

Analizador léxico del sistema de escritura alfabética de lengua de signos española

Memoria de suficiencia investigadora
Programa de doctorado: Aplicaciones de la Informática
Escuela Politécnica Superior
Universidad de Alicante

Encarnación Gómez Alcaraz

Diciembre 2011

Prefacio

Resumen

Esta memoria hace referencia a los cursos y trabajos que se realizaron durante los períodos de docencia e investigación tutelada del programa de doctorado *Aplicaciones de la Informática* del Departamento de Lenguajes y Sistemas Informáticos de la Universidad de Alicante.

El tema central de este trabajo es la investigación del sistema de escritura alfabética existente para la lengua de signos española, identificando sus componentes y forma de estructurarse para posteriormente realizar un analizador léxico que pueda determinar si la escritura en este sistema de un signo cumple las reglas estructurales del sistema.

Estructura

La memoria consta de dos partes. En la primera de ellas se muestra una ficha de las asignaturas cursadas durante el periodo de docencia, en ella se detalla el profesorado, los créditos, el tipo y la calificación obtenida y un resumen de los objetivos y tareas realizadas en la asignatura, además de un pequeño comentario sobre la aportación de sus contenidos a mi orientación investigadora. En la segunda parte, se detalla el trabajo que se ha realizado a lo largo del periodo de investigación tutelada, incluyendo un resumen con objetivos, resultados y conclusiones del trabajo. Finalmente, se añade la bibliografía de los artículos y libros más relevantes consultados para la realización del trabajo.

Agradecimientos

Con estas líneas quiero agradecer a quienes me han ayudado a hacer posible este trabajo. Empezando por mi director en la investigación, Mikel L. Forcada, por su paciencia, comprensión y sobre todo sus ideas y consejos para llevar a cabo este trabajo.

A Sergio Ortíz, Gema Ramírez y Miriam Antunes por sus conocimientos y dedicación personal para realizar este trabajo.

A Rubén Nogueira y Susana Tudela que me han transmitido toda su pasión hacia la lengua de signos y sin ellos no hubiera podido sacar este trabajo adelante.

A Antonio López por todo el cariño y apoyo que me ha dado día tras día y a toda mi familia que no se ha cansado de darme ánimos.

Índice

Prefacio.....	3
Parte I. Memoria del periodo de docencia.....	5
Cómo se Escriben y se Publican Trabajos Científicos.....	5
Tecnologías del Lenguaje Humano.....	5
Traducción Automática: Fundamentos y Aplicaciones.....	6
Técnicas Avanzadas en Traducción Automática	7
XML, Marcado de Textos y Bibliotecas Digitales	8
Parte II. Memoria del trabajo de investigación tutelado	9
1 Introducción.....	9
1.1 Lengua de signos.....	9
1.2 Lengua de signos española.....	10
1.3 Motivación y objetivos de la investigación.....	10
2 Estado de la cuestión.....	12
2.1 Sistemas de escritura.....	12
2.1.1 Sistemas no alfabéticos.....	12
2.1.2 Sistemas alfabéticos.....	15
2.2 Recursos lingüísticos existentes en SEA.....	17
2.2.1 Material de referencia.....	17
2.2.2 Textos.....	17
2.2.3 Diccionarios.....	20
2.2.4 Buscador de información en vídeos signados.....	21
3 Analizador léxico del sistema de escritura alfabética de lengua de signos española.....	23
3.1 Estudio y modificación de SEA.....	23
3.2 Desarrollo del analizador léxico de SEA para la identificación de los parámetros que conforman un signo codificados en SEA.....	29
3.3 Creación de un entorno colaborativo para el uso del analizador léxico de SEA, recursos lingüísticos y la inclusión de nuevo léxico en SEA.....	36
3.4 Creación de recursos lingüísticos para el par español → LSE de la plataforma Apertium...	44
4 Evaluación.....	46
5 Conclusiones.....	47
6 Trabajo futuro.....	48
7 Bibliografía.....	49
8 Anexo.....	50
A.Código del analizador del sistema SEA.....	50
B.Pasos para la obtención de las palabras a analizar.....	70
C.Pasos para probar/evaluar el analizador.....	70
D.Páginas web para obtener las palabras de uso frecuente.....	71
E.Signos no validados por el analizador léxico.....	73

Parte I. Memoria del periodo de docencia

El período de docencia se llevó a cabo durante los años académicos 2008–2009, cursando 20 créditos del programa de doctorado *Aplicaciones de la Informática* del Departamento de Lenguajes y Sistemas Informáticos de la Universidad de Alicante.

Los cursos realizados se detallan a continuación:

Cómo se Escriben y se Publican Trabajos Científicos

Profesorado: Forcada Zubizarreta, Mikel L. y Pérez Ortiz, Juan Antonio

Créditos: 4

Tipo: Fundamental

Calificación obtenida: 8,5

Una parte muy importante en la carrera de un investigador es difundir adecuadamente el trabajo realizado para que la ciencia y la técnica puedan seguir avanzando. Sin embargo, explicar lo que se ha hecho oralmente o en un documento que se quiere publicar no es una tarea sencilla. Este curso trata de iniciar a los investigadores en el proceso de escribir, publicar y comunicar los trabajos científicos de forma exitosa.

Los objetivos principales del curso son:

1. Comprender la necesidad de métodos especializados de documentación y comunicación para el trabajo científico y tecnológico.
2. Concienciarse de la importancia de los métodos de comunicación y documentación en el avance de la ciencia y la técnica, así como de la reproducibilidad de lo que se expone en una comunicación científica.
3. Conocer y asumir los mecanismos básicos de producción de documentos de naturaleza científica, así como los mecanismos de publicación de trabajos científicos.
4. Aprender a preparar y realizar presentaciones orales de trabajos científicos.
5. Conocer los mecanismos básicos de obtención de documentos científicos así como las fuentes bibliográficas básicas.

Con el curso aprendí cuáles son los diferentes tipos de publicaciones existentes (artículos en congresos y revistas, libros, *pósters*, etc.) y sus diferencias, cómo realizar la estructura básica de cualquier comunicación científica, cómo es el proceso de envío, revisión, aceptación (o no) de trabajos científicos a congresos, revistas, etc.

Como trabajo práctico se realizaron distintas presentaciones orales y *pósters* para reforzar los conocimientos adquiridos.

Considero este curso necesario y esencial para todo alumno que se inicie en la investigación científica, ya que a lo largo de su carrera deberá transmitir sus trabajos a través de publicaciones y presentaciones, por lo que tendrá que conocer los mecanismos convencionales para hacerlo.

Tecnologías del Lenguaje Humano

Profesorado: Ferrández Rodríguez, Antonio, Moreno Boronat, Lidia

Créditos: 4

Tipo: Fundamental

Calificación obtenida: 9

El procesamiento del lenguaje natural (PLN) es una parte esencial de la inteligencia artificial que investiga y formula mecanismos computacionalmente efectivos que faciliten la interrelación hombre/máquina y permitan una comunicación mucho más fluida y menos rígida que los lenguajes formales. Todo sistema de PLN intenta simular un comportamiento lingüístico humano; para ello debe tomar conciencia tanto de las estructuras propias del lenguaje, como del conocimiento general acerca del universo de discurso. De esta forma, una persona que participa en un diálogo sabe cómo combinar las palabras para formar una oración, conoce los significados de las mismas, sabe cómo éstos afectan al significado global de la oración y posee un conocimiento del mundo en general que le permite participar en la conversación. En este curso se realiza una breve introducción al PLN presentando la organización de los sistemas de comprensión del lenguaje natural (módulos de análisis léxico, sintáctico y semántico) y las aplicaciones del PLN que coexisten actualmente en este campo (traducción automática, acceso a bases de datos, extracción de información en bases de datos, recuperación o búsqueda de información, etc.).

Los objetivos principales de la asignatura son:

1. Introducir al alumno en el campo de la investigación de la comprensión del “lenguaje natural”
2. Formar al alumno en los fundamentos teóricos de la disciplina, preparándole de esta manera para el estudio y desarrollo de la tecnología apropiada a cada aplicación, así como para la asimilación de los cambios futuros.
3. Estudiar los módulos que componen cualquier aplicación de PLN: análisis léxico, sintáctico y semántico.
4. Estudiar las aplicaciones que necesitan del PLN: traducción automática, acceso a bases de datos, extracción de información en Bases de Datos, recuperación o búsqueda de información.

En la parte teórica del curso se explicaron los conceptos y problemas básicos que se derivan del procesamiento del lenguaje natural, así como sus aplicaciones, orígenes y tendencias actuales. Se realizó una introducción al análisis léxico, presentando conceptos generales sobre información léxica y su representación y se estudiaron algunos aspectos de la influencia de las restricciones semánticas en las entradas léxicas. Además, se presentaron los fundamentos del análisis sintáctico, modelos de representación de las estructuras sintácticas y algoritmos de análisis, y técnicas de análisis parcial y global. A continuación se vieron las características deseables en un lenguaje de especificación de la forma lógica con el fin de capturar e interpretar la semántica de una palabra, oración o texto. Y finalmente, se estudio el proceso de transformación de la forma lógica en la representación del significado final.

Para complementar la parte teórica del curso, en la parte práctica se vieron y utilizaron diferentes programas de cada una de las aplicaciones de PLN.

Con este curso he podido entender los objetivos que persigue el PLN y los mecanismos y aplicaciones que se utilizan para cada tarea en concreto.

Traducción Automática: Fundamentos y Aplicaciones

Profesorado: Forcada Zubizarreta, Mikel L. y Pérez Ortiz, Juan Antonio

Créditos: 4

Tipo: Fundamental

Calificación obtenida: 8

Esta asignatura es una introducción a la traducción automática (TA), en la que se ofrece una visión general de las tecnologías utilizadas y de los problemas que se presentan al abordar esta tarea, así como sus posibles soluciones.

Los objetivos principales del curso son:

1. Hacerse conscientes de las aplicaciones posibles de la traducción automática y de los problemas que comporta.
2. Conocer los conceptos lingüísticos básicos para discutir los problemas y las soluciones (morfología, sintaxis, semántica, etc.).
3. Conocer las operaciones y las estructuras lingüístico-computacionales sobre las que se basan los sistemas de TA (analizadores, diccionarios, gramáticas).
4. Conocer el estado actual de la cuestión y las soluciones más sobresalientes (sistemas de transferencia, de interlingua, estadísticos, híbridos).
5. Hacerse conscientes de la complejidad intrínseca de la traducción automática y de sus causas.
6. Ser capaces de implementar ejemplos de algunos de los módulos que forman parte de los sistemas de TA.

Al inicio del curso se realizó una introducción a qué es la traducción automática, el grado de dificultad de las tareas a realizar y sus posibles aplicaciones, como la asimilación, diseminación de información, preedición (adaptación) y postedición de textos, etc. Posteriormente, se describieron las estrategias básicas utilizadas por los programas de traducción automática, como son los sistemas de interlingua o los sistemas de transferencia. Para finalizar el curso, se presentaron algunas nociones básicas sobre cómo evaluar un sistema de traducción automática, y se evaluaron diversos traductores existentes.

Como trabajo práctico se estudiaron diversos sistemas de TA existentes y se hizo un informe sobre los principales problemas observados y la manera de solucionarlos. También se estudió el funcionamiento de una memoria de traducción como tecnología complementaria de la TA.

Este curso me ha aportado gran cantidad de conceptos y problemas que desconocía en un principio y me han ayudado a entender mejor la traducción automática, su aplicación y limitaciones.

Técnicas Avanzadas en Traducción Automática

Profesorado: Forcada Zubizarreta, Mikel L., Carrasco Jiménez, Rafael C.

Créditos: 4

Tipo: Fundamental

Calificación obtenida: 7

Este curso profundiza los conocimientos básicos sobre traducción automática que se ven en el curso de Traducción Automática: fundamentos y aplicaciones.

Los objetivos principales del curso son:

1. Conocer en profundidad algunas de las herramientas teóricas que subyacen a la implementación de los módulos de los sistemas de traducción automática.
2. Estudiar la implementación de estos módulos.
3. Ser capaz de implementar versiones reducidas de estos módulos en el laboratorio.

En este curso se realizaron varias sesiones sobre la teoría de lenguajes, autómatas y gramáticas, estadística y probabilidad. Estas sesiones fueron la base para la comprensión adecuada de las diferentes técnicas que se estudian en el curso y que son utilizadas para la implementación eficiente de los sistemas de TA, entre ellas:

- Transductores de estados finitos como técnica usada para cumplir las tareas de análisis, generación y búsqueda de equivalentes a partir de diccionarios de un sistema de TA.
- Desambiguadores léxicos estadísticos y basados en reglas para cumplir la tarea de desambiguación de las palabras homógrafas de una lengua.
- *Chunkers*, utilizados como detectores de secuencias de palabras, categorías, etc.
- Analizadores sintácticos, utilizados para generar una representación sintáctica de un sintagma.
- Alineamiento de textos bilingües y su aprovechamiento, por ejemplo, para la creación de memorias de traducción o para la traducción estadística.

Este curso además de complementarme los conceptos aprendidos en Traducción Automática: Fundamentos y Aplicaciones, me ha proporcionado una visión más completa del grado de dificultad y limitaciones de cada una de las fases por las que pasa un texto para ser traducido. Además, este curso me ha resultado útil para la realización del trabajo de investigación en el que he aplicado la técnica de transductores de estados finitos para el procesamiento léxico.

XML, Marcado de Textos y Bibliotecas Digitales

Profesorado: Carrasco Jiménez, Rafael C.

Créditos: Fundamental

Tipo: 4

Calificación obtenida: 6

En este curso se explica cómo se utilizan y se definen lenguajes de marcado así como las nuevas técnicas de búsqueda de transformación y de recuperación de la información. Teniendo en cuenta el desarrollo actual de las bibliotecas digitales (y de Internet), estas técnicas han adquirido una especial relevancia y permiten consultar, transformar y recuperar grandes volúmenes de información. El marcado de textos ha originado los sistemas HTML, SGML y XML, que constituyen, hoy en día, uno de los campos de desarrollo más activos.

Los objetivos principales del curso son:

1. Conocer cómo se utilizan y definen lenguajes de marcado así como nuevas técnicas de búsqueda, de transformación y recuperación de la información para este tipo de textos.
2. Comprender los trabajos y las líneas de investigación originadas a partir de estos temas.

En el curso se vieron métodos para la alineación de textos, recuperación de información, búsqueda de variaciones en textos y como comprimir documentos para poder buscar a posteriori información en ellos.

Los conceptos que he adquirido en este curso me han resultado bastante útiles en mi carrera profesional y para esta investigación. Entre los componentes relacionados con esta asignatura que trabajan con XML se cuentan los diccionarios y las reglas de transferencia estructural de Apertium.

Parte II. Memoria del trabajo de investigación tutelado

1 Introducción

1.1 Lengua de signos

La lengua de signos (LS), también denominada lengua de señas (A. Oviedo, 1997) es el principal sistema de comunicación utilizado por la comunidad de personas sordas.¹ Esta es una lengua visual y gestual, en lugar de oral-auditiva, una lengua natural de las personas sordas que les permite comunicarse y que no deriva de las lenguas orales, aunque sí que está influenciada por éstas igual que de otras LS.

Es una lengua que cuenta con los tradicionales niveles fonológicos, morfológico y sintáctico, y posee sus propias reglas gramaticales, que son diferentes a las de las lenguas orales (Rodríguez Ortiz 2005, p. 31), aunque de cualquier manera se pueden establecer analogías entre los niveles fonológicos de las lenguas orales y el nivel de articulación del signo en la lengua de signos. No es una lengua internacional (Pérez Cobacho et al 2001, p.14), aunque, en la actualidad, existe un Sistema de Signos Internacional (SSI), que es utilizado en reuniones internacionales como en el congreso de la Federación Mundial de Sordos (World Federation of the Deaf (WFD) congress) y en eventos como en los Juegos Olímpicos para Sordos. Su contenido se basa en signos propios, consensuados y de distintas lenguas. Además las lenguas de signos desarrollan dialectos y variedades cuando su uso se extiende. Según la Federación Mundial de Personas Sordas (WFD), hay más de 70 millones de personas sordas en el mundo y las distintas comunidades han ido creando, en cada país o en cada región, sistemas lingüísticos naturales que han dado lugar a las diferentes lenguas de signos existentes. Cada país e incluso dentro del mismo país existen diferentes lenguas de signos, como por ejemplo, en España existe la lengua de signos española y la lengua de signos catalana, y en otros países de habla hispana, aunque utilizan el español oral, tienen otras lenguas de signos, como por ejemplo la lengua de signos argentina o la lengua de signos colombiana.

El canal de emisión de los mensajes es la principal diferencia entre las lenguas de signos y las lenguas orales, donde en la LS su codificación se hace en el canal visual (los mensajes se codifican en señales gestuales) mientras que en las lenguas orales el canal utilizado es el aire (los mensajes se codifican en señales sonoras), y esta diferencia tiene una gran repercusión en las características del código de estas lenguas.

Las personas sordas utilizan la posición corporal, los gestos manuales, las expresiones faciales para transmitir significados, a esto es lo que le llamamos signo. Un signo es un complejo articulatorio que podemos dividir en diferentes elementos o parámetros. Estos parámetros son: configuración de la mano, orientación de la palma de la mano, lugar del espacio donde se realiza el signo, movimiento que realizan las manos para efectuar el signo, componentes expresivos no manuales que se ejecutan a la vez que el signo.

¹ La comunidad sorda es un colectivo de gente de los cuales la mayoría tienen deficiencias auditivas, totales o parciales.

1.2 Lengua de signos española

La lengua de signos española (LSE) constituye una lengua natural, completamente desarrollada y autónoma respecto a las lenguas orales y a otras lenguas signadas

En España es la lengua de signos adoptada a nivel estatal. La Ley 27/2007 de 23 de octubre,² reconoce y protege los derechos de la comunidad sorda a comunicarse o a recibir libremente información por cualquier medio de difusión, reconoce oficialmente las lenguas de signos españolas³ y regula los medios de apoyo a la comunicación oral de las personas sordas, con discapacidad auditiva y sordociegas.

Según la Encuesta de Discapacidad, Autonomía Personal y Situaciones de Dependencia del INE de 2008⁴, en España hay 1.064.100 personas sordas o con discapacidad auditiva y el número de personas que usan de lengua de signos en España supera las 400.000 personas.⁵

1.3 Motivación y objetivos de la investigación

Muchas personas con deficiencia acústica sufren analfabetismo funcional,⁶ tienen dificultad de entender correctamente el español, ya que no es su lengua materna, y la diferencia tipológica⁷ de la lengua de signos española respecto del español les impide tener un buen acceso a la comunicación, información y servicios prestados en nuestra sociedad. Y este acceso puede verse mejorado por las tecnologías existentes en la actualidad, ya que aunque la LSE es una lengua tradicionalmente gestual y visual, a día de hoy se dispone de los recursos necesarios para tratarla de manera computacional como una lengua. Esto es posible gracias a los esfuerzos invertidos en el diseño de sistemas de escritura para lenguas de signos, entre los que cabe destacar el sistema de escritura alfabético, en adelante SEA (Herrero et al., 1999) computacionalmente tratable por un ordenador. Éste sistema de escritura será una herramienta básica en este proyecto.

El objetivo de este proyecto es el crear una herramienta para procesar automáticamente la LSE y un portal web colaborativo con recursos lingüísticos en SEA. Se creará el primer analizador del sistema de escritura SEA, con el que se pretende identificar las diferentes partes que conforman un signo transcrito a esta notación y comprobar que este cumple los requisitos de formación por los que se rige este sistema de escritura. Con este analizador será posible una mayor difusión de este sistema de escritura, ya que con él se podrán crear recursos lingüísticos para facilitar su aprendizaje.

Asimismo, el uso del analizador resulta un elemento necesario a la hora de realizar diccionarios de lengua de signos codificada en SEA, ya que antes de incluir una entrada en estos diccionarios se podría validar su escritura utilizando el analizador a modo de corrector ortográfico.

2 <http://www.boe.es/boe/dias/2007/10/24/pdfs/A43251-43259.pdf>

Las referencias de URL que se presentan en este trabajo han sido revisadas en noviembre 2011.

3 Lengua de signos española (LSE), la lengua de las personas sordas, con discapacidad auditiva y sordociegas en España, según el texto de la ley 27/2007 y lengua de signos catalana (LSC) lengua de las personas sordas, con discapacidad auditiva y sordociegas en Cataluña, según el texto de la misma ley.

4 <http://www.ine.es/jaxi/menu.do?type=pcaxis&path=/t15/p418/a2008/hogares/p01/modulo1&file=pcaxis>

5 La lengua de signos catalana cuenta con unas 25.000 personas usuarias, de las que 12.000 son personas sordas. http://www.cnse.es/lengua_signos/lengua_sing_cata.html

6 El analfabetismo funcional, según la UNESCO, es una "inadecuada capacidad de alfabetización para las demandas de la vida cotidiana". Una persona analfabeta funcional sabe leer y escribir, pero no tiene un nivel suficiente para leer sin ayuda textos cotidianos en su vida.

7 La diferencia tipológica es al grado de semejanza o divergencia estructural, léxica, etc. que existe entre dos lenguas.

Dentro del proyecto se desarrollará un portal web colaborativo en el que diferentes usuarios podrán, por un lado, utilizar el analizador y varios recursos lingüísticos en SEA y por otro, crear nuevos datos para estos, ya que en él que se podrá indicar la notación en SEA de diferentes palabras de uso frecuente⁸ del español.

Además, dentro del portal colaborativo se incluirá el sistema de traducción automática de código abierto español→LSE escrita en SEA de la plataforma Apertium⁹ y la posibilidad de generar una plantilla de entradas para sus diccionarios monolingües y bilingüe para las palabras nuevas que se introducen en el portal colaborativo, y así poder mejorar la cobertura del citado traductor.

También se desarrollará un diccionario bilingüe español–lengua de signos en SEA dentro del portal en el que se podrán realizar búsquedas de las distintas palabras incluidas en los diccionarios de Apertium y las nuevas entradas introducidas y validadas del portal colaborativo, tanto en español como por las diferentes partes que conforman el signo.

Estas herramientas y recursos pretenden ser de utilidad para la enseñanza y aprendizaje de la lengua de signos española y contribuir en su normalización.

El proyecto que aquí se presenta contiene las siguientes tareas:

- Estudio y modificación de SEA¹⁰ (apartado 3.1)
- Desarrollo del analizador léxico de SEA para la identificación de los parámetros que conforman un signo codificados en SEA (apartado 3.2)
- Creación de un entorno colaborativo para el uso del analizador léxico de SEA, recursos lingüísticos y la inclusión de nuevo léxico en SEA (apartado 3.3)
- Creación de recursos lingüísticos para el par español→LSE de la plataforma Apertium (apartado 3.4)

8 Uso frecuente según diferentes páginas web, consultar en Anexo D

9 Apertium: Plataforma libre/de código fuente abierto para la traducción automática. <http://www.apertium.org>

10 SEA: sistema de escritura alfabética de la LSE

2 Estado de la cuestión

2.1 Sistemas de escritura

Aunque, como se dijo en el apartado 3.1, la lengua de signos es una lengua gestual y visual, a día de hoy existen varios sistemas de escritura para lenguas de signos. Existen dos grupos de sistemas para plasmar la lengua de signos sobre un papel: por un lado, los sistemas no alfabéticos de escritura de lenguas de signos, y por otro, los sistemas de escritura alfabética de la LSE.

En estos dos grandes grupos se representa el signo mediante una notación de base fonética donde se identifican las partes formales de los signos (configuración de las manos, orientación, movimiento, localización, punto de contacto con el cuerpo si hay, etc.) y son representados con un símbolo diferente o bien por caracteres gráficos o alfabéticos como si se tratara de escribir palabras de una lengua oral.

El uso de estos sistemas permite la comunicación escrita y están orientados a la transmisión de conocimientos. Las personas sordas podrán mediante la lectura representarse mentalmente los signos. Por ejemplo este tipo de escritura se puede utilizar en medios donde no es práctico o fácil introducir un canal visual-gestual, como puede ser el caso de la red de información de tiempo real Twitter,¹¹ donde utilizando un sistema de notación alfabética una persona con conocimientos de dicha escritura puede hacer un uso total de esta red. En la red social Twitter ya podemos encontrar a personas que utilizan la notación de escritura alfabética SEA.¹²

2.1.1 Sistemas no alfabéticos

Dentro del grupo de los sistemas no alfabéticos encontramos muchos sistemas, la mayoría icónicos,¹³ donde sus símbolos representan las formas de lo que están representando, son sistemas fonéticos que transcriben literalmente lo que se ve cuando se realiza el signo. Los más conocidos son Stokoe, Hamnosys y SignWriting.

- Stokoe

Stokoe es el promotor del primer sistema de transcripción que lleva su nombre y permite presentar los signos según unos parámetros gestuales independientes de la ordenación alfabética de la lengua oral.

Este sistema de escritura que se compone de 55 símbolos con los que se puede escribir cualquier signo. Donde 19 símbolos hacen referencia a la forma de la mano, 12 a lugares y 24 movimientos o actividades.

Los símbolos se escriben de forma lineal en un orden fijo; algunos símbolos tenían más de un significado dependiendo del lugar de la cadena en que se situasen.

Una de las ventajas y a la vez un inconveniente de este sistema es que fue diseñado para que pudiera escribirse con una máquina de escribir empleando un tipo de letra especial, y por ello está limitado a la hora de seleccionar los símbolos a utilizar.

11 <http://twitter.com/about>

12 <http://twitter.com/scheimiq>

13 Los sistemas icónicos están formados por símbolos que tienen una relación de semejanza con el objeto que representan.

Utiliza una simbología icónica respecto a la localización y el movimiento.¹⁴ Por ejemplo el símbolo '>' para indicar movimiento hacia la derecha, '<' hacia la izquierda, 'B' para la mano plana,¹⁵ 'C' para la mano curva, etc.

Una de las limitaciones que tiene este sistema es que no posee ningún elemento para la transcripción de la expresión facial y otros rasgos no manuales.

La notación Stokoe esta pensada para el análisis lingüístico y con representación lineal. Sus principales inconvenientes son su complejidad, que dificulta la interpretación y el tratamiento deficiente de los parámetros no manuales de los signos.(Guillem Masso, 2007)

Signo en BSL ¹⁶ de la palabra en inglés “What?”	
Visual http://www.signwriting.org/forums/linguistics/ling001.html	Stokoe
	

Tabla 1. Representación signada en BSL y en notación Stokoe de la palabra en inglés “what?”

- HamNoSys

El Sistema de Notación de Hamburgo (HamNoSys: Prillwitz, 1991) esta inspirado en el trabajo de Stokoe, pero con la diferencia de que este no esta limitado a los símbolos que se recogen en las teclas de la máquina de escribir, y que los símbolos son más icónicos.

Al igual que sucede con el sistema de Stokoe, los símbolos se ordenan linealmente. Y representan la localización, la configuración, el movimiento y la orientación de la mano.

La ventaja que tiene esta notación es que puede utilizarse con cualquier lengua de signos ya que es muy detallada y puede recoger las variaciones más pequeñas realizadas en los signos.

14 http://en.wikipedia.org/wiki/Stokoe_notation

15 Iconicidad respecto al alfabeto datilológico. El alfabeto dactilológico es la representación manual de cada una de las letras que componen el alfabeto de la lengua oral.

16 La lengua de signos británica (BSL) es la lengua de signos utilizada en Reino Unido.

Para tratarlo computacionalmente se ha creado un lenguaje de tipo XML llamado SiGML (Signing Gesture Mark-up Language), dentro del proyecto europeo ViSiCAST (Virtual Signing: Capture, Animation, Storage and Transmission, Bangham et al. 2000), y que ha superado algunas limitaciones del sistema (Elliott et al., 2000). Se pueden comprobar el resultado de la aplicación en la animación de los avatares desarrollados dentro del mismo proyecto como en su continuación eSign (Essential Sign Language Information on Government Networks, Elliott et al., 2007; Kennaway, Glauert y Zwitserlood, 2007).^{17 18}

Signo en LSE de la palabra en español “Avisar”		
Visual © www.sematos.eu	HamNoSys	SiGML
		<pre> <sigml> <hns_sign gloss="\$PROD:avisar"> <hamnosys_nonmanual> <hnm_mouthpicture picture="AVISAR"/> </hamnosys_nonmanual> <hamnosys_manual> <hamfinger2/> <hamfingerbendmod/> <hamextfingeru/> <hampalml/> <hamunderchin/> <hamclose/> <hammoveu/> <hamsmallmod/> <hammoveor/> <hamsmallmod/> <hamarcu/> </hamnosys_manual> </hns_sign> </sigml> </pre>

Tabla 2. Representación signada en LSE y en notación HamNosys y SiGML de la palabra en español “avisar”

- SignWriting

SignWriting¹⁹ (Sutton, 1995) es un sistema de escritura icónica que no sigue un orden lineal de los trazos.

Este sistema de escritura que se compone de símbolos que hacen referencia a la configuración de las manos, según la orientación de las palmas y los dedos, el movimiento de las manos, los dedos, el tipo de contacto existente y los parámetro faciales y corporales, que pueden indicar localización y símbolos de puntuación.

17 <http://www.visicast.co.uk/>

18 <http://www.sign-lang.uni-hamburg.de/esign/>

19 <http://www.signwriting.org>

Aunque por su carácter figurativo y no lineal hace costoso su procesamiento por ordenador, la creación del lenguaje SWML (SignWriting Markup Language, Costa, A.C.R. & Dimuro, G.P., 2001), en formato XML, ha facilitado la implementación de SignWriting en aplicaciones de procesamiento del lenguaje de signos. Es el que se utiliza en la herramienta VSigs²⁰ para crear la animación de signos a partir de texto (Papadogiorgaki et al., 2004).

<p>Signo en español “Avisar”</p>	
<p>Visual © www.sematos.eu</p>	<p>SignWriting http://www.signbank.org/SignPuddle1.5/searchword.php?ui=5&sgn=55&sid=845&search=avisar&type=any</p>
	

Tabla 3. Representación signada y en notación SignWriting del signo en español “avisar”

2.1.2 Sistemas alfabéticos

- Glosas

Las glosas son adaptaciones de la lengua de signos a la lengua oral escrita de los aspectos tanto gramaticales como léxico-semánticos que permiten profundizar en el análisis gramatical de las LS (Báez & Cabeza 2002: 669). Cada glosa representa el significado de la forma base, no marcada, del signo. Son utilizadas en lingüística para mostrar la gramática de una lengua a los usuarios de otra. Son muy útiles para estudiar la gramática de la lengua de signos, ya que en ellas se establece una relación directa entre el signo y el significante, pero cuenta como inconvenientes que en ellas no se describen los trazos del signo, ni cuentan con unos criterios claros y universales y se encuentran supeditadas a la lengua oral.

Español	Fui a Córdoba
Glosa	PASADO-YO-CORDOBA-IR-YA

Tabla 4. Ejemplo de oración en español y su correspondencia en glosa

²⁰ <http://vsigs.iti.gr:8080/Vsigs>

- SEA

El Sistema de escritura alfabética (SEA, Herrero et al., 1999) es una notación para escribir la LSE basada en el alfabeto latino.

El SEA esta basado en criterios fonológicos. Al igual que cuando escribimos una lengua oral sólo tomamos aquello que es fonológicamente distintivo (no escribimos igual que hablamos), la escritura mediante SEA no refleja todo lo que aparece en el espacio, sino aquello fonológicamente distintivo. Con este sistema se obtienen todas las ventajas que puede aportar la escritura alfabética, como son su procesamiento, la comunicación electrónica, la creación de corpus, la enseñanza de la lengua.

El sistema está compuesto por un conjunto de caracteres alfabéticos que, según una clasificación y su lugar dentro del signo escrito, hacen referencia a la bimanualidad del signo, su lugar, contacto, configuración de la mano, orientación de la mano, dirección, forma de movimiento, espacio gramatical y el lugar de la mano activa respecto a la pasiva.

Se ha diseñado para que sea aplicable a todas las lenguas de signos existentes en el mundo, de las que hay 120 lenguas reconocidas, dos de ellas en España.²¹

Signo en español “Avisar”	
Visual © www.sematos.eu	SEA
	chémawe

Tabla 5. Representación signada y en notación SEA del signo en español “avisar”

Como mencionamos en el apartado 1.3, éste es el sistema que utilizaremos en nuestro proyecto. Hemos decidido utilizar la notación SEA puesto que SEA no requiere ni un hardware ni un software especial para ser tratado, como puede ser HamNoSys, donde necesitaríamos un teclado especial para poder teclear sus símbolos, o un software adicional que nos ayudara a introducirlos, o el sistema SignWriting, que por su carácter figurativo y no lineal hace costoso su procesamiento por ordenador. Todos los símbolos utilizados en la notación SEA se pueden teclear tanto con un teclado convencional como con dispositivos móviles que tengan la posibilidad de introducir también caracteres acentuados.

²¹ <http://accesibilidadenlaweb.blogspot.com/2008/12/sistema-de-escritura-alfabtica-de-la.html>

Nos hemos decantado también por esta notación porque por su diseño, con mínimas modificaciones, se podrá aplicar a todas las LS existentes, y si en alguna de las LS se diera el caso de que falta por considerar alguna de sus característica bastaría con identificar las partes (orientación, posición, lugar, etc.) que no estén codificadas y, tras un análisis, asignarles una codificación alfabética.

que pensamos que en un futuro se podría ampliar el sistema de notación SEA para que fuera válido para todas las LS existentes, ya que, aunque esta diseñado para poderse utilizar con todas las lenguas de signos, en la actualidad los componentes que tiene codificados se basan este sistema no tiene detectadas todas las posibles . Basta con identificar las partes (orientación, posición, lugar, etc.) que no estén codificadas y, tras un análisis, asignarles una codificación alfabética.

2.2 Recursos lingüísticos existentes en SEA

2.2.1 Material de referencia

Un material básico de referencia sobre la escritura de SEA es el libro “*Escritura alfabética de la lengua de signos española: Once lecciones*”²², cuyo objetivo es el de explicar paso a paso el sistema de escritura de la lengua de signos española. Aunque, como ocurre con todas las lenguas, el sistema SEA ha evolucionado, por lo que, como es lógico, estos cambios no están reflejados en este libro, pero que seguramente estarán en próximas ediciones del libro.

2.2.2 Textos

Como se nombró en el apartado 2.1, el uso de sistemas como SEA permite la comunicación escrita y la transmisión de información en una lengua donde su canal principal de comunicación es el visual-gestual.

Se pueden encontrar textos en SEA en la red de información de tiempo real Twitter, donde sólo se permite el uso de texto (140 caracteres). En ella podemos comprobar como se comunican varias personas de la comunidad sorda y hacen un uso total de esta red en su lengua,²³ sin necesidad de incluir vídeos signados.

22 Ángel Herrero Blanco (2003). *Escritura alfabética de la lengua de signos española: Once lecciones*. Universidad de Alicante. Servicio de publicaciones.

23 <http://twitter.com/scheimiq>



Ilustración 1: Mensajes en Twitter utilizando el sistema de escritura SEA

Igualmente, podemos encontrar en la Biblioteca de Signos de la Biblioteca Virtual Miguel de Cervantes²⁴ las primeras lecciones de la “Gramática visual de la Lengua de Signos Española”,²⁵ que están disponibles tanto en vídeo signado como escritas con la notación SEA.²⁶

Como se verá a lo largo del proyecto, el sistema de escritura SEA ha ido evolucionando²⁷ respecto material de referencia nombrado en el apartado 2.2.1, esta es la razón por la que la notación en SEA que se aprecia en la Ilustración 1 y la Ilustración 3 tengan rasgos diferentes, como por ejemplo, la aparición del símbolo “:.”. Los recursos que se están nombrando tendrían que revisarse para adaptarse a la nueva notación.

24 <http://bib.cervantesvirtual.com/seccion/signos>

25 Lecciones signadas de morfología y sintaxis de la LSE

26 http://bib.cervantesvirtual.com/portal/signos/linguistica/gramatica_visual/morfologia/leccion01/SEA.shtml
http://bib.cervantesvirtual.com/portal/signos/linguistica/gramatica_visual/morfologia/leccion02/SEA.shtml
http://bib.cervantesvirtual.com/portal/signos/linguistica/gramatica_visual/morfologia/leccion03/SEA.shtml

27 En la actualidad, las actualizaciones en la notación de SEA están disponibles en http://wiki.apertium.org/wiki/Traductor_de_español_a_LSE escrita en colaboración con el equipo de diseño la primera versión del sistema de escritura.



- ▶ **Presentación**
- **¿Qué es la gramática?**
- **La gramática es como un juego**
- **Gramática e iconicidad**
- **Tres dimensiones de la iconicidad**
- **Morfología y sintaxis**
- **El modelo fonológico del signo**



Gramática Visual de la LSE.
Morfología. Lección 1.
Introducción.

Ilustración 2: Primera lección de la "Gramática visual de la Lengua de Signos Española" con video signado

Gramática visual de la Lengua de Signos Española. Lección 1. Presentación. Ángel Herrero y colaboradores

PRESENTACIÓN

Ahsemiuco òamaagre

Sòamehy eumiuc-miac sòamacy óqwraeheb sl jeimauweb, ¿aeilep sòamehy eumiuc-miac semaho emec sm òameadohe?

Slz omawraeheb ynaëmiagroc sòamacy óqwraeheb sm abomaullob.

¿T.iäcrela òameaviwub äemau-1maa eämaadohb?

Eylsòameaha Lcjrcie eylsòameaca omeljcie eylseha ezpejrcie sm abòamaupo sòamahy âecru ömaals eylsj ômecie sèiha vèimac UNIVERSALES, eylsòameaha Lcjrcie eylsòameaca omeljcie o-uma sòamacy óqwraeheb.

Eimaufahe o-uma sòamahy âecru sm eämapeb soáha ômeuku sòameaha Lcb eilsc oámyfahercye eylsj ômewreihorcie.

Sm òameadohe, sòameaha Lcb sv òamiadohe, sòameaha Lcb veimeuq Omawraihob slz omawraeheb, sm ymomuazu sm aëmiugrel y.òazpuld, eimaufahe eilsòameaha Lcb srb eimiugrewy sòameaha Lcb éímaudahery y.òazpuld.

¿QUÉ ES LA GRAMÁTICA?

Ilustración 3: Primera lección de la "Gramática visual de la Lengua de Signos Española" en SEA

2.2.3 Dicionarios

Los diccionarios son recursos lingüísticos importantes para contribuir a la difusión de una variedad estándar de la lengua de signos española.

En 2002 se plantea el proyecto DILSE III para editar el primer diccionario normativo de la LSE, *DILSE - Tesoro de la lengua de signos española*,²⁸ realizado por Fundación CNSE.²⁹ Es un diccionario descriptivo, de uso y bilingüe (LSE–castellano). Cada lema está acompañado de sus acepciones correspondientes y en cada una de ellas se puede acceder a los vídeos de signos correspondientes. Sus vídeos están signados por profesionales sordos expertos en LSE que tienen la lengua de signos española como lengua materna. Además, muchas de las definiciones que se pueden encontrar en el diccionario han sido cedidas por la Real Academia Española de la lengua (RAE), de sus diccionarios normativos: *DRAE*³⁰ y *Diccionario del Estudiante*.³¹ La búsqueda de los signos se puede hacer en castellano o a través de los parámetros formativos del signo. Adicionalmente, los signos se encuentran transcritos en SEA (sistema de escritura alfabética propia de la LSE, Herrero et al., 1999 y Herrero, 2003).

En 2008 se empezó a desarrollar el traductor de código abierto de español–lengua de signos escrita bajo la notación SEA³² desarrollado por la empresa Prompsit Language Engineering, SL³³ en colaboración con el profesor Ángel Herrero del Departamento de Filología Española, Lingüística General y Teoría de la Literatura de la Universidad de Alicante, utilizando la tecnología Apertium. Para su desarrollo ha sido necesario crear un bilingüe de español–lengua de signos en SEA, que en la actualidad tiene aproximadamente unas 3500 palabras de ámbito general.

2.2.4 Buscador de información en vídeos signados

La manera natural de comunicarse de una persona sorda se realiza mediante un canal lingüístico visual; la lengua de signos incluye diferentes componentes de carácter visual hechas con las manos, la cara, y el torso de la persona que signa,³⁴ para transmitir su significado. Por lo que la generación de vídeos signados es una manera de representación gestual visual para informar a las personas sordas signantes.

En la Biblioteca de Signos de la Biblioteca Virtual Miguel de Cervantes³⁵ se ha desarrollado un buscador de información en vídeos signados³⁶ que permite buscar signos que aparezcan en los distintos vídeos de las lecciones “*Gramática Visual de la LSE*” que se han mencionado en el apartado 2.2.2, para ello los signos a buscar pueden escribirse tanto en español como en SEA. Esto es posible porque los vídeos están transcritos a español y a SEA. En la Ilustración 4 podemos ver el resultado de la búsqueda de la palabra en español “*lengua*” y el fragmento del video donde se menciona.

28 <http://www.fundacioncnse.org/tesorlse>

29 La Fundación CNSE (Confederación Estatal de Personas Sordas) es organización no gubernamental creada en 1936 que representa a las personas sordas del estado español y sus familias a través de sus federaciones y asociaciones. Desde ella se viene luchando, trabajando y representando los derechos e intereses de las personas Sordas; actuando como coordinadora de todo el movimiento asociativo de este colectivo en España. La CNSE está especialmente vinculada a cualquier acción que directa o indirectamente mejore la calidad de vida de las personas Sordas en cualquiera de sus ámbitos: social, laboral e, indudablemente, educativo (CNSE 2004) <http://www.cnse.es>

30 Real Academia Española. (2001). *Diccionario de la lengua española* (22a. ed.). Madrid, España: Espasa Calpe

31 Real Academia Española. (2005). *Diccionario del estudiante*. Madrid, España: Santillana

32 <http://translator.prompsit.com/demosea/>

33 <http://www.prompsit.com>

34 Signar es realizar una comunicación no verbal, mediante signos, gestos, posturas y posiciones en lugar de palabras.

35 <http://bib.cervantesvirtual.com/seccion/signos>

36 http://bib.cervantesvirtual.com/portal/signos/materiales/buscador_signos/

BIBLIOTECA DE SIGNOS

Por favor, introduzca el término que desea buscar
Atención, recuerde introducir correctamente los términos de su bu

lengua

Resultados 1 - 4 de aproximadamente 64 palabras. (1.325 segundos)

[1.-Gramática Visual de la LSE. Morfología. leccion 1. Parte 1](#)
ESTE PROYECTO OBJETIVO CUAL? LENGUA-DE-SIGNOS ESPAÑOL GRAMÁTICA APRENDER.
 ¿aeilep sòamehy eumiuc-miac semaho emec sm òameadohe? Slz omawraeh eb ynaèmiagroc sòamacy óqwraeh eb sm abomaullob.
[\[reducir contexto\]](#)

[2.-Gramática Visual de la LSE. Morfología. leccion 1. Parte 1](#)
LENGUA TODO DIFERENTE
 Eylsòameaha Lcjr cie eylsòameaca omeljicie eylseha ezpejr cie
[\[ampliar contexto\]](#)

[3.-Gramática Visual de la LSE. Morfología. leccion 1. Parte 1](#)
LENGUA TODO HABER GRAMÁTICA.
 eyl sòameaha Lcjr cie eylsòameaca omeljicie o-uma sòamacy óqwraeh eb.
[\[ampliar contexto\]](#)

[4.-Gramática Visual de la LSE. Morfología. leccion 1. Parte 1](#)
DENTRO LENGUA GRUPO COINCIDIR.
 soáha òmeuku sòameaha Lcb eilsc oámyfahercye eylsj òmewreihorcie.
[\[ampliar contexto\]](#)

Páginas de resultado: 2 3 4 5 6 7 8 9 10 >>

Cargando...

ESTE PROYECTO OBJETIVO CUAL?
 LENGUA-DE-SIGNOS ESPAÑOL
 GRAMÁTICA APRENDER.
 ¿aeilep sòamehy eumiuc-miac semaho
 emec sm òameadohe? Slz omawraeh eb
 ynaèmiagroc sòamacy óqwraeh eb sm
 abomaullob.

Illustration 4: Buscador de información en vídeos signados de la Biblioteca de Signos

3 Analizador léxico del sistema de escritura alfabética de lengua de signos española

3.1 Estudio y modificación de SEA

Los signos tienen unas características determinadas que hacen que se diferencien uno de otro. Estos se realizan en un lugar (del espacio o del cuerpo, con o sin contacto con este), con una configuración y orientación determinada de las manos (en algunos signos de una mano y en otros de ambas), y además, los signos dinámicos se caracterizan también por su movimiento direccional o interno. Estas características constituyen los elementos mínimos por los que se compone un signo, y la variación de uno de estos elementos significa la aparición de un signo diferente. En el sistema de escritura SEA se han codificado y agrupado todas las características.³⁷ Durante el desarrollo de este trabajo, se han introducido una serie de cambios en la notación SEA, todos ellos en colaboración con el equipo que diseñó dicho sistema, y asumidas para posteriores versiones y que serán publicados en un futuro. Algunos de estos cambios se deben a que se han detectado algunas características que no se habían considerado previamente, lo que ha hecho que SEA haya evolucionado para incluirlas (como por ejemplo, la clasificación H que se verá más adelante). Otros cambios realizados se deben a que los caracteres utilizados no estaban disponibles en todos los dispositivos móviles, por lo que se ha estudiado, junto con el equipo que diseñó SEA, otro tipo de codificación y se han sustituido por otros que sí que se podían teclear tanto por un teclado convencional como por un dispositivo móvil (como ejemplo tenemos el carácter 'ö' que se ha sustituido por 'õ'). Otro cambio, el más significativo, se ha producido para evitar ambigüedades y que el texto sea computacionalmente tratable. En SEA se hacía la distinción entre signos monomanuales y bimanuales³⁸ utilizando el carácter espacio ' ' entre sus componentes cuando era bimanual. La utilización de este carácter daba lugar a confusión, puesto que podía entenderse como la marca de espacio entre signos. Una persona humana puede intuir esta distinción, pero a la hora de procesarlo computacionalmente, no se podía crear fácilmente una regla sencilla para determinar cada uno de los casos, por lo que se optó por sustituir el carácter de espacio por el carácter '!'. Se puede pensar que este carácter podría tener un problema similar al anterior, pero el carácter '!' siempre va seguido de un espacio cuando indica puntuación, mientras que en este caso iría seguido de un carácter alfabético.

SEA antiguo	SEA evolucionado
sc òamegreb	sc:òamegreb

Tabla 6. Representación del signo en español “libro” en SEA antes y después de la modificación del símbolo que indica la bimanualidad

Como se pudo ver en el apartado 2.2.2, los cambios producidos ya se están empezando a aplicar.

Una vez incluidos los cambios en SEA, este ha clasificado todas las características de los signos en 10 grupos y han de escribirse siguiendo un orden que es el mismo que el de su articulación.

Los grupos en los que se clasifican estas características son:

37 En otros estudios se han incluido los componentes no manuales como uno de los parámetros formativos de los signos, pero en la notación SEA no se ha tipificado.

38 Los signos pueden ser articulados con una o ambas manos, por lo que se pueden clasificar como signos monomanuales aquellos signos que se articulan con una sola mano, sobre la que recae la información pertinente; y los signos bimanuales, los que se articulan con ambas manos; dentro de estos, podemos encontrar dos grupos, lo que se pueden realizar con las dos manos simétricas (con todos sus parámetros formativos iguales) y los que se realizan manteniendo una mano fija (articulador pasivo) y la otra en movimiento.

- S: Bimanualidad.
Este grupo incluye las codificaciones³⁹ que indican que el signo es de tipo bimanual (en el que intervienen las dos manos) y representa el tipo de simetría en las manos
- L: Lugar.
En este grupo están las codificaciones de los diferentes lugares del cuerpo, como por ejemplo la cabeza, el ojo, la nariz, etc.
- C: Contacto.
En este caso solo existe una codificación dentro de este grupo, en la que si existe es que hay contacto entre las manos y el cuerpo y su ausencia indica que no hay contacto.
- Q: Configuración.
Dentro de este grupo se encuentran las codificaciones de las configuraciones que puede tomar la mano, entre ellas, si la mano está abierta o cerrada, con los dedos juntos o separados, etc.
- O: Orientación.
En este grupo se incluye la codificación de las diferentes orientaciones que puede tomar los dedos de la mano.
- D: Dirección.
Dentro de este grupo están las codificaciones de las diferentes direcciones que se realiza en el signo. Dentro de este grupo existe una serie de codificaciones que solamente pueden aparecer en signos bimanuales.⁴⁰
- F: Forma de movimiento.
En este grupo están las codificaciones de las distintas formas de movimiento que se realizan en el signo. Dentro de este grupo existe una serie de codificaciones que solamente pueden aparecer en signos bimanuales.⁴¹
- G: Espacio gramatical.
En este grupo están las codificaciones de los diferentes lugares gramaticales.
- H: Lugar de la mano activa respecto a la pasiva.
Dentro de este grupo están las codificaciones de los diferentes lugares que toma la mano activa respecto a la mano pasiva.
- NM: No manual
Dentro de este grupo están las codificaciones de los diferentes de acciones no manuales que se realizan en el signo.

Junto con esta clasificación y codificación⁴² de las distintas características del signo, en SEA se establece una serie de reglas que indican el orden en el que deben de aparecer cada una de estas codificaciones. Estas reglas se pueden dividir en dos grandes grupos, las que afectan a los signos monomanuales (los que se signan con una sola mano, que es la mano activa⁴³) y los signos bimanuales (los signos en los que interactúan las dos manos, la activa y la pasiva⁴⁴)

39 Estas codificaciones se puede encontrar en el código del analizador implementado para este trabajo y que esta disponible en el Anexo A dentro de los paradigmas Simbolos_X donde X hace referencia al nombre del grupo.

40 Las codificaciones del grupo D que empiezan por 'j' solo puede aparecer cuando el signo es bimanual, ya que indica que en el movimiento las manos se mueven juntas hacia el mismo lugar.

41 Las codificaciones del grupo F que empiezan por 'j' solo puede aparecer cuando el signo es bimanual, ya que indica que en el movimiento las manos se mueven juntas hacia el mismo lugar.

42 Las codificaciones de las características, y a qué grupo pertenecen, están disponibles en http://wiki.apertium.org/wiki/Traductor_de_español_a_LSE y en el libro de Ángel Herrero Blanco (Herrero, 2003)

43 La mano activa es sobre la que recae la información del signo, y en los signos bimanuales, es la que tiene la condición de dominante.

44 La mano pasiva es la mano que tiene la condición de no dominante en los signos bimanuales.

A continuación se muestra el orden general en el que tiene que escribir un signo, y esta representado por el nombre del grupo por el cual se clasifican.

- Signos monomanuales
 - LCQODF
 - GQODF
- Signos bimanuales

En estos signos, las estructuras que preceden al carácter ':' hacen referencia a la mano pasiva y las estructuras posteriores al carácter ':' hacen referencia a la mano activa

 - Signos bimanuales de mano activa
 - S:⁴⁵ LCQODF
 - S : GQODF
 - GS : QODF
 - GS : GQODF
 - Signos bimanuales de mano pasiva
 - S1⁴⁶ QOH : LCQODF
 - GS1 QOH : QODF
 - L1⁴⁷ S1 QOH : LQODF
 - L2⁴⁸ S1 QOH : LQODF
 - L2 H : LQODF
 - L2 CQODF

Se ha mostrado el orden general de las estructuras; hay que tener en cuenta que no todos los signos se componen de todas las estructuras, ya que, por ejemplo, hay signos que no tienen movimiento, por lo que estos no tendrían en su escritura la estructura F.

Existen una serie de normas y excepciones dentro de cada orden general, y son las que se comentan a continuación:

- Las reglas de simplificación de SEA (omisiones). En el sistema de escritura SEA se pueden omitir algunas de las partes, para que no quede un signo escrito muy largo, por ejemplo, QUERER es “ynòazy” donde se ha omitido la orientación puesto que esta sería la orientación natural de la mano, el signo sin omisión sería el siguiente “ynòamizy”.
 - El lugar lateral con el brazo flexionado no hace falta escribirlo (lugar eb)
 - Por ejemplo la palabra “pastor” que en SEA es “ebömewraeheb” (G: eb Q: ö O: me F: wraehe F: b) puede simplificarse en “ömewraeheb” (Q: ö O: me F: wraehe F: b)
 - La orientación de la palma hacia el cuerpo no hace falta escribirla (orientación natural)
 - Por ejemplo, la palabra “cariño” que en SEA es “km.òmazub” (L: km C: . Q: ò O: ma D: zu F: b), puede simplificarse en “km.òzub” (L: km C: . Q: ò D: zu F: b)
 - La orientación de la palma en continuidad con el brazo flexionado tampoco (ma, me)

45 Símbolo que indica bimanualidad, a la izquierda del símbolo se representan las características de la mano pasiva, y a su derecha las características de la mano activa.

46 S1 hace referencia a un subconjunto de S (este subconjunto solo tiene un elemento que es 's' que hace referencia a que existe una simetría en el plano vertical)

47 L1 hace referencia a un subconjunto de L. (este es un subconjunto esta formado por las codificaciones que hacen referencia a la muñeca y al codo de la mano activa)

48 L2 hace referencia a un subconjunto de L (este es un subconjunto esta formado por las codificaciones que hacen referencia a la muñeca, antebrazo, brazo y codo de la mano pasiva)

- Por ejemplo, la palabra “tú” que en SEA es “eleme” (G: el Q: e O: me) puede simplificarse en “ele” (G: el Q: e)
- Todas los grupos son opcionales (pueden estar o no representados en el signo) a excepción del grupo Q de la mano activa que debe estar presente siempre, y se puede dar el caso de ser la única para identificar un signo, como ocurre con la palabra “jefe” que en SEA es “J”⁴⁹, una de las codificaciones del grupo Q.
- Las codificaciones de los diferentes componentes del grupo Q se dividen en tres subgrupos, codificaciones en mayúscula (QMAYUS, letras mayúsculas que representan el alfabeto dactilológico), codificaciones de número (QNUM números del 0 al 9 y QNUMORDINAL números ordinales) y el resto (QR). Dentro de un signo, el grupo Q puede aparecer repetida en los siguientes casos:
 - QNUM seguido de QR
 - “líder”⁵⁰ → somigy:3iãmejway donde
S:s Q:o O:mi H:gy SB:⁵¹ QNUM: 3 QR: iã O:me D: jway
 En este caso la configuración QNUM indica que el dedo medio (el tercero) de la mano activa se encuentra estirado y la configuración QR que los dedos pulgar y medio de la mano pasiva se cierran, y al aparecer estas dos configuraciones juntas se indica que las dos manos se entrelazan juntando ambas configuraciones.
 - QMAYUS puede estar repetido varias veces. Se utiliza sobre todo en los signos que representan palabras deletreadas.
 - “lse” → LSE donde
QMAYUS: L QMAYUS: S QMAYUS: E
 Este signo lo que realiza es un deletreo de la palabra lse utilizando la configuración de la mano de acuerdo la forma que tiene esta en el alfabeto dactilológico.
 - QR seguido de uno o varios QMAYUS
 - “ojo” → jeJO donde
L: j QR: e QMAYUS: J y QMAYUS: O
 Este signo se realiza en el lugar del ojo derecho, su configuración de la mano es con el dedo índice estirado y el deletreo de las letras J y O según el alfabeto dactilológico.
 - QNUM puede estar repetido varias veces. Se utiliza sobre todo en los signos que representan los números.
 - “92” → 92 donde **QNUM: 9 y QNUM: 2**
 Estas configuraciones hacen referencia al signado del número cardinal en lengua de signos.

49 En la web del proyecto Spread the sign <http://www.spreadthesign.com/es/> se puede comprobar que el signo que se realiza para la palabra jefe es haciendo con la mano derecha la letra J del alfabeto dactilológico.

50 Se puede ver esta formación de configuración en el video correspondiente a este signo en <http://www.spreadthesign.com/es/>

51 SB es un conjunto con un único elemento que es ':' que hace referencia al símbolo de separación de bimanualidad.

- QNUMORDINAL seguido de QR
 - “santa alianza” → somimucy:4°eãmèzi donde
S:s Q: o O: mimu H:cy SB: : QNUMORDINAL:4° QR:eã O: me D: zi
 Este signo es bimanual, donde la configuración de la mano pasiva indica que es con la mano abierta y los dedos estirado, la orientación de la palma de la mano en continuidad con el brazo es hacia la izquierda y hacia abajo y la situación de la mano activa respecto de la pasiva es con contacto a la izquierda; y la configuración de la mano activa es con la forma del numero ordinal indicado y el dedo indice estirado y el pulgar en paralelo a este, con una orientación de la palma de la mano en continuidad con el brazo hacia el frente y el tipo de dirección que se realiza es con contacto arrastrado hacia la izquierda.
- El grupo C contiene a las codificaciones que representan el contacto con el cuerpo, por lo que si forma parte del signo, debe de estar detrás de la estructura L, nunca puede aparecer sin esta estructura. No puede haber contacto si no se especifica el lugar donde tiene que producirse. Es decir, la estructura CQ no puede existir, esta debería ser LCQ
 - serio → z.olleu donde
L: z C:. Q:o F:llu
 Este escritura del signo indica que este se realiza en la nariz con contacto y con los dedos estirados y separados y que la forma de movimiento que realiza es flexionando los dedos hacia abajo.
- El grupo D puede aparecer una o varias veces seguidas.
 - “gracias a” → s:chòawepob donde
S:s SB:: L:ch Q:òà D: we F:po F:b
 En este signo solamente se realiza en una dirección, que en este caso es un movimiento curvo.
 - “naufragio” → sc:òãmègruewu donde
S:sc PB:: Q: òã O: mè D: grue y D: wu
 En este signo se eralizan dos tipos de direcciones, primero se mueve en una dirección curva de 90° hacia abajo y al frente seguido de un cambio de dirección de 180° hacia abajo.
- El grupo F puede aparecer una o varias veces seguidas.
 - “no importar” → s:ömamacre donde
S: s SB: : Q: ö O: mama F: cre
 En este caso solamente hay una forma de movimiento que es sin desplazamiento en el espacio, con un sólo giro de muñeca de 180° hacia el frente.
 - “peinar” → cm.aëmamafaheb donde
L:cm C:. Q: aë O: mama F: fahe y F: b
 En este signo la primera codificación de forma indica que el signo realiza un abatimiento de muñeca en dirección al frente y la segunda forma indica que este abatimiento se realiza dos veces.
 - “color” → ch.omamylszp donde
L: cj C: . Q: o O: mamy F: ls, F: z y F: p
 El signo realiza un cambio en la forma de la configuración donde se flexionan sucesivamente los dedos juntamente con un roce y un punteado en el aire.

- Los grupos F y D pueden aparecer intercambiados en el orden, por ejemplo, dentro de la regla L C Q O D F, esta podría aparecer como L C Q O F D.
 - “sartén” → aēmeweb donde
 - **Q:** aē **O:** me **D:** we **F:** b
En estos casos se realiza primeramente la codificación relativa a la dirección seguida de lo indicado en la forma.
 - “aburrido” → chòamigrewu **F:** gre **D:** wu
En estos casos se realiza primeramente la codificación relativa a la forma seguida de lo indicado en la dirección.

- Los grupos D y F, además de poder aparecer intercambiados, pueden aparecer repetidos dentro de estos cambios. Estas son las opciones posibles documentadas que se pueden formar:

◦ DFF	◦ DDFFF	◦ FFFD
◦ DFFF	◦ FDD	◦ FFDF
◦ DDF	◦ FFD	◦ FDF
◦ DDFF	◦ FFDD	

- Cuando el espacio no es L sino G, no va seguido de C, pues G se compone de los espacios lateral y central, no de lugares corporales, por lo que no puede haber contacto con dicho espacio. Esto sucede tanto en la parte del signo de la mano pasiva como de la mano activa.
- Cuando el espacio se antepone a los símbolos de bimanualidad, C no es posible excepto en el caso siguiente, en el que el espacio sigue siendo un lugar L del cuerpo.
 - L1 S Q O H : L Q O D F
- Los casos de signos bisilábicos⁵² se representan mediante un guión y las estructuras que se modifican. Después del guión, sólo representamos los parámetros que han cambiado, que pueden ser uno o varios (si son varios, estos seguirán el mismo orden descrito arriba). Si el signo
 - Por ejemplo, si un signo presenta una configuración Q inicial y otra final diferente, la estructura que tendría sería: S:LQODF-Q
 - “azul” → chòamimuzy-auma donde
L: ch **Q:** àa **O:** mimu **F:** zy **SS:**⁵³ - **Q:** au **O:** ma
el signo tiene una configuración “àa” inicial y una final “au”, además, también cambia la orientación que inicialmente era “mimu” y cuando cambia la configuración cambia a “ma”
 - La configuración no es obligatoria detrás del guión, así que esta estructura también es posible S:LQODF-F
 - “bien” → eimemacro-cre donde
Q: ei **O:** mema **F:** cro **SS:** - **F:** cre

52 Por convenio y por analogía a la escritura de las lenguas orales, se utiliza el concepto de sílabas para indicar distintas partes que forman la escritura de un signo en SEA. La primera sílaba del signo será aquella que va delante del símbolo del guión y la posterior, segunda sílaba, y así sucesivamente. Un signo puede estar formado de varias sílabas.

53 SS es un conjunto con un único elemento que es '!' que hace referencia al símbolo de separación de sílaba o signo compuesto.

el signo realiza dos formas de movimiento pero una después de la otra, en este caso un realiza un movimiento sin desplazamiento en el espacio con un giro en la muñeca de 180° hacia el cuerpo y después otro igual pero hacia el frente.

- Los signo compuesto se representan mediante un guión, al igual que los signos bisilábicos, pero la diferencia de estos es que después del guión se escribe de nuevo todo el signo, siguiente la estructura general S:LQODF-S:LQODF

- “incógnita” → alémamucuywu-olememupe donde

G: al **Q:** é **O:** mamu **D:** cuy **D:** wu **SS:** - **G:** ol **Q:** e **O:** memu **F:** pe

- Cuando el espacio se antepone a los símbolos de bimanualidad, C no es posible excepto en el caso siguiente, en el que el espacio sigue siendo un lugar L del cuerpo.

L1 S1 Qp Op H : L Q O D F

- “golf” → ohsaëmicy:aëmegriu **L:**ohs **Q:**aë **O:**mi **H:**cy: **Q:**aë **O:**me **D:**griu

3.2 Desarrollo del analizador léxico de SEA para la identificación de los parámetros que conforman un signo codificados en SEA

Un programa de análisis automático del sistema SEA (sistema de escritura alfabética de la lengua de signos española) sería un recurso provechoso para comprobar si los signos escritos en SEA cumplen los requisitos de formación por los que se rige este sistema de escritura, y con ello poder generar diccionarios de signos en SEA. Con la elaboración de un analizador léxico de SEA se podría identificar cada uno de los parámetros que conforman el signo y su orden, y si este se valida correctamente, se podría incluir dentro de un diccionario.

Estudiando las reglas de formación del SEA vistas en el apartado 3.1, podemos decir que SEA es un lenguaje formal, puesto que es un lenguaje cuyos símbolos primitivos y reglas para unir esos símbolos están formalmente especificados. El alfabeto⁵⁴ Σ de este lenguaje es:

$\Sigma = \{a, b, c, d, e, f, g, h, i, j, k, l, m, n, ñ, o, p, q, r, s, t, u, v, w, x, y, z, A, B, C, D, E, F, G, H, I, J, K, L, M, N, Ñ, O, P, Q, R, S, T, U, V, W, X, Y, Z, à, á, â, ã, ä, å, è, é, ê, ë, ì, í, î, ï, ð, ó, ô, õ, ö, ù, ú, ü, û, 1, 2, 3, 4, 5, 6, 7, 8, 9, 0, ^, \cdot, -, \cdot, \cdot\}$

Y además, se trata de un lenguaje regular porque puede ser representado y reconocido por un autómata finito determinista⁵⁵ (AFD) o un autómata finito indeterminista (AFI). Hemos utilizado el AFI porque permite representaciones más compactas. En la Ilustración 5 se puede observar uno de los casos, donde en el estado O existen dos transiciones posibles para la el símbolo del alfabeto 'z', por lo que dada la cadena de entrada 'z' existen dos posibles estados en los que se puede encontrar el autómata tras la lectura.

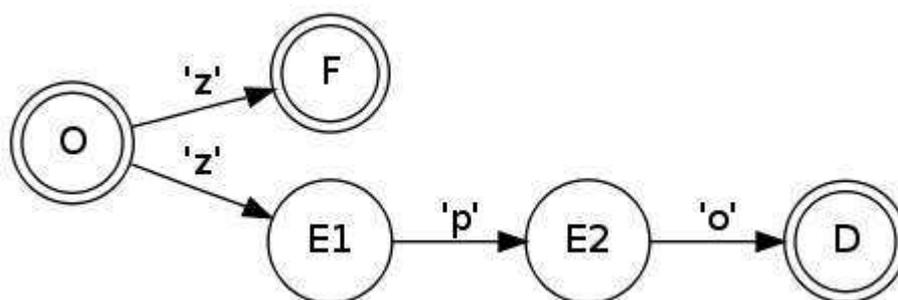


Ilustración 5: El estado O tiene dos posibles transiciones para el símbolo “z”

54 Un alfabeto Σ es un conjunto finito y ordenado de símbolos.

55 Un autómata finito determinista (AFD) es aquel que solo permiten una transición desde un estado. Un autómata finito indeterminista (AFI) es la generalización de los AFD en la que se permite que de un estado parta más de una transición etiquetada con el mismo símbolo o que exista más de un estado inicial. Es posible transformar un autómata indeterminista en uno determinista.

Para realizar el analizador léxico se ha realizado el autómata finito indeterminista, en la siguiente ilustración se muestra el AFI con las transiciones representadas por grupos de cadenas de entrada.

Las transiciones del autómata son:

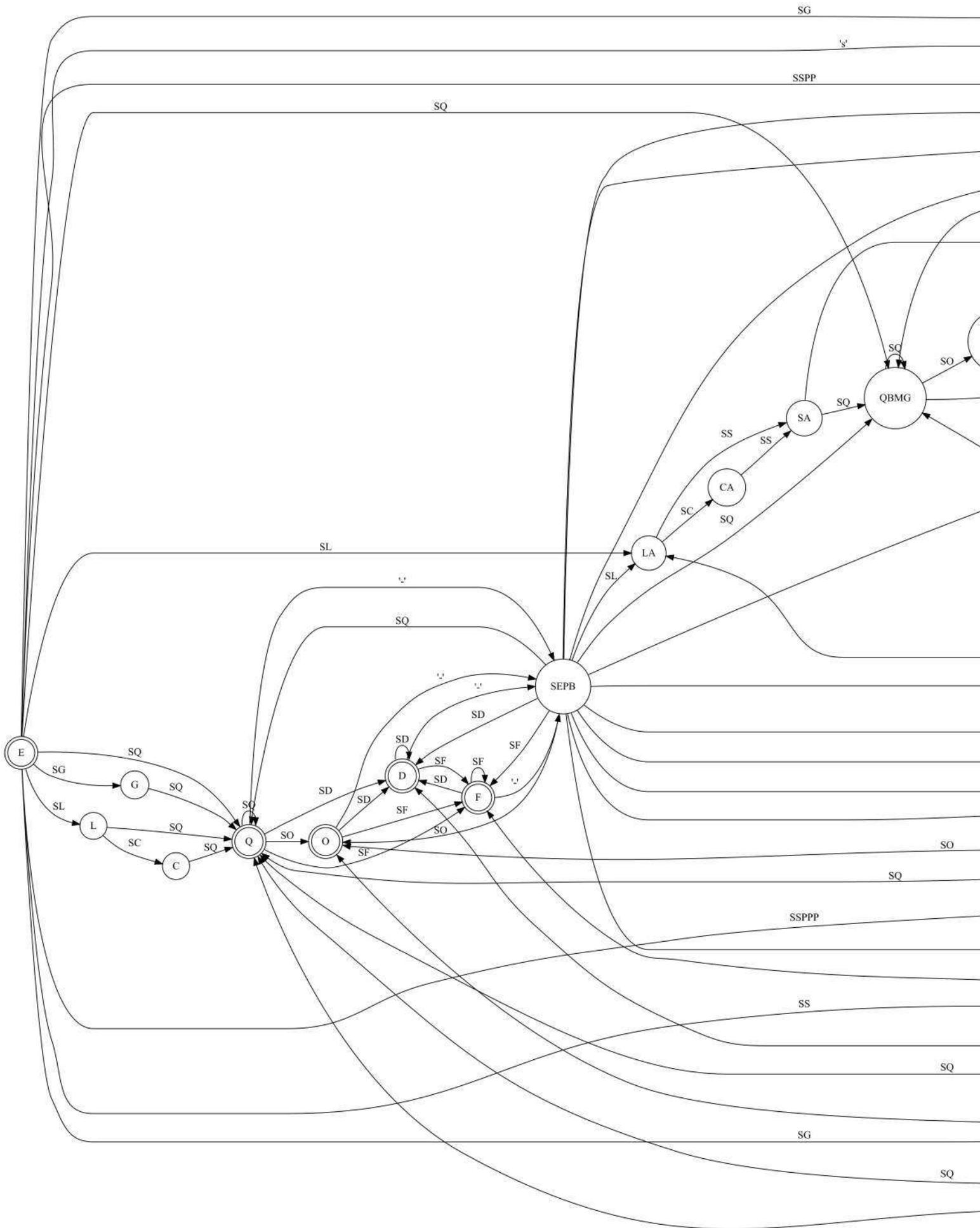
- El símbolo '-' que representa la parte bisilábica o compuesta de un signo.
- El símbolo 's' que indica bimanualidad (es un subconjunto del grupo S)
- Las cadenas de SL y SLB hacen referencia a las codificaciones del grupo L.
- Las cadenas de SC hacen referencia a las codificaciones del grupo C.
- Las cadenas de SD hacen referencia a las codificaciones del grupo D a excepción de aquellas que empiezan por el carácter 'j' (codificaciones que solamente pueden aparecer en signos bimanuales).
- Las cadenas de SDJ hacen referencia a las codificaciones del grupo D (incluidas aquellas que empiezan por el carácter 'j').
- Las cadenas de SF hacen referencia a las codificaciones del grupo F a excepción de aquellas que empiezan por el carácter 'j' (codificaciones que solamente pueden aparecer en signos bimanuales).
- Las cadenas de SFJ hacen referencia a las codificaciones del grupo F (incluidas aquellas que empiezan por el carácter 'j').
- Las cadenas de SG y SGB hacen referencia a las codificaciones del grupo G.
- Las cadenas de SH hacen referencia a las codificaciones del grupo H.
- Las cadenas de SO hacen referencia a las codificaciones del grupo O.
- Las cadenas de SQ hacen referencia a las codificaciones del grupo Q.
- Las cadenas de SS hacen referencia a las codificaciones del grupo S.
- Las cadenas de SO hacen referencia a las codificaciones del grupo O.
- Las cadenas de SSPP hacen referencia a las codificaciones 'ah', 'oh', 'sah', 'soh', 'sih', 'suh'.
- Las cadenas de SSPPP es un subconjunto de SSPP y hace referencia a las codificaciones 'sah', 'soh', 'sih', 'suh'.
- Las cadenas de SREP son las codificaciones que indican repetición.
- Las cadenas de SNM son las codificaciones no manuales.

Cada uno de los estados representa uno de los grupos de la notación SEA vistos en el apartado 3.1 o son estados a los que se llega después de haber leído uno de los símbolos de bimanualidad o de separación de sílaba.

El etiquetado de los estados hace referencia al tipo grupo de la notación SEA que han aceptado

- El estado E es el estado de inicio del autómata.
- Los estados C, CA, CB y CBM son los estados a los que se llega tras la lectura de una de las codificaciones del grupo de contacto (C).
- El estado DJBM es al que se llega tras leer una codificación del grupo de dirección (D) y al estado D cuando se ha leído el un subgrupo de codificaciones que solamente pueden aparecer en signos bimanuales. Ambos estados son estados de aceptación de la cadena.
- El estado FJBM es al que se llega tras leer una codificación del grupo de forma (F) y al estado F cuando se ha leído el un subgrupo de codificaciones que solamente pueden aparecer en signos bimanuales. Ambos estados son estados de aceptación de la cadena.
- Los estados G, GB, GBM, GBMQ, GBMS son aquellos a los que se llega tras la lectura de una de las codificaciones del grupo de espacio gramatical (G). Además, el estado GB es un estado de aceptación.
- Los estados HQ, HS son los estados a los que se llega tras leer una de las codificaciones del grupo de lugar de la mano activa respecto a la pasiva (H).

- Los estados L, LA, LB, LBM son los estados a los que se llega tras leer una de las codificaciones del grupo de lugar (L). El estado LB, es además, un estado de aceptación.
- Los estados O, OBMQ, OBMS, OJBM son aquellos a los que se llega tras la lectura de una de las codificaciones del grupo de orientación (O). Los estados O y OJBM son estados de aceptación.
- Los estados Q, QBM, QBMG, QBMS son aquellos a los que se llega tras la lectura de una de las codificaciones del grupo de configuración (Q). De ellos, los estados Q y QBM son estados de aceptación.
- Los estados S, SA y SG son aquellos a los que se llega tras la lectura de una de las codificaciones del grupo de bimanualidad (S).
- Los estados SP y SPQ son aquellos a los que se llega cuando se acepta un subconjunto de S que está formado por una única codificación que es el carácter 's'.
- El estado SPP es al que se llega cuando se acepta un código del conjunto 'ah', 'oh', 'sah', 'soh', 'sih', 'suh'.
- El estado SPPP es al que se llega cuando se acepta un código del conjunto 'sah', 'soh', 'sih', 'suh'.
- Los estados SEPB y SEPBB son aquellos a los que se llega cuando se ha leído la codificación de separación de sílaba.
- Los estados SEPG, SEPQ y SEPS son aquellos a los que se llega cuando se ha leído la codificación de separación de sílaba.
- El estado REP es al que se llega cuando se lee un código del conjunto 'bs', 'rs' que indican repetición del signo. Además, este estado es de aceptación.
- El estado NM es el estado al que se llega cuando se lee un código del grupo de no manuales (NM) y este estado es de aceptación.





En este proyecto se ha utilizando la tecnología Apertium para generar el analizador.

Como se vio en el apartado 3.1, la notación del sistema SEA incluye una serie de códigos que se pueden agrupar según las características a las que representan; estas se han identificado en unidades léxicas y junto con una serie de entradas que identifican a las transiciones que puede haber (estas cumplen las reglas de formación de SEA), además se han establecido unas entradas de excepción para contemplar unos signos que no cumplen con las normas de SEA al ser estos especiales. Estos signos son protocolo → s:n-fòãmale y pañal → s:b-g.òamumawa⁵⁶ en los que el símbolo '-' no significa que el signo es bisilábico ni compuesto, sino indica que el lugar L puede estar codificado como "n.-f" (lugar nuca-frente, la mano pasiva estaría en el lugar de la nuca y la mano activa en la frente) y "b-g." (glúteo-ingle, la mano pasiva estaría en el lugar del glúteo y la mano activa en la ingle) pero solamente en estos casos, por ello no se ha puesto en el conjunto de codificaciones de L para que no se creen falsos signos con este lugar. Todos estos códigos y entradas se han plasmado en un diccionario en el formato XML de los diccionarios de Apertium.

Se ha utilizado la librería Ittoolbox para construir el autómeta. La librería Ittoolbox es una colección de herramientas que sirve para construir transductores de letras a partir de diccionarios definidos en un formato particular XML. Los programas más importantes son It-comp (compilador), que lee los diccionarios y los transforma en transductores de letras mínimos, y It-proc (procesador), que procesa ficheros con los transductores de letras para generar una salida.

El compilador It-comp se apoya en el uso de agrupaciones de cadenas (paradigmas) que son convenientes tanto para gestionar el diccionario de forma abreviada, como para realizar la construcción y la minimización de los transductores de forma eficiente.

El procesador It-proc usa los transductores generados por It-comp para procesar un fichero de entrada y generar la salida resultante de la transducción. Utiliza diversos algoritmos que se pueden seleccionar mediante *flags*. Por ejemplo, -a, la opción por defecto, se usa para análisis morfológico (generación de formas léxicas a partir de formas superficiales), -g para generación morfológica (el proceso inverso, la generación de formas superficiales a partir de las formas léxicas). En este trabajo se ha usado la opción -a combinada con la opción -c para que no aplique ninguna gestión particular con las letras mayúsculas y minúsculas, puesto que en SEA las letras mayúsculas están incluidas dentro de su alfabeto, siendo un tipo de codificación para uno de sus grupos, en concreto de Q, que representa que la configuración de la mano toma la forma de la letra según el alfabeto dactilológico. La salida del transductor es la clasificación en de la cadena en los distintos grupos de SEA. En la tabla 7 podemos ver el resultado de analizar la palabra en español avisar que en SEA se escribe "chémawe".

```
$ echo "chémawe" | lt-proc -c analizador_sea.bin
^chémawe/(lugar=ch)(configuración=é)(orientación=ma)(dirección=we)$
```

Tabla 7. Ejecución y salida del analizador para la palabra en español "avisar" escrita en SEA

Del resultado obtenido se puede observar que la cadena ha sido validada correctamente, la estructura que sigue este signo es LQOD y su interpretación es:

Lugar: ch Mentón.

Configuración: e El dedo índice flexionado y los demás cerrados.

Orientación: ma La orientación de la palma de la mano en continuidad con el brazo hacia arriba.

Dirección: we Recto hacia el frente.

56 En <http://www.sematos.eu/lse-p-pa%C3%B1al-6048-es.html> puede verse signada la palabra diplomático y en <http://www.sematos.eu/lse-p-pa%C3%B1al-6048-es.html> la palabra pañal

Como se verá en el apartado 3.3, el analizador puede utilizarse dentro del portal colaborativo creado en este proyecto. En él se podrá escribir una o varias palabras escritas en SEA y se mostrará para cada una de ellas si la palabra está escrita correctamente y la estructura que se ha detectado con el analizador (notar que para un signo se puede haber detectado dos o más posibles estructuras, ya que al tratarse de un AFI, puede estar en varios estados finales, se mostrarán todas las que se hayan detectado).

Illustration 6: Analizador de SEA

El código del analizador léxico está disponible en el Anexo A y en el paquete del traductor automático de español–lengua de signos de Apertium y los pasos para la obtención de los signos y probarlos con el analizador están especificados en los Anexos B y C.

3.3 Creación de un entorno colaborativo para el uso del analizador léxico de SEA, recursos lingüísticos y la inclusión de nuevo léxico en SEA

Uno de los objetivos por el que se desarrolla el entorno colaborativo es para ofrecer a los usuarios de lengua de signos la posibilidad de introducir nuevos signos en SEA que aún no están recogidos en los diccionarios del traductor español → LSE de Apertium.

Otro objetivo es el de facilitar el aprendizaje de la notación SEA, por lo que se han realizado una serie de interfaces de usuario para acceder a diversos recursos, como es el caso del analizador y así identificar las diferentes partes que conforman un signo escrito en esta notación y comprobar que este cumpla los requisitos de formación por los que se rige este sistema de escritura. O bien la interfaz de diccionario en la que se pueden buscar tanto las palabras en español como en lengua de signos.

Este entorno colaborativo será de acceso restringido, por lo que será necesario darse de alta. Dentro del entorno existen dos roles para los usuarios (*expertos* y *anónimos*). Esta distinción es para permitir o denegar el acceso a diversas partes del entorno, puesto que los “*expertos*” serán usuarios experimentados en la notación SEA, mientras que los usuarios “*anónimos*” serán aquellos que conocen o quieren aprender la notación SEA. En la Ilustración 7 se puede ver la base de datos que se ha utilizado para realizar el portal colaborativo. Para gestionar los usuarios y sus perfiles se utilizan las tablas *clave*, *rol* y *usuario*;⁵⁷ y para la gestión de las palabras y su sea correspondiente se utilizan las tablas de *palabra*, *sea*, *estado*, *categoría*, *idioma* e *histórico_palabra_sea*.⁵⁸ Y en la Ilustración 8 la pantalla de entrada al entorno colaborativo.

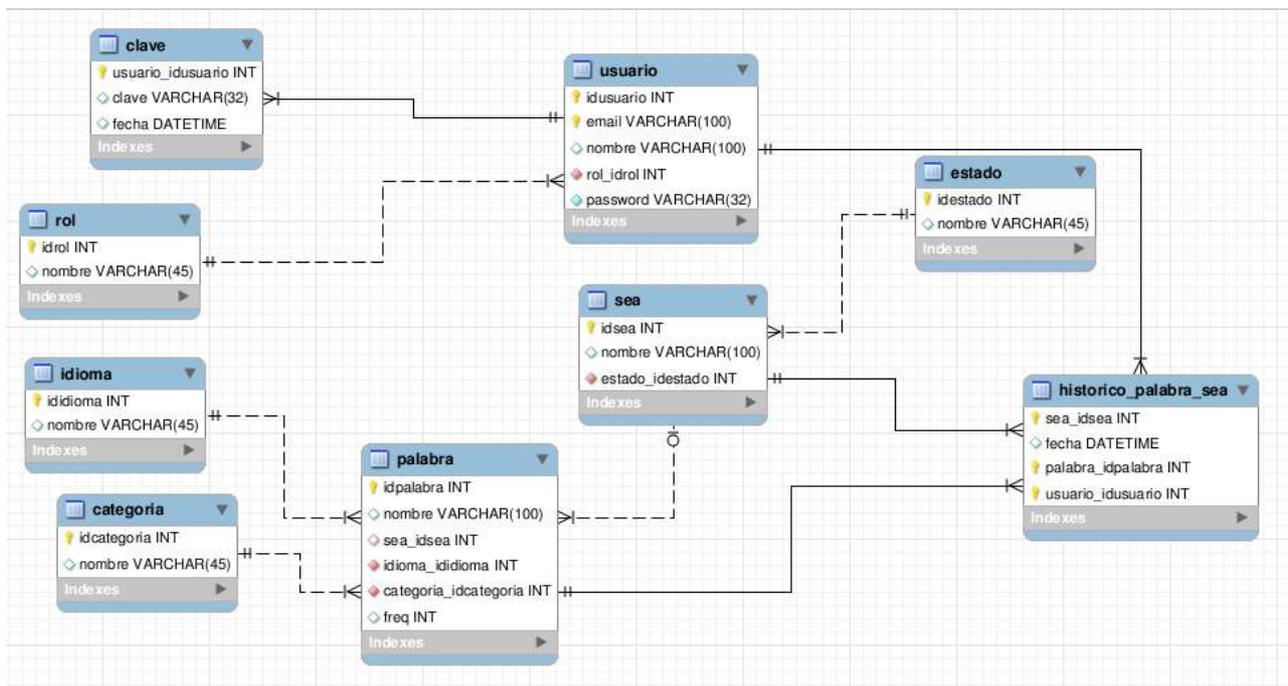


Ilustración 7: Base de datos del entorno colaborativo



Recursos lingüísticos para la Lengua de Signos Española (LSE)

Entrada

Autenticación de usuario

Nombre	<input type="text"/>
Contraseña	<input type="password"/>

Entrar

Illustration 8: Entrada al entorno colaborativo

57 Un usuario tendrá una clave para entrar al entorno colaborativo, y si incorpora una nueva entrada (asigna o modifica un SEA para una palabra) se registrará en la tabla de históricos.

58 Una palabra tiene una categoría, un idioma y un SEA, y en el histórico estarán registrados todas las asignaciones de SEA que ha tenido y qué usuario lo ha introducido. Además, un SEA puede corresponder a varias palabras.

Como se ha comentado, existen dos tipos de usuarios; esto es debido a que los usuarios “*expertos*” tendrán acceso a diferentes menús y además opciones avanzadas en determinadas secciones del entorno, como por ejemplo, acceso a la plantilla para obtener las entradas que se han validado y no están recogidas en el diccionario de Apertium o la opción de poder dar como válida, descartar o modificar la asignación de una escritura SEA para una palabra en español.

En el entorno colaborativo se ha integrado el traductor de Apertium de español–lengua de signos en SEA, en el que se puede traducir automáticamente un texto en español a lengua de signos en SEA.

En este trabajo se han realizado varios cambios dentro de este traductor, entre ellos, se han actualizado los diccionarios monolingües de lengua de signos y el bilingüe, porque varios de los signos se han visto afectados por los cambios que se han producido en la notación SEA (cambios comentados en el apartado 3.1). Y otro de los cambios que se ha realizado dentro del traductor de Apertium ha sido en las reglas de transferencia estructural, ya que al traducir, cuando se trataba de inicio de frase las reglas ponían la primera letra en mayúscula, lo que no era válido en la algunos casos. Es decir, si traducíamos “Susana come arroz” nos devolvía “SUSANA eyle sòamahy:ifehezpb v.ôpob”; en este caso la “S” estaba correctamente en mayúscula, ya que hace referencia al grupo de configuración Q, mientras que si realizamos la traducción “algunos de mis hermanos” nos mostraba que su traducción era “Ye s:ememueb oãmamals” donde la “Y” debería de ser en minúscula. Ahora⁵⁹ ya no transforma la primera letra en mayúscula, la escribe conforme está en los diccionarios.



Illustration 9: Página del traductor español–lengua de signos en SEA de Apertium integrado en el entorno colaborativo

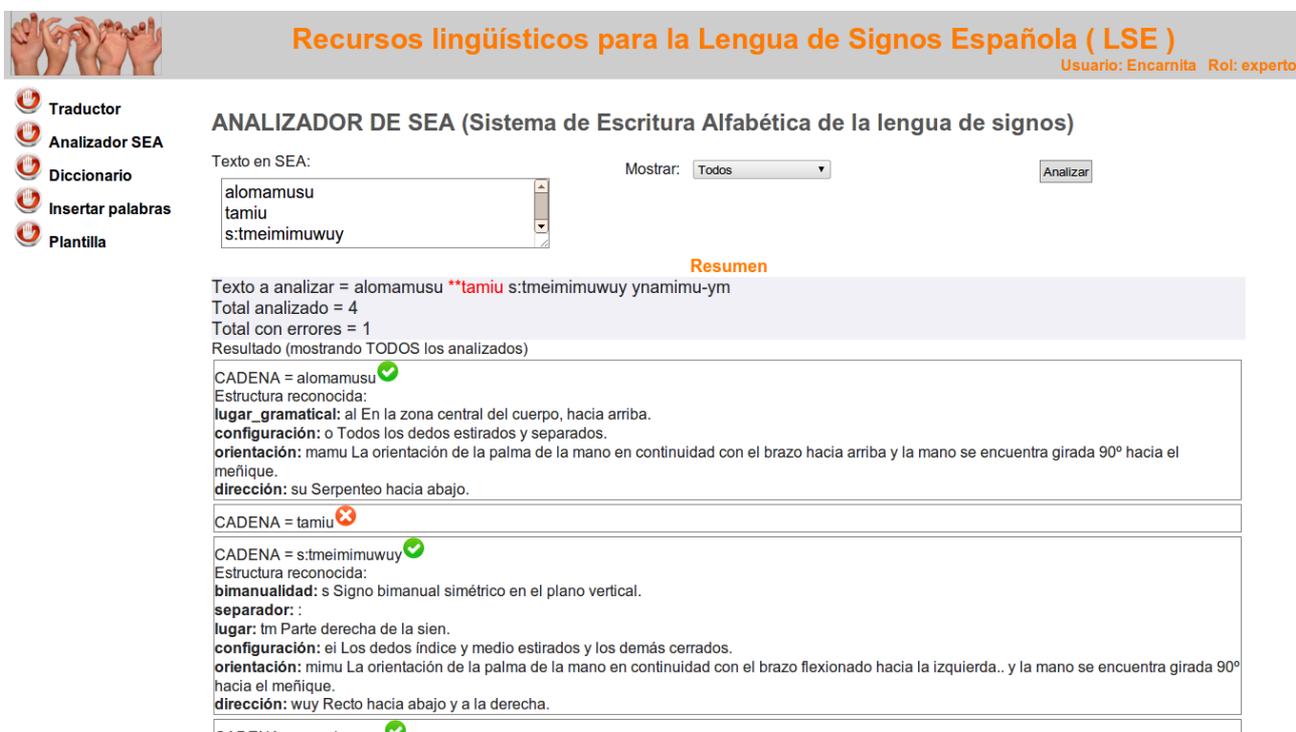
En este trabajo, se han identificado todas las codificaciones de los diferentes grupos de la notación SEA y se ha definido el significado de cada una de ellas. En total se han realizado 1753 codificaciones y transcrito en español sus definiciones; en la Tabla 7 se puede ver un pequeño ejemplo de estas definiciones.

⁵⁹ Se puede descargar la nueva versión en <https://apertium.svn.sourceforge.net/svnroot/apertium/nursery/apertium-es-ssp/>

L	fn	Parte izquierda de la frente.
L	jz	Entrecejo.
L	jmm	Comisura derecha del ojo derecho.
Q	âi	Los dedos pulgar y medio cerrados con contacto y los demás cerrados.
Q	èi	Los dedos índice y medio flexionados y los demás cerrados.
O	ma	La orientación de la palma de la mano en continuidad con el brazo hacia arriba.
D	wa	Recto hacia arriba.
D	soi	Serpenteo hacia el cuerpo y a la izquierda.
D	khu	Las manos se entrecruzan hacia abajo y con movimiento de ida y vuelta.
F	jfah	Abatimiento de la muñeca. Las manos se mueven juntas hacia el mismo lugar.
F	jthu	Temblor con movimiento hacia abajo y con movimiento de ida y vuelta. Las manos se mueven juntas hacia el mismo lugar.

Tabla 7. Ejemplo de distintas codificaciones de los grupos de la notación SEA y su significado

Desde la interfaz del analizador de SEA se pueden introducir una o varias palabras escritas en SEA para identificar las diferentes partes que conforman un signo escrito en esta notación y comprobar que este cumpla los requisitos de formación por los que se rige este sistema de escritura. Se mostrará en primer lugar un resumen en el que se podrá ver si en el texto existe algún signo escrito en SEA que no se cumple alguna de las normas de la notación, aparecerá dicho signo marcado previamente con dos asteriscos (**) y en rojo; el número de signos en SEA que se han detectado como erróneos y el número de signos correctos. Y en segundo el signo o signos con un símbolo que advierte si el signo estaba correcto o no, y en caso de ser correcto, se mostrará la estructura identificada para este, además de una explicación de lo que significa la codificación detectada. Si se analizan varios signos, se da la opción de mostrar todos, o bien solamente aquellos que han proporcionado algún tipo de anomalía o solamente los que han sido validados correctamente.



Recursos lingüísticos para la Lengua de Signos Española (LSE)
 Usuario: Encamita Rol: experto

ANALIZADOR DE SEA (Sistema de Escritura Alfabética de la lengua de signos)

Texto en SEA:

Mostrar:

Resumen

Texto a analizar = alomamusu ****tamiu** s:tmeimimuwuy ynamimu-ym
 Total analizado = 4
 Total con errores = 1
 Resultado (mostrando TODOS los analizados)

CADENA = alomamusu
 Estructura reconocida:
lugar_gramatical: al En la zona central del cuerpo, hacia arriba.
configuración: o Todos los dedos estirados y separados.
orientación: mamu La orientación de la palma de la mano en continuidad con el brazo hacia arriba y la mano se encuentra girada 90° hacia el meñique.
dirección: su Serpenteo hacia abajo.

CADENA = tamiu

CADENA = s:tmeimimuwuy
 Estructura reconocida:
bimanualidad: s Signo bimanual simétrico en el plano vertical.
separador: :
lugar: tm Parte derecha de la sien.
configuración: ei Los dedos índice y medio estirados y los demás cerrados.
orientación: mimu La orientación de la palma de la mano en continuidad con el brazo flexionado hacia la izquierda.. y la mano se encuentra girada 90° hacia el meñique.
dirección: wuy Recto hacia abajo y a la derecha.

CADENA = ynamimuy

Illustration 10: Página de análisis de signos escritos en SEA

En la Ilustración 10 se puede ver la interfaz de diccionario en la que se pueden buscar tanto las palabras en español como en lengua de signos. La búsqueda se realiza sobre un diccionario interno de la aplicación. Este diccionario está codificado en el formato XML de los diccionarios de Apertium y esta formado por las entradas del diccionario bilingüe de español–lengua de signos de Apertium y por las palabras que han sido introducidas por los distintos usuarios del entorno y han sido validadas por algún usuario “*experto*”. Este diccionario contiene cada una de las palabras del diccionario bilingüe de Apertium, junto con su categoría y la estructura analizada de las palabras.

```

<dictionary>
  <e>
    <p>
      <l>
        abeja
        <s n="n"/>
      </l>
      <r>
        âemamuwraiheb
        <s n="n"/>
        <estructura>
          <grupo g='Q'>âe</grupo>
          <grupo g='O'>mamu</grupo>
          <grupo g='F'>wraihe</grupo>
          <grupo g='F'>b</grupo>
        </estructura>
      </r>
    </p>
  </e>

  <e>
    <p>
      <l>abajo
      <s n="adv"/>
      </l>
      <r>
        olemumawu
        <estructura>
          <grupo g='G'>ol</g>
          <grupo g='Q'>e</g>
          <grupo g='O'>muma</g>
          <grupo g='D'>wu</g>
        </estructura>
      </r>
    </p>
  </e>
</dictionary>

```

Tabla 8. Estructura interna del diccionario del entorno colaborativo

Se ha optado por añadir la opción de visualizar la categoría de cada palabra, porque una palabra en español puede ser signada de diferente manera según su categoría gramatical, como ocurre con la palabra en español “bien”, que si es un adverbio su escritura en SEA es *v.ôle*, mientras que si es un nombre su signo sería *seméha:emèzpery*.

En la Ilustración 10 se muestra el resultado de la búsqueda de signos que sean bimanuales y simétricos en el plano vertical, por ello se ha seleccionado la codificación de simetría sea 's' y además, aquellos que su lugar de signado sea la zona central del pecho, por lo que la codificación del lugar del signo es 'y'. Han sido 5 los signos resultantes y en cada uno de ellos se muestra la información en español, en SEA y su categoría gramatical. y además se proporciona la opción de acceder directamente a ver el análisis de esta palabra, junto al análisis léxico aparecerá la palabra en español a la que corresponde.

Recursos lingüísticos para la Lengua de Signos Española (LSE)
 Usuario: Encamita Rol: experto

DICCIONARIO ESPAÑOL-LENGUA DE SIGNOS EN SEA DE APERTIUM

Español SEA
 Bimanualidad (S) Lugar activa-pasiva (H) Lugar (L)
 Configuración (Q) Espacio gramatical (G) Orientación (O)
 Dirección (D) Forma de movimiento (F) No manual (NM)

Español	SEA	Categoría	
quizá	y.éázpugrol s:òmamupo	advervio	Ver análisis
biquini	s:yáémumazy-s:stáézy	nombre	Ver análisis
sensación	y.âemumyzpub s:òmemazjwreih	nombre	Ver análisis
chaqueta	s:pmame-y	nombre	Ver análisis
expresar	s:y.òmèmale	verbo	Ver análisis

[Información SEA](#)

Ilustración 11: Diccionario del entorno colaborativo

El entorno colaborativo sugiere palabras en español cuya traducción a SEA no está aún disponible en los diccionarios. Estas palabras se puede ver en la Ilustración 12 y Ilustración 13, que muestran un listado de palabras en español de los signos utilizados más frecuentemente en diversas páginas web de uso cotidiano por personas pertenecientes a la comunidad sorda y que no se encuentran registradas en los diccionarios del traductor español–lengua de signos de Apertium. Para obtener estas palabras se ha tomado como referencia el sitio web de la Confederación Estatal de Personas Sordas (CNSE), y además, los sitios web de miembros asociados a la CNSE y miembros colaboradores de la CNSE, ya que estos son sitios web de referencia de cualquier persona perteneciente al colectivo sordo. En un principio solamente se habían tenido en cuenta las páginas web de las asociaciones, pero no se obtuvieron muchos resultados porque la mayoría de la información esta signada en vídeos, de los que no se puede obtener texto escrito para saber las palabras más utilizadas, así que se optó por consultar todas las páginas webs referenciadas en el sitio web de la CNSE. Consultar en Anexo D. Una vez se obtuvieron todas las palabras, se obtuvo de ellas su lema y categoría haciendo uso del paquete morfológico del traductor de español–catalán de Apertium, puesto que este es uno de los más completos de esta plataforma, y al tratarse de palabras desconocidas, no se iba a poder obtener la categoría de cada palabra con el morfológico de lengua de signos. Con todas las palabras identificadas con su categoría, utilizando el paquete morfológico del traductor de español–lengua de signos para identificar las palabras en español desconocidas en los diccionarios de Apertium. Al final se obtiene un listado de todas las palabras que no existen en los diccionarios de Apertium ordenadas según la frecuencia de aparición en las distintas webs analizadas.



- Traductor
- Analizador SEA
- Diccionario
- Insertar palabras
- Plantilla

Listado de palabras para insertar

Estado de las palabras: Todos Mostrar

Buscar:

Frecuencia	Palabra	Idioma	Categoría	SEA	Estado	
3196	centro	español	nombre	chemawe	aceptado	Ficha/Editar
3042	salud	español	nombre	t.emab	aceptado	Ficha/Editar
1588	sanitario	español	adjetivo	chemawe	pendiente	Ficha/Editar
1499	ciudadano	español	adjetivo	uama	pendiente	Ficha/Editar
997	enfermedad	español	nombre	ynmawe	pendiente	Ficha/Editar
994	prestación	español	nombre	paemi	pendiente	Ficha/Editar
796	portal	español	nombre			Ficha/Editar
649	dependencia	español	nombre			Ficha/Editar
565	autónomo	español	adjetivo			Ficha/Editar
558	consultorio	español	nombre			Ficha/Editar
529	complejo	español	adjetivo			Ficha/Editar
518	municipio	español	nombre			Ficha/Editar
511	sanidad	español	nombre			Ficha/Editar
508	profesional	español	adjetivo			Ficha/Editar
506	mental	español	adjetivo			Ficha/Editar

Illustration 12: Listado de palabras para insertar (vista del usuario “experto”)



- Traductor
- Analizador SEA
- Diccionario
- Insertar palabras

Listado de palabras para insertar

Buscar:

Frecuencia	Palabra	Idioma	Categoría	
796	portal	español	nombre	Ficha/Editar
649	dependencia	español	nombre	Ficha/Editar
565	autónomo	español	adjetivo	Ficha/Editar
558	consultorio	español	nombre	Ficha/Editar
529	complejo	español	adjetivo	Ficha/Editar
518	municipio	español	nombre	Ficha/Editar
511	sanidad	español	nombre	Ficha/Editar
508	profesional	español	adjetivo	Ficha/Editar
506	mental	español	adjetivo	Ficha/Editar
505	publicación	español	nombre	Ficha/Editar
504	contactar	español	verbo	Ficha/Editar
504	organización	español	nombre	Ficha/Editar
503	estadístico	español	adjetivo	Ficha/Editar
503	web	español	nombre	Ficha/Editar
499	normativo	español	adjetivo	Ficha/Editar
499	provincia	español	nombre	Ficha/Editar

Illustration 13: Listado de palabras para insertar (vista del usuario “anónimo”)

Para el usuario “anónimo” solamente se muestran las 100 primeras; esto es debido a que se quiere intentar que las primeras palabras a las que se les añade la escritura SEA sean de uso frecuente, y que cuando estas sean validadas y se añadan a los diccionarios tanto del entorno colaborativo como del traductor de Apertium, de esta manera se podrá incrementar significativamente su cobertura en palabras más utilizadas, y aportar más beneficio para todos los usuarios del entorno colaborativo.

El signo tendrá dos fases de verificación, una inmediata, donde se comprobará que tiene la estructura de un signo en SEA, para ello se hará uso del analizador y se mostrará al usuario un mensaje diciendo si es correcto o no; y el segundo paso de verificación será por parte de un experto en SEA, quien revisará la validez de la escritura en SEA del signo y posteriormente se procederá a la inclusión de este nuevo signo los distintos diccionarios. La segunda fase es imprescindible, por un lado porque se permite al usuario grabar el SEA escrito para una palabra tanto si esta ha sido verificada afirmativamente por el analizador o no, esto es así porque puede darse el caso de que el signo sea una excepción (como los vistos en el apartado 3.2) y no está aún considerada en el analizador, o porque se haya formado con una nueva codificación debido a la evolución que esta sufriendo la notación SEA y la cual aún no exista en el analizador. Y, por otro lado, porque el signo introducido en SEA puede cumplir la estructura válida de esta notación pero no ser correcta para el signo, por ejemplo, un usuario puede introducir el signo en SEA para la palabra en español “cueva” como “sl:òmemugruy”, que cumple la estructura de un signo escrito en SEA, pero que es incorrecto puesto que la codificación de simetría utilizada (“sl”) indica que el signo tiene una simetría vertical pero con movimiento contrario y no es la que corresponde para el signo “cueva”, esta debería de ser “sc” que es simetría con contacto inicial, por lo que el signo en SEA válido sería “sc:òmemugruy”.

Cada vez que se asigna un SEA para la palabra, esta cambia de estado a “pendiente” o al que le asigne el “experto”, y esta desaparece del listado. Sólo volverá a aparecer en dicho listado si el “experto” ha descartado la escritura en SEA para esa palabra.

También se guardará un archivo histórico para cada palabra en español, la palabra en SEA que se le ha asignado, la fecha y el usuario que ha asignado el SEA, el objetivo de guardar este archivo histórico es para poder obtener información de las asignaciones que se han hecho anteriormente a una palabra antes de darla como válida.

En la Ilustración 14 y la Ilustración 15 se pueden ver las diferentes fichas para introducir el signo en SEA de una palabra en español.

Recursos lingüísticos para la Lengua de Signos Española (LSE)
 Usuario: Encarnita Rol: anónimo

Traductor
 Analizador SEA
 Diccionario
 Insertar palabras

Ficha de la palabra
 Información de la palabra

Frecuencia	796
Palabra	portal
Categoría	nombre
Idioma	español
SEA	<input type="text"/>

Comprobar

Enviar

Información SEA

Illustration 14: Ficha de palabra para introducir escritura en SEA (vista del usuario “anónimo”)

Illustration 15: Ficha de palabra para introducir escritura en SEA (vista del usuario “experto”)

El usuario “experto” podrá validar las escrituras en SEA de las palabras nuevas introducidas, en la Ilustración 16 podemos ver un ejemplo para validar la palabra en español “portal”, en la que se muestra el análisis léxico del signo, su significado y si este cumple las estructuras de formación de SEA. Si es valida, el usuario cambiaría el estado de esta palabra de pendiente a valido, si por el contrario no lo es, se marcaría como descartado.

Illustration 16: Ficha de palabra con SEA insertado y pendiente de confirmar su validación (vista del usuario “experto”)

3.4 Creación de recursos lingüísticos para el par español → LSE de la plataforma Apertium

Uno de los objetivos de este proyecto es la creación de recursos lingüísticos para el traductor de español–lengua de signos que existe en la plataforma Apertium. Este traductor se ha integrado dentro del entorno colaborativo, como hemos visto en el apartado 3.3, en el que se puede usar Apertium para traducir automáticamente un texto en el sentido español hacia la lengua de signos en SEA.

Este traductor hace uso de varios diccionarios, uno monolingüe de español, otro monolingüe de lengua de signos y uno bilingüe, todos ellos con una estructura interna determinada. Para facilitar la inclusión del léxico en estos diccionarios, se ha creado dentro del entorno colaborativo la opción de generar la plantilla para las entradas que se deben introducir en estos diccionarios, teniendo en cuenta las palabras a las que se les ha asignado el signo en SEA y han sido validadas por un experto en la notación SEA. Cuando se introduzcan estas entradas en los diccionarios hay que completarlas con los paradigmas correspondientes.

Recursos lingüísticos para la Lengua de Signos Española (LSE)
Usuario: Encarnita Rol: experto

- Traductor
- Analizador SEA
- Diccionario
- Insertar palabras
- Plantilla

Plantilla para los diccionarios de Apertium

1
Monolingüe español: <e lm="lema-es"><i>salud</i></pardef n="paradigma__n"></e>
Bilingüe español-lengua de signos: <e><p><i>salud<s n="n"/></i></p></t.emab<s n="n"/></r></p></e>
Monolingüe lengua de signos: <e lm="lema-ssp"><i>t.emab</i></pardef n="paradigma__n"></e>

2
Monolingüe español: <e lm="lema-es"><i>centro</i></pardef n="paradigma__n"></e>
Bilingüe español-lengua de signos: <e><p><i>centro<s n="n"/></i></p></chemawe<s n="n"/></r></p></e>
Monolingüe lengua de signos: <e lm="lema-ssp"><i>chemawe</i></pardef n="paradigma__n"></e>



Illustration 17: Plantilla para los diccionarios de Apertium

4 Evaluación

Hemos evaluado⁶⁰ el funcionamiento del analizador léxico de SEA teniendo en cuenta las palabras ya incluidas en los diccionarios de Apertium, ya que estas han sido revisadas por expertos en SEA (las entradas de los diccionario contienen las modificaciones realizadas en SEA⁶¹). Se ha obtenido un listado con todos los signos escritos bajo la notación SEA únicos (una misma escritura en SEA puede pertenecer a varias palabras en español). De un total de 2592 palabras analizadas,⁶² el analizador da como correctas 2582 y con errores 9, es decir, daría como validas en SEA el 99,65%.

De estos datos se podría deducir que el analizador tiene una tasa de error bastante baja, puesto que falla en el 0,35% de los casos.

Al estudiar los casos que no han sido validados⁶³ por el analizador se ha comprobado que:

- En tres de los casos se debe a que las codificaciones de simetría (S) utilizadas no están recogidas en la notación SEA. Actualmente se está comprobando si estos signos están escritos correctamente y si esta codificación ha de incluirse en la notación de SEA, o por el contrario, hay que modificar la escritura de estos signos dentro de los diccionarios de Apertium.
- Y los otros cinco se deben a que la estructura que tienen no cumple ninguna de las establecidas por el sistema SEA. Al igual que con los casos anteriores, se está estudiando si estas estructuras son válidas y se pueden generalizar. Si esto fuera así, se tendría que modificar el analizador para que validara estos nuevos casos, o por contra, escribir correctamente estos signos en los diccionarios de Apertium.

60 Pasos para evaluar el analizador léxico de SEA en el Anexo C

61 Aunque se han detectado algunos signos que no cumplen con los requisitos de la notación SEA. Esto puede deberse a que se han ido realizando cambios en la notación y puede que no se haya reflejado en todos los signos del diccionario.

62 Pasos para la obtención de las palabras a analizar en el Anexo B

63 Los signos y el tipo de error por lo que no se han validado en el analizador están detallados en el Anexo E

5 Conclusiones

Con este trabajo se ha pretendido crear una herramienta que sea útil para generar herramientas y recursos para procesar el sistema de escritura alfabético SEA. Para ello se ha desarrollado el analizador léxico⁶⁴ para dicho sistema de escritura, con el que se pretende identificar las diferentes partes que conforman un signo transcrito a esta notación y comprobar que este cumple los requisitos de formación por los que se rige este sistema de escritura. Este analizador resultará útil a la hora de crear entradas en diccionarios de lengua de signos transcritos en SEA, ya que puede usarse a modo de corrector ortográfico; y crear recursos para el aprendizaje de esta notación.

En trad-lse con la misma licencia que Apertium

Además, en el proyecto se ha creado un portal colaborativo, en el que se podrá utilizar este analizador junto con otros recursos basados en SEA, como es el traductor de español–lengua de signos en SEA, un diccionario bilingüe, y también, crear nuevas entradas para estos.

Teniendo en cuenta las entradas del diccionario bilingüe de lengua de signos de Apertium, se ha obtenido que el analizador léxico de SEA da como válidas el 100% de ellas. De este dato se deduce que el analizador tiene una tasa nula

64 El código del analizador léxico se ha liberado junto con el traductor español–lengua de signos de Apertium en el mismo lugar <https://apertium.svn.sourceforge.net/svnroot/apertium/nursery/apertium-es-ssp/> y bajo la misma licencia.

6 Trabajo futuro

Tomando como base este analizador léxico se puede abrir una línea de investigación para desarrollar las instrucciones necesarias para generar el movimiento de un personaje virtual, haciendo que este realice con éxito los signos, ya que se tienen identificadas cada una de las partes que conforman un signo, y de esta manera, la comunidad sorda podría obtener por el canal visual-gestual toda la información. La salida del analizador se podría convertir a los formatos SiGML o al lenguaje SWML mencionados en el apartado 2.1.1 utilizados para que un personaje virtual animado pueda signar la entrada recibida en SEA o en español si se hiciera uso del traductor automático español-lengua de signos de Apertium.

7 Bibliografía

- A. Oviedo (1997). "¿Lengua de señas", "¿lenguaje de signos?", "lenguaje gestual", "lengua manual"? Razones para escoger una denominación". El bilingüismo de los sordos (2), Bogotá: INSOR, págs. 7-11
- Rodríguez Ortiz, Isabel de los Reyes. 2005. Comunicar a través del silencio: las posibilidades de la lengua de signos española. Sevilla: Publicaciones de la Universidad, pág. 31
- Pérez Cobacho, Juan et al. 2001. Introducción a la lengua de signos española: una experiencia piloto. Murcia: Publicaciones Diego Marin, pág. 14
- Herrero et al., 1999. Fonología y escritura de la lengua de signos española. ELUA, 13, 1999, págs. 89-116
- Massó Sanabre, Guillem, 2007. Traducció automàtica aplicada a les llengües de signes : estat de la qüestió. <http://hdl.handle.net/2072/5002>
- Prillwitz, S, 1991. Der lange Weg zur Zweisprachigkeit Gehörloser im deutschen Sprachraum. In Gebärdensprache in Forschung und Praxis, eds. S. Prillwitz, Siegmund, and T. Vollhaber, 17-34. Hamburg: Signum
- Bangham et al., 2000. Virtual Signing: Capture, Animation, Storage and Transmission – an Overview of the ViSiCAST Project. En IEE Seminar on Speech and Language processing for disabled and elderly people. London. England.
- Elliott et al., 2000. The development of language processing support for the Visicast Project. 4th International ACM SIG-CAPH Conference on Assitive Technologies. New York
- Elliott et al., 2007. Linguistic modelling and lanuage- processing technologies for avatar-based sign language presentation. Universal Access in the Information Society, 6(4):375–391.
- Kennaway, Glauert y Zwitserlood, 2007. Providing signed content on the internet by synthesized animation. ACM Transactions on Computer Human Interaction, 14(15):1–29.
- Sutton, Valerie y Gleves, Richard, 1995. SignWriter Computer Program Versión 4.3, La Jolla, The Deaf Action Committee For SignWriting
- Costa, A. C. R. & Dimuro, G. P., 2001. Supporting Deaf Sign Languages on the Web. The SignWriting Journal, v.1, n.0, July 2011

Papadogiorgaki, M., Grammalidis, N., Makris, L., Sarris, N. y Strintzis, M. G., 2004. VSigns - A Virtual Sign Synthesis Web Tool. En: Proceedings of COST 276 Works-hop on Information and Knowledge Management for Integrated Media Communication, Tesalónica, Grecia.

Báez Montero, Inmaculada C. & Carmen Cabeza Pereiro, 2002: 669. Métodos de trabajo de campo en lingüística de las lenguas de señas. Alberto Bernabé y otros (eds.): Presente y futuro de la lingüística en España. La sociedad de lingüística, 30 años después, Madrid, SEL, 665-670.

Ángel Herrero Blanco, 2003. Escritura alfabética de la lengua de signos española: Once lecciones. Universidad de Alicante. Servicio de publicaciones.

DILSE - Tesoro de la lengua de signos española <http://www.fundacioncse.org/tesorolse>

Mikel L. Forcada, Francis M. Tyers, Gema Ramírez-Sánchez, 2009. The Apertium machine translation platform: five years on, in Juan Antonio Pérez-Ortiz, Felipe Sánchez-Martínez, Francis M. Tyers, ed., Proceedings of the First International Workshop on Free/Open-Source Rule-Based Machine Translation (Alacant, 02-03.11.2009) , págs. 3-10

8 Anexo

A. Código del analizador del sistema SEA

El analizador esta implementado en un fichero con el formato XML de los diccionarios de Apertium.

Se han definido dos tipos de entrada, una para reconocer los signos que se han identificado como excepciones y otra para el resto.

En los paradigmas identificados con el nombre “X_general” se han definido una serie de entradas que especifican las reglas de estructuración válidas de la escritura SEA.

En los paradigmas identificados con el nombre “Simbolos_X” se han definido las codificaciones de cada uno de los grupos de la clasificación de las características de los signos en SEA, donde X es el nombre del grupo.

Y en los paradigmas identificados como “X_interno” se encuentra definido por separado cada carácter del abecedario del sistema de escritura SEA, donde X hace referencia a dicho carácter.

```
<dictionary> <!-- -*- nxml -*- -->
<alphabet>abcdefghijklmnopqrstuvwxyzABCDEFGHIJKLMNÑOPQRSTUVWXYZáááááâââââéééééîîîîîóóóóó
úúúúú1234567890°</alphabet>

<pardefs>
  <pardef n="a_interno">
    <e><i>a</i></e>
  </pardef>
  <pardef n="b_interno">
    <e><i>b</i></e>
  </pardef>
  <pardef n="c_interno">
```

```
<e><i>c</i></e>
</pardef>
<pardef n="d_interno">
  <e><i>d</i></e>
</pardef>
<pardef n="e_interno">
  <e><i>e</i></e>
</pardef>
<pardef n="f_interno">
  <e><i>f</i></e>
</pardef>
<pardef n="g_interno">
  <e><i>g</i></e>
</pardef>
<pardef n="h_interno">
  <e><i>h</i></e>
</pardef>
<pardef n="i_interno">
  <e><i>i</i></e>
</pardef>
<pardef n="j_interno">
  <e><i>j</i></e>
</pardef>
<pardef n="k_interno">
  <e><i>k</i></e>
</pardef>
<pardef n="l_interno">
  <e><i>l</i></e>
</pardef>
<pardef n="m_interno">
  <e><i>m</i></e>
</pardef>
<pardef n="n_interno">
  <e><i>n</i></e>
</pardef>
<pardef n="ñ_interno">
  <e><i>ñ</i></e>
</pardef>
<pardef n="o_interno">
  <e><i>o</i></e>
</pardef>
<pardef n="p_interno">
  <e><i>p</i></e>
</pardef>
<pardef n="q_interno">
  <e><i>q</i></e>
</pardef>
<pardef n="r_interno">
  <e><i>r</i></e>
</pardef>
<pardef n="s_interno">
  <e><i>s</i></e>
</pardef>
<pardef n="t_interno">
  <e><i>t</i></e>
</pardef>
<pardef n="u_interno">
  <e><i>u</i></e>
</pardef>
<pardef n="v_interno">
  <e><i>v</i></e>
</pardef>
<pardef n="w_interno">
  <e><i>w</i></e>
</pardef>
<pardef n="x_interno">
  <e><i>x</i></e>
</pardef>
```

```

<pardef n="y_interno">
  <e><i>y</i></e>
</pardef>
<pardef n="z_interno">
  <e><i>z</i></e>
</pardef>

<pardef n="agrave_interno">
  <e><i>à</i></e>
</pardef>
<pardef n="aaccento_interno">
  <e><i>á</i></e>
</pardef>
<pardef n="adieresis_interno">
  <e><i>ä</i></e>
</pardef>
<pardef n="atilde_interno">
  <e><i>ã</i></e>
</pardef>
<pardef n="acircunflejo_interno">
  <e><i>â</i></e>
</pardef>

<pardef n="egrave_interno">
  <e><i>è</i></e>
</pardef>
<pardef n="eaccento_interno">
  <e><i>é</i></e>
</pardef>
<pardef n="edieresis_interno">
  <e><i>ë</i></e>
</pardef>
<pardef n="etilde_interno">
  <e><i>ẽ</i></e>
</pardef>
<pardef n="ecircunflejo_interno">
  <e><i>ê</i></e>
</pardef>

<pardef n="igrave_interno">
  <e><i>ì</i></e>
</pardef>
<pardef n="iacento_interno">
  <e><i>í</i></e>
</pardef>
<pardef n="idieresis_interno">
  <e><i>ï</i></e>
</pardef>
<pardef n="itilde_interno">
  <e><i>ĩ</i></e>
</pardef>
<pardef n="icircunflejo_interno">
  <e><i>î</i></e>
</pardef>

<pardef n="ograve_interno">
  <e><i>ò</i></e>
</pardef>
<pardef n="oaccento_interno">
  <e><i>ó</i></e>
</pardef>
<pardef n="odieresis_interno">
  <e><i>ö</i></e>
</pardef>
<pardef n="otilde_interno">
  <e><i>õ</i></e>
</pardef>
<pardef n="ocircunflejo_interno">

```

```

    <e><i>ô</i></e>
</pardef>

<pardef n="ugrave_interno">
    <e><i>ù</i></e>
</pardef>
<pardef n="uacento_interno">
    <e><i>ú</i></e>
</pardef>
<pardef n="udieresis_interno">
    <e><i>ü</i></e>
</pardef>
<pardef n="utilde_interno">
    <e><i>ÿ</i></e>
</pardef>
<pardef n="ucircunflejo_interno">
    <e><i>û</i></e>
</pardef>

<pardef n="sepbisilabica_interno">
    <e><i>-</i></e>
</pardef>
<pardef n="sepbimanual_interno">
    <e><i>:</i></e>
</pardef>
<pardef n="punto_interno">
    <e><i>.</i></e>
</pardef>

<pardef n="CODIFICACIONESPACIO_interno">
    <e><par n="a_interno"/></e>
    <e><par n="e_interno"/></e>
    <e><par n="i_interno"/></e>
    <e><par n="o_interno"/></e>
    <e><par n="u_interno"/></e>
    <e><par n="y_interno"/></e>
</pardef>

<pardef n="SEPBIMANUAL_interno">
    <e><par n="sepbimanual_interno"/></e>
</pardef>

<pardef n="SEPBISILABICA_interno">
    <e><par n="sepbisilabica_interno"/></e>
</pardef>

<pardef n="Simbolos_S">
    <e><par n="s_interno"/></e>
    <e><par n="s_interno"/><par n="l_interno"/></e>
    <e><par n="s_interno"/><par n="c_interno"/></e>
    <e><par n="s_interno"/><par n="g_interno"/></e>
    <e><par n="s_interno"/><par n="k_interno"/></e>
    <e><par n="s_interno"/><par n="r_interno"/></e>
    <e><par n="s_interno"/><par n="v_interno"/></e>
    <e><par n="s_interno"/><par n="z_interno"/></e>
    <e><par n="s_interno"/><par n="v_interno"/><par n="c_interno"/></e>
    <e><par n="s_interno"/><par n="g_interno"/><par n="l_interno"/></e>
    <e><par n="s_interno"/><par n="r_interno"/><par n="l_interno"/></e>
    <e><par n="s_interno"/><par n="c_interno"/><par n="l_interno"/></e>
    <e><par n="s_interno"/><par n="v_interno"/><par n="l_interno"/></e>
    <e><par n="s_interno"/><par n="v_interno"/><par n="r_interno"/></e>
    <e><par n="s_interno"/><par n="h_interno"/><par n="a_interno"/></e>
    <e><par n="s_interno"/><par n="h_interno"/><par n="e_interno"/></e>
    <e><par n="s_interno"/><par n="h_interno"/><par n="o_interno"/></e>
    <e><par n="s_interno"/><par n="h_interno"/><par n="u_interno"/></e>
    <e><par n="s_interno"/><par n="c_interno"/><par n="a_interno"/></e>
    <e><par n="s_interno"/><par n="c_interno"/><par n="e_interno"/></e>
    <e><par n="s_interno"/><par n="c_interno"/><par n="o_interno"/></e>

```



```

<e><par n="adieresis_interno"/><par n="e_interno"/><par n="i_interno"/></e>
<e><par n="a_interno"/><par n="egrave_interno"/><par n="igrave_interno"/></e>
<e><par n="e_interno"/></e>
<e><par n="e_interno"/><par n="i_interno"/></e>
<e><par n="e_interno"/><par n="u_interno"/></e>
<e><par n="e_interno"/><par n="acircunflejo_interno"/></e>
<e><par n="e_interno"/><par n="adieresis_interno"/></e>
<e><par n="e_interno"/><par n="atilde_interno"/></e>
<e><par n="egrave_interno"/><par n="igrave_interno"/></e>
<e><par n="eacento_interno"/><par n="a_interno"/></e>
<e><par n="eacento_interno"/></e>
<e><par n="eacento_interno"/><par n="aacento_interno"/></e>
<e><par n="eacento_interno"/><par n="atilde_interno"/></e>
<e><par n="eacento_interno"/><par n="iacento_interno"/></e>
<e><par n="i_interno"/></e>
<e><par n="i_interno"/><par n="acircunflejo_interno"/></e>
<e><par n="i_interno"/><par n="adieresis_interno"/></e>
<e><par n="i_interno"/><par n="atilde_interno"/></e>
<e><par n="iacento_interno"/><par n="aacento_interno"/></e>
<e><par n="o_interno"/></e>
<e><par n="o_interno"/><par n="adieresis_interno"/></e>
<e><par n="o_interno"/><par n="atilde_interno"/></e>
<e><par n="ograve_interno"/></e>
<e><par n="ograve_interno"/><par n="a_interno"/></e>
<e><par n="ograve_interno"/><par n="atilde_interno"/></e>
<e><par n="oacento_interno"/></e>
<e><par n="oacento_interno"/><par n="a_interno"/></e>
<e><par n="ocircunflejo_interno"/></e>
<e><par n="ocircunflejo_interno"/><par n="atilde_interno"/></e>
<e><par n="ocircunflejo_interno"/><par n="e_interno"/><par n="u_interno"/></e>
<e><par n="ocircunflejo_interno"/><par n="a_interno"/></e>
<e><par n="otilde_interno"/></e>
<e><par n="otilde_interno"/><par n="a_interno"/></e>
<e><par n="otilde_interno"/><par n="a_interno"/><par n="e_interno"/><par
n="u_interno"/></e>
<e><par n="otilde_interno"/><par n="aacento_interno"/></e>
<e><par n="otilde_interno"/><par n="atilde_interno"/></e>
<e><par n="odieresis_interno"/></e>
<e><par n="u_interno"/></e>
</pardef>

<pardef n="QMAYUS_interno">
  <e><p><l/><r> (configuración=</r></p><re>[A-ZÑÁÉÍÓÚÛÇÀÈÌÒÙÀÈÌÒÙ]
+</re><p><l/><r></r></p></e>
</pardef>

<pardef n="QNUM_interno">
  <e><re>[0-9]+</re></e>
</pardef>

<pardef n="ORDINAL_interno">
  <e><i>º</i></e>
</pardef>

<pardef n="QNUMORDINAL_interno">
  <e><par n="QNUM_interno"/><par n="ORDINAL_interno"/></e>
</pardef>

<pardef n="Osentido_interno">
  <e><par n="a_interno"/></e>
  <e><par n="agrave_interno"/></e>
  <e><par n="aacento_interno"/></e>
  <e><par n="e_interno"/></e>
  <e><par n="egrave_interno"/></e>
  <e><par n="eacento_interno"/></e>
  <e><par n="i_interno"/></e>
  <e><par n="igrave_interno"/></e>
  <e><par n="iacento_interno"/></e>

```



```

n="SEPBIMANUAL_interno"/><p><l/><r></r></p></e>
</pardef>

<pardef n="SEPBISILABICA">
  <e><p><l/><r>(bisilábico-compuesto=</r></p><par
n="SEPBISILABICA_interno"/><p><l/><r></r></p></e>
</pardef>

<pardef n="S">
  <e><p><l/><r>(bimanualidad=</r></p><par n="Sb_general"/><p><l/><r></r></p></e>
</pardef>

<pardef n="SP">
  <e><p><l/><r>(sp=</r></p><par n="SP_interno"/><p><l/><r></r></p></e>
</pardef>

<pardef n="SPP">
  <e><p><l/><r>(spp=</r></p><par n="SPP_interno"/><p><l/><r></r></p></e>
</pardef>

<pardef n="SPPP">
  <e><p><l/><r>(sppp=</r></p><par n="SPPP_interno"/><p><l/><r></r></p></e>
</pardef>

<pardef n="H">
  <e><p><l/><r>(situación_h=</r></p><par n="Simbolos_H"/><p><l/><r></r></p></e>
</pardef>

<pardef n="G">
  <e><p><l/><r>(lugar_gramatical=</r></p><par
n="Simbolos_G"/><p><l/><r></r></p></e>
</pardef>

<pardef n="L">
  <e><p><l/><r>(lugar=</r></p><par n="Simbolos_L"/><p><l/><r></r></p></e>
</pardef>

<pardef n="C">
  <e><p><l/><r>(contacto=</r></p><par n="Simbolos_C"/><p><l/><r></r></p></e>
</pardef>

<pardef n="Q">
  <e><p><l/><r>(configuración=</r></p><par n="Simbolos_Q"/><p><l/><r></r></p></e>
  <e><p><l/><r>(configuración=</r></p><par n="Simbolos_Q"/><p><l/><r></r></p><par
n="QMAYUS_interno"/></e>
  <e><par n="QMAYUS_interno"/></e>
  <e><p><l/><r>(configuración=</r></p><par
n="QNUM_interno"/><p><l/><r></r></p></e>
  <e><p><l/><r>(configuración=</r></p><par
n="QNUM_interno"/><p><l/><r></r></p><p><l/><r>(configuración=</r></p><par
n="Simbolos_Q"/><p><l/><r></r></p></e>
  <e><p><l/><r>(configuración=</r></p><par
n="QNUMORDINAL_interno"/><p><l/><r></r></p></e>
  <e><p><l/><r>(configuración=</r></p><par
n="QNUMORDINAL_interno"/><p><l/><r></r></p><p><l/><r>(configuración=</r></p><par
n="Simbolos_Q"/><p><l/><r></r></p></e>
</pardef>

<pardef n="O">
  <e><p><l/><r>(orientación=</r></p><par n="O_interno"/><p><l/><r></r></p></e>
</pardef>
<pardef n="REP">
  <e><p><l/><r>(repetición=</r></p><par n="Simbolos_REP"/><p><l/><r></r></p></e>
</pardef>

<pardef n="NM">
  <e><p><l/><r>(no_manual=</r></p><par n="Simbolos_NM"/><p><l/><r></r></p></e>
</pardef>

```

```

<pardef n="D">
  <e><p><l/><r>(dirección=</r></p><par n="Simbolos_D"/><p><l/><r></r></p></e>
</pardef>

<pardef n="DD">
  <e><par n="D"/><par n="D"/></e>
</pardef>

<pardef n="DJ">
  <e><p><l/><r>(dirección=</r></p><par n="DJ_interno"/><p><l/><r></r></p></e>
</pardef>

<pardef n="DJDJ">
  <e><par n="DJ"/><par n="DJ"/></e>
</pardef>

<pardef n="F">
  <e><p><l/><r>(forma=</r></p><par n="Simbolos_F"/><p><l/><r></r></p></e>
</pardef>

<pardef n="FF">
  <e><par n="F"/><par n="F"/></e>
</pardef>

<pardef n="FFF">
  <e><par n="F"/><par n="FF"/></e>
</pardef>

<pardef n="FFFF">
  <e><par n="F"/><par n="FFF"/></e>
</pardef>

<pardef n="FJ">
  <e><p><l/><r>(forma=</r></p><par n="FJ_interno"/><p><l/><r></r></p></e>
</pardef>

<pardef n="FFJ">
  <e><par n="FJ"/><par n="FJ"/></e>
</pardef>

<pardef n="FFFJ">
  <e><par n="FJ"/><par n="FFJ"/></e>
</pardef>

<pardef n="FFFFJ">
  <e><par n="FJ"/><par n="FFFJ"/></e>
</pardef>

<pardef n="FD">
  <e><par n="F"/><par n="D"/></e>
</pardef>

<pardef n="FDD">
  <e><par n="F"/><par n="DD"/></e>
</pardef>

<pardef n="FFD">
  <e><par n="FF"/><par n="D"/></e>
</pardef>

<pardef n="FFDF">
  <e><par n="FF"/><par n="D"/><par n="F"/></e>
</pardef>

<pardef n="FFDD">
  <e><par n="FF"/><par n="DD"/></e>
</pardef>

```

```

<pardef n="FFFD">
  <e><par n="FFF"/><par n="D"/></e>
</pardef>
<pardef n="FDF">
  <e><par n="F"/><par n="D"/><par n="F"/></e>
</pardef>

<pardef n="FDFJ">
  <e><par n="FJ"/><par n="DJ"/><par n="FJ"/></e>
</pardef>

<pardef n="FDJ">
  <e><par n="FJ"/><par n="DJ"/></e>
</pardef>

<pardef n="FDJDJ">
  <e><par n="FJ"/><par n="DJDJ"/></e>
</pardef>

<pardef n="FFDJ">
  <e><par n="FFJ"/><par n="DJ"/></e>
</pardef>

<pardef n="FFDFJ">
  <e><par n="FFJ"/><par n="DJ"/><par n="FJ"/></e>
</pardef>

<pardef n="FFDJDJ">
  <e><par n="FFJ"/><par n="DJDJ"/></e>
</pardef>

<pardef n="FFFDJ">
  <e><par n="FFFJ"/><par n="DJ"/></e>
</pardef>

<pardef n="DF">
  <e><par n="D"/><par n="F"/></e>
</pardef>

<pardef n="DFF">
  <e><par n="D"/><par n="FF"/></e>
</pardef>

<pardef n="DFFF">
  <e><par n="D"/><par n="FFF"/></e>
</pardef>

<pardef n="DDF">
  <e><par n="DD"/><par n="F"/></e>
</pardef>

<pardef n="DDFF">
  <e><par n="DD"/><par n="FF"/></e>
</pardef>

<pardef n="DDFFF">
  <e><par n="DD"/><par n="FFF"/></e>
</pardef>

<pardef n="DJF">
  <e><par n="DJ"/><par n="FJ"/></e>
</pardef>

<pardef n="DJFF">
  <e><par n="DJ"/><par n="FFJ"/></e>
</pardef>

```

```

<pardef n="DJFFF">
  <e><par n="DJ"/><par n="FFFJ"/></e>
</pardef>

<pardef n="DJDJF">
  <e><par n="DJDJ"/><par n="FJ"/></e>
</pardef>

<pardef n="DJDJFF">
  <e><par n="DJDJ"/><par n="FFJ"/></e>
</pardef>

<pardef n="DJDJFFF">
  <e><par n="DJDJ"/><par n="FFFJ"/></e>
</pardef>

<pardef n="D_general">
  <e><par n="D"/></e>
  <e><par n="DD"/></e>
  <e><par n="DF"/></e>
  <e><par n="DFD"/></e>
  <e><par n="DFDD"/></e>
  <e><par n="DFDDF"/></e>
  <e><par n="DFDDFF"/></e>
</pardef>

<pardef n="DJ_general">
  <e><par n="DJ"/></e>
  <e><par n="DJDJ"/></e>
  <e><par n="DJF"/></e>
  <e><par n="DJFF"/></e>
  <e><par n="DJFFF"/></e>
  <e><par n="DJDJF"/></e>
  <e><par n="DJDJFF"/></e>
  <e><par n="DJDJFFF"/></e>
</pardef>

<pardef n="F_general">
  <e><par n="F"/></e>
  <e><par n="FF"/></e>
  <e><par n="FFF"/></e>
  <e><par n="FFFF"/></e>
  <e><par n="FD"/></e>
  <e><par n="FDD"/></e>
  <e><par n="FFD"/></e>
  <e><par n="FFDD"/></e>
  <e><par n="FFFD"/></e>
  <e><par n="FFDF"/></e>
  <e><par n="FDF"/></e>
</pardef>

<pardef n="FJ_general">
  <e><par n="FJ"/></e>
  <e><par n="FFJ"/></e>
  <e><par n="FFFJ"/></e>
  <e><par n="FFFFJ"/></e>
  <e><par n="FDJ"/></e>
  <e><par n="FDJDJ"/></e>
  <e><par n="FFDJ"/></e>
  <e><par n="FFDJDJ"/></e>
  <e><par n="FFFJDJ"/></e>
  <e><par n="FFDFJ"/></e>
  <e><par n="FDFJ"/></e>
</pardef>

<pardef n="O_general">
  <e><par n="O"/></e>

```

```

    <e><par n="O"/><par n="D_general"/></e>
    <e><par n="O"/><par n="F_general"/></e>
</pardef>

<pardef n="OJ_general">
    <e><par n="O"/></e>
    <e><par n="O"/><par n="DJ_general"/></e>
    <e><par n="O"/><par n="FJ_general"/></e>
</pardef>

<pardef n="Q_general">
    <e><par n="Q"/></e>
    <e><par n="Q"/><par n="O_general"/></e>
    <e><par n="Q"/><par n="D_general"/></e>
    <e><par n="Q"/><par n="F_general"/></e>
</pardef>

<pardef n="QBM_general">
    <e><par n="Q"/></e>
    <e><par n="Q"/><par n="OJ_general"/></e>
    <e><par n="Q"/><par n="DJ_general"/></e>
    <e><par n="Q"/><par n="FJ_general"/></e>
</pardef>

<pardef n="C_general">
    <e><par n="C"/><par n="Q_general"/></e>
</pardef>
<pardef n="CBM_general">
    <e><par n="C"/><par n="QBM_general"/></e>
</pardef>

<pardef n="L_general">
    <e><par n="L"/><par n="Q_general"/></e>
    <e><par n="L"/><par n="C_general"/></e>
</pardef>

<pardef n="LB_general">
    <e><par n="L"/></e>
    <e><par n="L"/><par n="C"/></e>
    <e><par n="L_general"/></e>
    <e><par n="L"/><par n="O_general"/></e>
    <e><par n="L"/><par n="D_general"/></e>
    <e><par n="L"/><par n="F_general"/></e>
</pardef>

<pardef n="LBM_general">
    <e><par n="L"/><par n="QBM_general"/></e>
    <e><par n="L"/><par n="CBM_general"/></e>
</pardef>

<pardef n="G_general">
    <e><par n="G"/><par n="Q_general"/></e>
</pardef>

<pardef n="GB_general">
    <e><par n="G"/></e>
    <e><par n="G"/><par n="Q_general"/></e>
</pardef>

<pardef n="GBMS_general">
    <e><par n="G"/><par n="QBM_general"/></e>
</pardef>

<pardef n="SEPS">
    <e><par n="SEPBIMANUAL"/><par n="LBM_general"/></e>
    <e><par n="SEPBIMANUAL"/><par n="QBM_general"/></e>
    <e><par n="SEPBIMANUAL"/><par n="GBMS_general"/></e>
</pardef>

```

```

<pardef n="SEPG">
  <e><par n="SEPBIMANUAL"/><par n="GBMS_general"/></e>
  <e><par n="SEPBIMANUAL"/><par n="QBM_general"/></e>
</pardef>

<pardef n="SEPQ">
  <e><par n="SEPBIMANUAL"/><par n="QBM_general"/></e>
</pardef>

<pardef n="S_general">
  <e><par n="S"/><par n="SEPS"/></e>
</pardef>

<pardef n="SG_general">
  <e><par n="S"/><par n="SEPG"/></e>
</pardef>

<pardef n="GBM_general">
  <e><par n="G"/><par n="SG_general"/></e>
</pardef>

<pardef n="HS_general">
  <e><par n="H"/><par n="SEPQ"/></e>
  <e><par n="H"/><par n="SEPS"/></e>
</pardef>

<pardef n="OBMS_general">
  <e><par n="O"/><par n="HS_general"/></e>
  <e><par n="O"/><par n="SEPQ"/></e>
</pardef>

<pardef n="QBMS_general">
  <e><par n="Q"/><par n="OBMS_general"/></e>
</pardef>

<pardef n="SP_general">
  <e><par n="SP"/><par n="QBMS_general"/></e>
  <e><par n="SP"/><par n="OBMS_general"/></e>
</pardef>

<pardef n="SPP_general">
  <e><par n="SPP"/><par n="SP_general"/></e>
</pardef>

<pardef n="SPPP_general">
  <e><par n="SPPP"/><par n="HS_general"/></e>
</pardef>

<pardef n="HQ_general">
  <e><par n="H"/><par n="SEPQ"/></e>
</pardef>

<pardef n="OBMQ_general">
  <e><par n="O"/><par n="HQ_general"/></e>
  <e><par n="O"/><par n="SEPQ"/></e>
</pardef>

<pardef n="QBMG_general">
  <e><par n="Q"/><par n="OBMQ_general"/></e>
  <e><par n="Q"/><par n="HQ_general"/></e>
</pardef>

<pardef n="SPQ_general">
  <e><par n="SP"/><par n="QBMG_general"/></e>
</pardef>

<pardef n="GBMQ_general">

```

```

    <e><par n="G"/><par n="SPQ_general"/></e>
</pardef>

<pardef n="SA_general">
    <e><par n="S"/><par n="QBMG_general"/></e>
    <e><par n="S"/><par n="SEPQ"/></e>
</pardef>

<pardef n="LA_general">
    <e><par n="L"/><par n="SA_general"/></e>
    <e><par n="L"/><par n="C"/><par n="SA_general"/></e>
</pardef>
<pardef n="SIGUIENTESILABA">
    <e><par n="REP"/></e>
    <e><par n="NM"/></e>
    <e><par n="D_general"/></e>
    <e><par n="F_general"/></e>
    <e><par n="O_general"/></e>
    <e><par n="Q_general"/></e>
    <e><par n="GB_general"/></e>
    <e><par n="LB_general"/></e>
    <e><par n="S_general"/></e>
    <e><par n="GBM_general"/></e>
    <e><par n="GBMQ_general"/></e>
    <e><par n="SPPP_general"/></e>
    <e><par n="SPP_general"/></e>
    <e><par n="SP_general"/></e>
    <e><par n="QBMG_general"/></e>
    <e><par n="LA_general"/></e>
</pardef>

<pardef n="SIGUIENTESILABA2">
    <e><par n="REP"/></e>
    <e><par n="NM"/></e>
    <e><par n="DJ_general"/></e>
    <e><par n="FJ_general"/></e>
    <e><par n="O_general"/></e>
    <e><par n="Q_general"/></e>
    <e><par n="GB_general"/></e>
    <e><par n="LB_general"/></e>
    <e><par n="S_general"/></e>
    <e><par n="GBM_general"/></e>
    <e><par n="GBMQ_general"/></e>
    <e><par n="SPPP_general"/></e>
    <e><par n="SPP_general"/></e>
    <e><par n="SP_general"/></e>
    <e><par n="QBMG_general"/></e>
    <e><par n="LA_general"/></e>
</pardef>

<pardef n="SEPBB">
    <e><par n="SEPBISILABICA"/><par n="SIGUIENTESILABA"/></e>
</pardef>
<pardef n="SEPBB2">
    <e><par n="SEPBISILABICA"/><par n="SIGUIENTESILABA2"/></e>
</pardef>

<pardef n="EXCEPCIONES">
    <e><p><l/><r>(bimanualidad=</r></p><par n="s_interno"/><p><l/><r></r></p>
    <p><l/><r>(separador=</r></p><par
n="sepbimanual_interno"/><p><l/><r></r></p>
    <p><l/><r>(lugar=</r></p><par
n="sepbisilabica_interno"/><par n="f_interno"/><p><l/><r></r></p>
    <p><l/><r>(configuración=</r></p><par
n="atilde_interno"/><p><l/><r></r></p>
    <p><l/><r>(orientación=</r></p><par
n="a_interno"/><p><l/><r></r></p>
    n="n_interno"/><par
    n="ograve_interno"/><par
    n="m_interno"/><par

```

```

        <p><l/><r> (forma=</r></p><par n="l_interno"/><par
n="e_interno"/><p><l/><r></r></p>
        </e>
        <e><p><l/><r> (bimanualidad=</r></p><par n="s_interno"/><p><l/><r></r></p>
        <p><l/><r> (separador=</r></p><par
n="sepbimanual_interno"/><p><l/><r></r></p>
        <p><l/><r> (lugar=</r></p><par n="b_interno"/><par
n="sepbisilabica_interno"/><par n="g_interno"/><par
n="punto_interno"/><p><l/><r></r></p>
        <p><l/><r> (configuración=</r></p><par n="ograve_interno"/><par
n="a_interno"/><p><l/><r></r></p>
        <p><l/><r> (orientación=</r></p><par n="m_interno"/><par n="u_interno"/><par
n="m_interno"/><par n="a_interno"/><p><l/><r></r></p>
        <p><l/><r> (dirección=</r></p><par n="w_interno"/><par
n="a_interno"/><p><l/><r></r></p>
        </e>
</pardef>

<pardef n="SILABAM">
    <e><par n="Q_general"/></e>
    <e><par n="G_general"/></e>
    <e><par n="L_general"/></e>
</pardef>

<pardef n="SILABAB">
    <e><par n="S_general"/></e>
    <e><par n="GBM_general"/></e>
    <e><par n="SP_general"/></e>
    <e><par n="SPP_general"/></e>
    <e><par n="SPPP_general"/></e>
    <e><par n="GBMQ_general"/></e>
    <e><par n="QBMG_general"/></e>
    <e><par n="LA_general"/></e>
</pardef>

<pardef n="SIGNO">
    <e><par n="SILABAM"/></e>
    <e><par n="SILABAM"/><par n="SEPB"/></e>
    <e><par n="SILABAM"/><par n="SEPB"/><par n="SEPB"/></e>
    <e><par n="SILABAM"/><par n="SEPB"/><par n="SEPB"/><par n="SEPB"/></e>

    <e><par n="SILABAB"/></e>
    <e><par n="SILABAB"/><par n="SEPB2"/></e>
    <e><par n="SILABAB"/><par n="SEPB2"/><par n="SEPB2"/></e>
    <e><par n="SILABAB"/><par n="SEPB2"/><par n="SEPB2"/><par n="SEPB2"/></e>
</pardef>

</pardefs>

<section id="main" type="standard">
    <e><par n="SIGNO"/></e>
    <e><par n="EXCEPCIONES"/></e>
</section>

</dictionary>

```

B. Pasos para la obtención de las palabras a analizar

Expandir el diccionario bilingüe de español–lengua de signos de apertium

```
$ lt-expand apertium-es-ssp.es-ssp.dix > es-ssp-dix.exp
```

Una vez expandido el diccionario, obtenemos el listado de palabras en lengua de signos únicas y ordenadas alfabéticamente.

```
$ cat es-ssp-dix.exp | gawk '{gsub(";<",";"); gsub("<","\t<"); gsub(";", "\t"); gsub(/<[a-z]*>/,""); gsub(/>/,""); gsub("\t\t","\t"); gsub("#",""); if(!index($0,"REGEXP")){print $0;}}'| gawk -F '\t' '{print $2;}' | gawk '{gsub(/ /, "\n"); print $0;}' | sort | uniq
```

C. Pasos para probar/evaluar el analizador

Compilar con tl-comp el diccionario

```
$ lt-comp lr analizador_sea.dix analizador_sea.bin
```

Procesar el signo con lt-proc y la opción -c para que no se aplique ninguna gestión de mayúsculas y minúsculas.

```
$ echo "sòamémaca:olda" | lt-proc -c analizador_sea.bin  
^sòamémaca:olda/(sp=s) (configuración=òa) (orientación=méma) (situación_h=ca)  
(separador=:) (configuración=o) (forma=lda)$
```

Si el signo que se está analizando es incorrecto, este aparecerá marcado con un asterisco.

```
$ echo "chmawe" | lt-proc -c analizador_sea.bin  
^chmawe/*chmawe$
```

El analizador no analiza solamente un signo, puede analizar varios. Para comprobarlo, este es el comando para analizar el listado de signos escritos en SEA del diccionario bilingüe de Apertium.

```
$ lt-expand apertium-es-ssp.es-ssp.dix | gawk '{gsub(";<",";"); gsub("<","\t<"); gsub(";", "\t"); gsub(/<[a-z]*>/,""); gsub(/>/,""); gsub("\t\t","\t"); gsub("#",""); if(!index($0,"REGEXP")){print $0;}}'| gawk -F '\t' '{print $2;}' | gawk '{gsub(/ /, "\n"); print $0;}' | sort | uniq | lt-proc -c analizador_sea.bin
```

D. Páginas web para obtener las palabras de uso frecuente

Los sitios web analizados para la obtención de palabras de uso frecuente:

- Confederación Estatal de Personas Sordas (CNSE) www.cnse.es
- Federación Andaluza de asociaciones de Personas Sordas (FAAS): www.faas.es
- Federación de Personas Sordas del Principado de Asturias (FESOPRAS): www.fesopras.org
- Federación de Personas Sordas de Cantabria (FESCAN): www.fescan.es
- Federación de Personas Sordas de Castilla la Mancha (FESORMANCHA): www.fesormancha.org
- Federación de asociaciones de Personas Sordas de Castilla y León(FAPSCL): www.fapscl.org
- Federació de Persones Sordes de Catalunya (FESOCA): www.fesoca.org
- Federación de Personas Sordas de la Comunidad de Madrid (FESORCAM): www.fesorcam.org
- Federación de Personas Sordas de la Comunidad Valenciana (FESORD CV): www.fesord.org
- Federación Extremeña de Asociaciones de Personas Sordas (FEXAS): www.fexas.es
- Federación de Asociaciones de Personas Sordas de Galicia (FAXPG): www.faxpg.es
- Federación de Asociaciones de Personas Sordas de Islas Canarias (FASICAN): www.fasican.org
- Federación de Asociaciones de Personas Sordas de la Región de Murcia (FESORMU): www.fesormu.org
- Federación de Asociaciones de Personas Sordas (Euskalgorrak): www.euskal-gorrak.org
- Agrupación de Personas Sordas de Zaragoza y Aragón: www.asza.net
- Fundación Andaluza Accesibilidad y Personas Sordas: www.fundacionaccesible.org
- Fundación Flor Lara: www.fundacionflorlara.com
- Asociación de Familias de personas sordas de Guipúzcoa (Arans-gi): www.aransgi.org
- Associació de pares de nens sords de Catalunya (APANSCE): www.apansce.org
- Colegio Gaudem: www.gaudem.es
- EUD: www.eud.eu
- WFD: www.wfdeaf.org
- Ministerio de Trabajo e Inmigración: www.tt.mtin.es
- Ministerio de Sanidad y Política Social: www.msc.es
- Ministerio de Educación: www.educacion.es
- Ministerio de Igualdad: www.migualdad.es
- Ministerio de Industria, Turismo y comercio: www.mityc.es
- Ministerio de Presidencia: www.mpr.es
- Ministerio de Cultura: www.mcu.es
- Centro de Referencia Estatal de Autonomía Personal y Ayudas Técnicas (CEAPAT): www.ceapat.org
- Centro Español de Subtitulado y Audio descripción (CESYA): www.cesya.es
- Imsero: www.seg-social.es
- Instituto de la Mujer: www.migualdad.es

- Real Patronato sobre Discapacidad: www.rpd.es
- Consejo de la Juventud de España: www.cje.org
- Instituto de la Juventud: www.injuve.migualdad.es
- Real Academia Española: www.rae.es
- Instituto Cervantes: www.cervantes.es
- Plataforma de ONG de Acción Social: www.plataformaongs.org
- Comité Español de Representantes de Personas con discapacidad (CERMI): www.cermi.es
- Fundación ONCE: www.fundaciononce.es
- ONCE: www.once.es
- Fundación ONCE para la atención de personas con sordoceguera (FOAPS): www.foaps.es
- Federación española de deportes para sordos: www.feds.es
- Asociación de Sordociegos de España (ASOCIDE): www.asocide.org
- Federación Española de Intérpretes de Lengua de signos y guías-intérpretes (FILSE): www.filse.org
- Biblioteca Virtual Miguel de Cervantes: www.cervantesvirtual.com
- UNIDIS: Centro de Atención a Universitarios con Discapacidad. www.portal.uned.es
- Vodafone España: www.vodafone.es
- Telefónica España: www.telefonica.es
- Asisa: www.asisa.es
- Fundación Luis Vives: www.fundacionluisvives.org
- Fundación Vodafone España: www.fundacion.vodafone.es
- Ibermutuamur: www.ibermutuamur.es
- Obra Social CajaMadrid: www.obrasocialcajamadird.es
- CCOO: www.ccoo.es
- UGT: www.ugt.es , www.ugt.es/ejb
- www.webvisual.tv
- www.diariosigno.com
- www.zeinu.tv

E. Signos no validados por el analizador léxico

Signos que no han sido validados por el analizador léxico y su razón.

- Utilizan un tipo de simetría no vertical que no está codificada en el sistema SEA
 - boquiabierto → scal:ómemufaho donde su codificación de simetría es “scal”
 - polémico → sherl:ómèmuwraheb donde su codificación de simetría es “sherl”
 - torero → shy:òmèfoho donde su codificación de simetría es “shy”
- No cumplen ninguna de las estructuras de formación de SEA
 - En los siguientes casos la codificación del grupo L se encuentra delante del símbolo de bimanualidad, no existe ninguna estructura con esta característica.
 - arriar → semacoah:òmizu tiene la estructura SQOHL:QOD
 - sòmémucaah:emèzhoy cañón tiene la estructura SQOHL:QOD
 - Tiene lugar gramatical para la mano pasiva y lugar corporal para la mano activa, no hay ninguna estructura establecida donde puedan existir estos dos lugares.
 - guitarra → ilsđamymahy:st.óazpub tiene la estructura GSOMH:LCODF
 - En los siguientes casos no existe la estructura S de simetría en la primera parte del signo (antes del símbolo de bimanualidad). Esta estructura se ha escrito en la segunda parte, dando a entender que la mano considerada normalmente como pasiva pasa a actuar como mano activa.
 - jarabe → aëmiha:sđámegro aëmiwa-v. tiene la estructura QOH:SQOF QOD-LC
 - derecha → òmemyhi:sòamacb tiene la estructura QMH:SQOFF
 - a mano derecha → òmemyhi:sòamecb tiene la estructura QOH:SQOFF