamiento final, como se comentó en el apartado 4.2.

Los resultados para cada tipo de alineamiento pueden verse en la Tabla 3 y en la Figura 5.

A tenor de estos resultados se pueden extraer varias conclusiones. En primer lugar, el mejor resultado se obtiene teniendo en cuenta sólo el

alineamiento en sentido **destino-origen (D0)** (LSE-castellano, en este caso).

	WER	PER	BLEU	NIST
U	48,92	44,61	0,4417	6,2037
CDP	43,11	35,96	0,5193	6,9982
OD	40,99	34,43	0,5707	7,4017
CD	36,47	30,56	0,6393	7,9843
C	35,41	29,65	0,6653	8,2681
I	31,01	26,36	0,7044	8,597
DO	**28,29**	**20,76**	**0,7385**	**8,9927**

Tabla 3: Resultados con cada tipo de alineamiento en el modelo de traducción, ordenados de peor a mejor resultados.

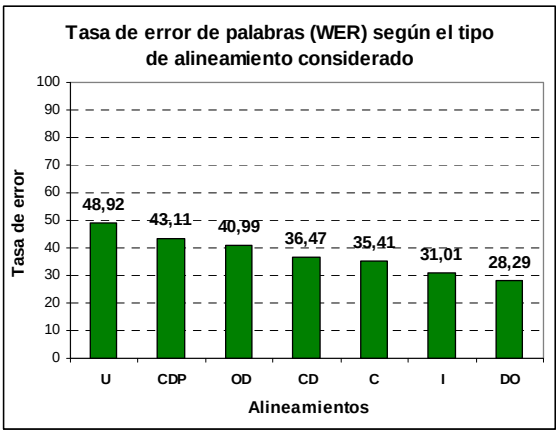

Figura 5: Comparativa de la tasa de error obtenida en traducción con cada tipo de alineamiento.

La traducción entre estas lenguas presenta unas características muy peculiares que apoyan estos resultados:

- En primer lugar, el número de signos por frase es mucho menor que el número de palabras en castellano: 4,4 y 6, respectivamente, en el corpus considerado en estos experimentos.
- El poder descriptivo de la LSE es más reducido.
- Por otro lado, al analizar las traducciones en LSE, se observa que el proceso de traducción de castellano a LSE es muy similar al proceso de comprensión o extracción semántica de una frase en castellano. La extracción semántica depende principalmente de un conjunto de palabras clave, mientras que existen otras palabras (palabras función) que no aportan significado y, por tanto, no tienen representación en la traducción a LSE.

Estas características avalan el hecho de que los alineamientos en los que el proceso está guiado por los signos (pudiendo quedar palabras sin alinear), reflejan mucho mejor la problemática de traducción de castellano a LSE. Por otro lado, si quisiéramos plantearnos este sistema de traducción para realizar un módulo de extracción semántica deberíamos configurarlo de esta misma manera: utilizando el alineamiento destino-origen.

Los resultados obtenidos con la **intersección (I)** de los alineamientos son también buenos, pero la diferencia con los mejores es estadísticamente significativa. En este caso, se eliminan algunos alineamientos (respecto el caso anterior) que pueden ser importantes a la hora de definir el modelo de traducción.

La combinación de alineamientos que utiliza la intersección más otros puntos de alineamiento de la unión (alineamiento con **crecimiento (C) y crecimiento diagonal (CD)**) presenta unos resultados intermedios dentro del estudio presentado. En este caso, muchos de los puntos de alineamiento rescatados de la unión serán de los obtenidos en el alineamiento destino-origen (DO).

Finalmente, a medida que vamos incluyendo los puntos de alineamiento definidos en la estrategia **origen-destino (OD)**, los resultados van empeorando sensiblemente, obteniéndose los peores resultados para el caso de utilizar la **unión (U)** de todos los puntos de alineamiento.

7.3 Resumen de Resultados

	WER	PER	BLEU	NIST
Inicio	43,44	36,01	0,5168	6,986
Fin	**28,29**	**20,76**	**0,7385**	**8,9927**

Tabla 4: Resumen de resultados con ajustes

Tras realizar los distintos ajustes en la arquitectura basada en subfrases, en la Tabla 4 puede verse un resumen de resultados, donde se muestra en sombreado los resultados iniciales y en negrita los conseguidos tras el proceso de ajuste. La mejora conseguida ha sido de una reducción relativa de la tasa de error del 35%

8 Conclusiones

Este artículo ha presentado una serie de ajustes de los parámetros de configuración de un sistema de traducción estadística basada en subfrases entre el castellano y la Lengua de Signos Española (LSE). La principal aportación ha sido el estudio del efecto que tiene el tipo de alineamiento en la calidad del modelo de traducción generado. Este tipo de alineamiento

es la base para generar el modelo de traducción basado en subfrases. Se ha demostrado que, utilizando únicamente el alineamiento en sentido destino-origen, mejoran notablemente los resultados (se consigue bajar la tasa de error de 43,11% a 28,29%). Esto es así porque, al traducir a LSE una frase en castellano, lo que se está haciendo principalmente es extraer información semántica, y esto se consigue mejor con el alineamiento destino-origen (DO). Este resultado es interesante, ya que demuestra que se podría utilizar un sistema de traducción estadístico basado en subfrases configurado de esta manera para hacer extracción semántica.

Es importante tener en cuenta que este comportamiento se da con un corpus limitado de 4080 frases. Al aumentar el corpus, los resultados podrían variar, ya que al disponer de más datos podría resultar mejor la intersección. No obstante, no se esperan cambios importantes, ya que, aunque aumente el número de datos, también lo harán las palabras desalineadas.

Agradecimientos

Este trabajo ha sido financiado por: Plan Avanza Exp Nº: PAV-070000-2007-567, ROBONAUTA (DPI2007-66846-c02-02) y SD-TEAM (TIN2008-06856-C05-03).

Bibliografía

Bungeroth J., Ney, H.,: Statistical Sign Language Translation. In Workshop on Representation and Processing of Sign Languages, LREC 2004, 105-108.

Cox, S.J., Lincoln M., Tryggvason J., Nakisa M., Wells M., Mand Tutt, and Abbott, S., 2002 "TESSA, a system to aid communication with deaf people". In ASSETS 2002, pages 205-212, Edinburgh, Scotland, 2002

Dreuw P., Neidle C., Athitsos V., Sclaroff S., and Ney H. 2008a. "Benchmark Databases for Video-Based Automatic Sign Language Recognition". In International Conference on Language Resources and Evaluation (LREC), Marrakech, Morocco, May 2008.

Dreuw, P., D. Stein, T. Deselaers, D. Rybach, M. Zahedi, J. Bungeroth, and H. Ney. 2008b "Spoken Language Processing Techniques for Sign Language Recognition and Translation. Journal Technology and Dissability. Volume 20 Pages 121-133. ISSN 1055-4181.

Dreuw, P., Stein D., and Ney H. 2009. "Enhancing a Sign Language Translation System with Vision-Based Features". LNAI, number 5085, pages 108-113, Lisbon, Portugal, January 2009.

Efthimiou E., and Fotinea, E., 2008 "GSLC: Creation and Annotation of a Greek Sign Language Corpus for HCI" LREC 2008.

Johnston T., 2008. "Corpus linguistics and signed languages: no lemmata, no corpus". 3rd Workshop on the Representation and Processing of Sign Languages, June 1. 2008.

Herrero, Ángel. 2004 "Escritura alfabética de la Lengua de Signos Española" Universidad de Alicante.

Koehn, Philipp. "*A Beam-Search Decoder for Factored Phrase-Based Statistical Machin Translation Models*". User Manual and Code Guide. Universidad de Edimburgo.

Koehn, Philipp., 2004. Training manual GIZA++.

Herrero, Ángel. 2004. "*Escritura alfabética de la Lengua de Signos Española*" Universidad de Alicante.

Morrissey S., and Way A., 2005. "*An example-based approach to translating sign language*". In Workshop Example-Based Machine Translation (MT X–05), pages109–116, Phuket, Thailand.

Morrissey S., Way A., Stein D., Bungeroth J., and Ney H., 2007 "*Towards a Hybrid Data-Driven MT System for Sign Languages. Machine Translation Summit (MT Summit)*", pages 329-335, Copenhagen, Denmark.

Och J., Ney H., 2000. "Improved Statistical Alignment Models". Proc. of the 38th Annual Meeting of the Association for Computational Linguistics, pp. 440-447, Hongkong, China.

San-Segundo R., Barra R., Córdoba R., D'Haro L.F., Fernández F., Ferreiros J., Lucas J.M., Macías-Guarasa J., Montero J.M., Pardo J.M, 2008. "*Speech to Sign Language translation system for Spanish*". Speech Communication, Vol 50. 1009-1020.

Stolcke A., 2002. "SRILM – An Extensible Language Modelling Toolkit".