

# PRINCIPIOS LINGÜÍSTICOS APLICADOS A LA SÍNTESIS AUTOMÁTICA DE LA LENGUA DE SIGNOS ESPAÑOLA

FERNANDO LÓPEZ-COLINO, JORDI PORTA Y JOSÉ COLÁS  
Universidad Autónoma de Madrid. Grupo HCTLab  
fj.lopez@uam.es  
jordi.porta@uam.es  
jose.colas@uam.es

## Resumen

Este trabajo presenta la base lingüística utilizada en un sintetizador de lengua de signos española (LSE). Los aspectos fundamentales tratados son la fonología de la LSE y una aproximación para describir mensajes signados. En relación a la fonología se describen los parámetros fonológicos utilizados, el modelo fonológico en el que se basa este trabajo y la aproximación para almacenar esta información. Por último se presentan los resultados obtenidos en pruebas de evaluación con nativos de LSE.

PALABRAS CLAVE: LSE, fonología, síntesis automática.

## Abstract

This work presents the linguistic basis of a Spanish Sign Language (LSE) synthesizer. This paper focuses on the phonologic foundations of LSE and an approach to describe signed messages. Related to phonology, we will discuss the phonologic parameters used during the synthesis, the phonologic model which has been used in this work and the proposed approach to store all this information. Finally, we present the results of the synthesizer's evaluations conducted by LSE natives.

KEY WORDS: LSE, phonology, automatic synthesis.

## 1. Introducción

En este trabajo se plantea una aproximación lingüística a la lengua de signos española (LSE) algo distinta a la tradicional. Esta aproximación se ha centrado en la aplicación y adaptación de una gramática descriptiva de la LSE. Esto es debido a que la definición debe hacerse para que un ordenador sea capaz de generar una animación de un personaje virtual (avatar). Esta definición debe o bien, ser absolutamente completa para que el ordenador tenga toda la información necesaria para la síntesis, o bien, estar complementada con un conjunto de reglas de inferencia que permitan al sistema extraer la información no definida. En este proceso de formalización, sobre todo en el aspecto fonológico del trabajo, se realiza una reflexión sobre aspectos que

otras formalizaciones o notaciones delegan al conocimiento, experiencia e intuición de un signante humano que interprete dicha definición. Esto no quiere decir que dichas notaciones o aproximaciones no sean correctas, sólo pone de manifiesto que cuando se idearon no se enfocaron a la síntesis por ordenadores.

El estudio lingüístico de las lenguas de signos (LS) tiene una trayectoria, a nivel internacional, de unos cincuenta años (especialmente la lengua de signos americana, ASL). Si nos centramos en el estudio de la LSE, este periodo se reduce hasta los veinte años. Este periodo de tiempo resulta muy pequeño si lo comparamos con el periodo de tiempo que los lingüistas han dedicado a las lenguas orales. No sólo el menor tiempo dedicado a las LS es un factor crucial para el estadio de desarrollo de la lingüística de las LS, también la falta de una representación escrita unificada y común, junto con los múltiples canales que intervienen en la comunicación signada aumentan la complejidad del estudio. Esto se agrava si consideramos que la LSE no ha sido oficial en España hasta el año 2007 (Gobierno de España, 2007), en el que se reconocen los derechos de las personas que desean comunicarse en esta lengua.

El procesado automático de contenidos para su traducción y síntesis automáticas comienza a nivel internacional hace algo más de diez años. Por fortuna, en este aspecto, la aplicación de las nuevas tecnologías a la traducción y síntesis de contenidos en LSE no tardó en aparecer (Ibáñez, 2006; López *et alii* 2006; Baldassarri y Royo Santas, 2007; San Segundo *et alii*, 2008a, 2008b). Aunque la aproximación de estos sistemas o bien se basa en soluciones desarrolladas para otras LS o bien no contemplan una base fonológica, centrándose en la definición manual de los signos.

En este trabajo se tratan dos aspectos del mensaje signado, ambos relevantes para la síntesis automática: la fonología de los signos y la descripción del contenido lingüístico. Finalmente, y de manera más breve, se presenta una aplicación informática que permite registrar las descripciones fonéticas de los signos y los resultados obtenidos de una evaluación por parte de signantes de LSE, de los contenidos sintéticos que se generan con este sistema.

### 1.1. Sistema de traducción de español a LSE

Antes de entrar en los detalles del sintetizador, es conveniente describir el contexto del mismo: Un sistema de traducción desde español a LSE cuenta con tres módulos independientes (ver ilustración 1). Estos tres módulos son: (i) Un reconocedor de voz, que convierte la señal acústica en texto, generando una transcripción. (ii) Un traductor de español a LSE, que convierte la

transcripción antes generada, o bien un texto, a una representación intermedia de la LSE. Esta notación utilizada para la representación intermedia recibe el nombre de HLSML (López-Colino y Colás, 2009a). (iii) Por último, el sintetizador de LSE, que toma la descripción del mensaje y genera una animación representando dicho mensaje en LSE.

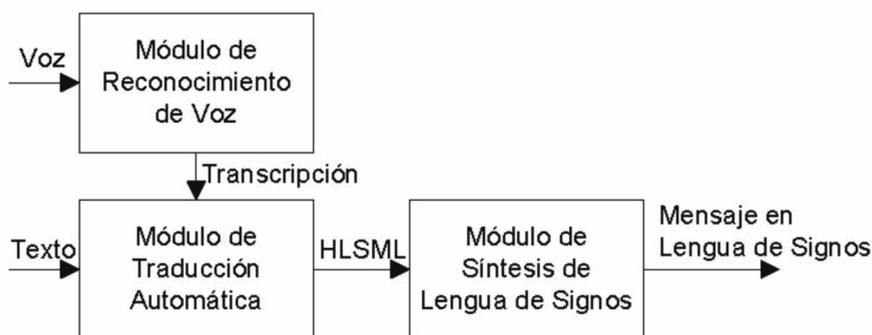


Ilustración 1. Diagrama de bloques del sistema de traducción automática de español a LSE (Ma TSyLSE)

En este trabajo nos vamos a centrar en el mensaje HLSML, que permite describir de manera escrita un mensaje en LSE, y en la base fonológica que utiliza el sintetizador.

## 2. Fonología del sintetizador de LSE

### 2.1. Parámetros fonológicos

La realización de un signo implica la articulación de sus elementos más básicos, sus parámetros fonológicos. Los parámetros que definen un signo difieren según el autor. Desde el trabajo de Fernández Villalibre (1851) en el que los signos son considerados unidades indivisibles, hasta el trabajo de Muñoz Baell (1999) quien defiende la existencia de siete parámetros formativos, han sido varios los autores que han afrontado el estudio fonológico de las LS.

El trabajo de Stokoe (1960) resulta pionero como estudio lingüístico de las LS. En su trabajo considera que el signo está compuesto por tres parámetros distintos *tab* (lugar de signado), *dez* (forma y orientación de la mano) y *sig* (cambio en *tab* o *dez*). Posteriormente, Battison (1974) describió la independencia de la orientación de la mano y la forma que adopta. Es esta formalización de cuatro parámetros (Configuración, Orientación, Lugar y

Movimiento) la que Herrero Blanco (2003) plantea para la descripción de la componente manual del signo. En cambio, Rodríguez González (1990) considera que el lugar de signado debe describirse como composición de dos parámetros independientes: la posición de la mano en un plano frontal (Localización), utilizando referencias anatómicas para su definición, y la distancia horizontal de la mano al cuerpo (Plano). Asimismo, considera que el Componente no manual (CNM) debe ser incluido en la definición fonológica del signo. También Herrero Blanco (2009: 67) considera el CNM parte del código lingüístico junto con los componentes articulatorios manuales del signo (los cuatro antes mencionados), pero no parte de la definición fonológica del signo. Por último, Muñoz Baell (1999) considera que el Punto de Contacto debe ser considerado como un parámetro más en la definición de un signo.

Estos estudios plantean distintas aproximaciones y teorías fonológicas, con distinto número de parámetros requeridos para la descripción de un signo, pero todas ellas concuerdan que la realización de los distintos parámetros es independiente. Recordemos, que este trabajo versa sobre la aplicación de las definiciones lingüísticas anteriores para lograr que un ordenador, de manera automática, genere contenidos signados con la suficiente calidad. Por lo tanto, la definición fonológica que utilizamos se regirá por los principios expuestos en la introducción de este trabajo: toda la información necesaria debe estar disponible o ser inferible utilizando reglas formalizables. Por esta razón hemos utilizado el planteamiento de Muñoz Baell (1999) definiendo siete parámetros distintos, aunque hemos considerado la necesidad de plantear alguna modificación que incluye similitudes con la aproximación expuesta por Herrero Blanco (2009).

### 2.1.1. *Configuración y Orientación de la mano*

Los parámetros *Configuración* y *Orientación*, salvo en el trabajo de Stokoe que los agrupaba en un único parámetro, son definidos de igual manera por todas las teorías fonológicas de las LS. La configuración de la mano viene definida por la flexión de las articulaciones de cada uno de los dedos. Para la LSE, Herrero define cuarenta y dos configuraciones fonológicas, treinta configuraciones del diccionario dactilológico y cinco configuraciones numerales. La orientación de la mano indica, para cada mano, el par que define la dirección de los dedos extendidos y la dirección de la palma. Herrero indica que la dirección que puede tomar la mano contempla seis direcciones ortogonales en el espacio y para cada una de estas direcciones se describen cuatro rotaciones de la palma de la mano; describiendo un total de veinticuatro posibles

orientaciones fonológicas. Las variaciones alofónicas que pueden sufrir estos dos fonemas, deben ser contempladas en el momento en el que nos enfrentamos a la síntesis de LSE. Las variaciones fonéticas, expresivas y gramaticales hacen que el número de unidades que se deben contemplar aumente. Actualmente, el sintetizador contempla ochenta alófonos de la Configuración y las variaciones alofónicas de la Orientación son doscientas ocho unidades si contemplamos variaciones de cuarenta y cinco grados en todas las direcciones del espacio y en las rotaciones de la muñeca.

### 2.1.2. Posición y contacto de la mano

Cuando nos centramos en el modo de definir la posición de las manos, nos encontramos con distintas soluciones: De una parte está la aproximación de Herrero Blanco (2009) al definir el *Lugar* como un único parámetro para definir la posición de las manos; y por otro lado, Rodríguez González (1992) que considera dos parámetros independientes, la *Localización* y el *Plano*. En ambos casos es relevante la existencia o no de contacto de la mano definida con la Localización. En este caso, las aproximaciones son distintas, Herrero define una modificación en el parámetro Lugar para señalar la existencia del contacto. Muñoz Baell (1999) considera la existencia de un nuevo parámetro, el *Punto de Contacto*, que define la parte de la mano activa que realiza el contacto; su aparición en la definición indica la existencia de este contacto. En la aproximación que se plantea en este trabajo, consideramos que ambas aproximaciones se pueden compatibilizar y extender: En este sistema se ha elegido como base el sistema fonológico de Muñoz Baell que define la posición de la mano utilizando los parámetros de Localización y Plano. Esta aproximación presenta una ventaja técnica, ya que reduce el número de unidades que se deben definir a priori, dado que la combinación de una Localización y un Plano para definir la posición de la mano, puede realizarse de manera automática. Esta aproximación presenta un problema a considerar: una definición que requiera contacto pero el Plano esté definido de tal manera que dicho contacto sea imposible. Si nos paramos a analizar un signo en el que la mano dominante entre en contacto con la mano pasiva u otra parte del cuerpo, ¿cuál es la distancia al cuerpo que permite este contacto? La respuesta es: la que sea necesaria en función de los valores de Configuración y Orientación de la mano dominante, la posición espacial de la Localización y la parte de la mano que realice ese contacto (Punto de Contacto). Por lo tanto, en el caso de un signo con contacto, el valor del Plano no puede ser definido utilizando el conjunto discreto de valores, sino que viene dado por la Localización. Es

decir, la posición espacial que antes se generaba utilizando los valores de Localización y Plano, en el caso de existir un contacto, la misma Localización se utiliza para definir la distancia horizontal de la mano al cuerpo. En este punto se plantea también la siguiente reflexión: la mano dominante no va a situarse en la misma posición si el contacto lo realizan distintas partes de la mano. Este aspecto pone de manifiesto que el valor del fonema del Punto de Contacto, parámetro definido por Muñoz Baell, incluye cierta información relevante para definir la posición final de la mano y, por tanto, en la realización del signo.

*(1) Supongamos la definición de un signo monomaneal y monosilabo. La configuración de la mano es con todos los dedos extendidos con máxima apertura entre los dedos. La orientación de la mano es vertical, los dedos apuntan hacia el techo, y la palma orientada hacia el signante. La localización se define en la punta de la nariz y el plano es frente al cuerpo, pero no existe contacto entre la mano dominante y el cuerpo. El gesto resultante es diferente si situamos frente a la nariz el extremo del meñique o el extremo del pulgar (fonema definido para el parámetro Punto de Contacto). En un caso tapamos el lado derecho de la cara y en el otro es el izquierdo, luego este valor debe ser definido o inferido para la síntesis.*

En el ejemplo anterior se pone de manifiesto la necesidad de indicar el punto más relevante de la mano cuando se va a colocar en la posición definida por la Localización y el Plano, independientemente de que haya contacto. Por esta razón, la solución fonológica que se utiliza en el sintetizador incluye la definición del Punto de Contacto en todos los signos.

### 2.1.3. Movimiento

El parámetro fonológico *Movimiento* se ha incluido desde las primeras aproximaciones a la definición fonológica de un signo. La diferencia se encuentra en el alcance de éste según los autores. Todos los autores coinciden en que el desplazamiento de las manos existe, la divergencia se presenta en si define un parámetro fonológico por sí mismo. Nosotros consideramos la existencia de dos tipos de desplazamiento (Corina 1996), un movimiento derivado de la necesidad de transición entre dos Lugares y un *Movimiento* como parámetro fonológico. Éste requiere la definición de la trayectoria a seguir por las manos (lineal, curva, línea ondulada, angular, etc.) siendo relevantes la velocidad y aceleración del movimiento.

En lo que difieren los distintos trabajos es en considerar la variación del valor de la forma de la mano y orientación de la muñeca (Configuración y Orientación) o bien como un parámetro distinto (Movimiento fonológico) o una variación de los valores que toman estos parámetros fonológicos. En el modelo de Stokoe

(1960), el parámetro *sig* define tanto una variación de posición como de la Configuración y Orientación de la mano. Esta aproximación es también utilizada por Herrero Blanco al definir que el parámetro *Movimiento* se compone tanto de movimiento de desplazamiento como de movimientos formales (los que modifican la Configuración y Orientación). Por el contrario, Muñoz Baell considera que en el *Movimiento*, sólo intervienen los movimientos de desplazamiento, considerando que la variación del resto de parámetros es un fenómeno local y debe ser gestionado dentro de dicho parámetro. Esta segunda aproximación es conceptualmente más sencilla (el cambio de la forma de la mano dentro de un signo corresponde a una transición entre dos fonemas de dicho parámetro fonológico). El inconveniente de esta aproximación está en que un signo con movimiento lineal puede ser definido tanto mediante una variación en los fonemas descritos para la Localización y el Plano, como mediante la inclusión de un fonema en el *Movimiento*. Esta diferencia es relevante desde el punto de vista lingüístico, ya que la definición de un signo no es única; pero para un sintetizador no es importante, ya que el gesto que realizará el avatar será equivalente. Por lo tanto, se ha optado por esta segunda aproximación para el sintetizador de LSE.

#### 2.1.4. *Componente no manual del signo*

Los estudios de LS se centraron inicialmente en la comprensión de la parte manual de los signos, como puede desprenderse de los estudios iniciales de Stokoe y Battison. Pero pronto quedó de manifiesto que el CNM forma parte del discurso signado. Según los autores Muñoz Baell y Herrero Blanco, el CNM está dentro del código lingüístico, pero difieren en si pertenece a los componentes articulatorios del signo.

A veces se ha considerado que los elementos kinésicos no manuales tienen en la LSE un carácter fonológico [...]. Esta confusión deriva de la idea de que los fonemas o parámetros formativos del signo son simultáneos. Hoy en día todos los investigadores están de acuerdo en que son secuencias, y en que el valor fundamental de los elementos no manuales es, en cambio, prosódico, actuando sobre todo el signo. Pero [...] algunos componentes no manuales pueden tener un valor discriminativo, y actuar sobre un segmento del signo [...] (Herrero Blanco, 2009: 68).

También queda de manifiesto la importancia del CNM en el diccionario normativo de la LSE (Fundación CNSE, 2008). En algunos de los signos recogidos se observan realizaciones del CNM. Las signaciones del diccionario corresponden a signos aislados, en los que la prosodia de la frase no está patente. Herrero reconoce en el CNM una función diacrítica que permite diferenciar dos signos con el mismo componente manual (Herrero Blanco, 2009:

81). Por lo tanto, contemplar el CNM en la definición de los signos sintetizables es necesario; si bien, la aproximación a esta definición no sigue los mismos criterios que para los parámetros manuales.

El CNM está compuesto por distintos canales independientes que pueden definirse ya sea tanto para funciones léxicas, morfológicas y sintácticas. Por lo tanto, la síntesis de LSE contempla la posibilidad de movimiento en las cejas, los párpados, las mejillas, la dirección de la mirada, la lengua, la boca<sup>1</sup>, los giros de cabeza, movimientos de los hombros<sup>2</sup> y movimientos del tronco; todos ellos gestionados de manera independiente.

### 2.1.5. Multiplicidad de canales durante el signado

Una vez descritos los parámetros fonológicos que se contemplan en el sintetizador, es necesario contemplar en su gestión la existencia de múltiples canales durante el signado. Estos múltiples canales hacen referencia a cada una de las manos, y el CNM (con la división en subcanales descrito en el apartado anterior). Los distintos canales pueden ser independientes (ej. un signo bimanual simétrico) o bien relacionados (ej. la mano dominante debe contactar con la pasiva, luego la posición de la primera está supeditada a la segunda). Esta relación debe contemplarse tanto en la definición de un signo como en construcciones sintácticas complejas.

## 2.2 Desde el modelo fonológico de *Movement-Hold* a la sílaba, evolución temporal del signo

En la sección anterior se han presentado los parámetros fonológicos que se deben contemplar en la definición del signo. En este apartado nos centraremos en cómo describir la evolución temporal de dichos elementos para la realización del signo.

El modelo fonológico de Liddell y Johnson (1989) divide el signo en segmentos, cada uno de estos segmentos está compuesto por dos elementos: *Hold*, que define una postura en la que todos los elementos articulatorios están estáticos, y *Movement*, que define un periodo de tiempo en el que algu-

1 La definición de la boca aúna el valor de la apertura de la boca, la posición de los labios, mostrar los dientes como posiciones estáticas que puede adoptar. Asimismo, las fonaciones y oralizaciones, corresponden a movimientos de la boca con cierta reminiscencia visémica.

2 Sobre los elementos aquí descritos, se debe permitir asignar valores independientes a cada una de las cejas, mejillas, párpados y hombros, al igual que ocurre con un signante humano.

no de esos elementos articulatorios está en transición. Esto implica que los fonemas considerados, Configuración, Orientación, Punto de Contacto y Lugar, deben estar sincronizados en cada *Hold*. La descripción de un mensaje signado requerirá el uso en paralelo de segmentos, uno por cada una de las manos y otro por cada elemento del CNM. Sandler (1989) y Sandler y Lillo-Martin (2006) consideran una extensión del modelo anterior en el que la configuración de la mano debe ser definida independientemente de los segmentos de desplazamiento de la mano. Este modelo recibe el nombre de *Hand Tier*, y simplifica la descripción de un signo eliminando estructuras redundantes. Esta extensión es seguida tanto por Corina (1996) como Hulst y Mills (1996), que consideran que la Orientación es un fonema que puede sufrir variaciones independientes durante la realización de un signo.

La idea de segmentos planteada por Liddell y Jonhson evoluciona hacia la inclusión del concepto de sílaba aplicada a la LS desde la LO (como ya lo fue el concepto de fonema). Así pues, una primera definición de sílaba sería:

Criteria for counting the number of syllables in surface forms

- a) The number of sequential phonological movements in a string equals the number of syllables in that string. When several shorter dynamic elements co-occur with a single dynamic element of longer duration, the single movement defines the syllable.
- b) If a structure is a well-formed syllable as an independent word, it must be counted as a syllable word-internally. (Brentari, 1996)

que posteriormente, Herrero Blanco concretaría de la siguiente manera:

Entendemos por sílaba cada una de las articulaciones autónomas (es decir, no ligadas por movimientos fonológicos independientes) que se producen en la realización de un signo. (Herrero Blanco, 2009)

Esta definición conlleva la implicación de que una sílaba en un signo requiere de una variación de uno o más de los parámetros que lo conforman. Pero, ¿es realmente necesario el concepto de sílaba en un sistema de síntesis? El concepto de sílaba es lingüísticamente relevante, pero puede ser omitido para la aproximación del sintetizador. La aproximación que hemos decidido tomar es la de maximizar la flexibilidad en las definiciones de los signos. Si consideramos una aproximación básica, podemos asegurar que todos los parámetros pueden estar en dos estadios diferentes: uno estático, en el que permanecen inalterables definido por un fonema, y uno dinámico, caracterizado por una transición entre dos estados estáticos. Además, si maximizamos la flexibilidad en la aproximación, debemos permitir que estos segmentos estáticos puedan ser simultáneos entre dos parámetros cualesquiera o no.

Esta aproximación implica considerar diferentes los movimientos internos y externos. Los primeros se plantean como una transición entre dos fonemas (pasar de una configuración de mano abierta a mano cerrada). Siendo los segundos una variación en la posición de la mano. Ahora bien, esta definición plantea un conflicto entre una variación del Lugar de articulación<sup>3</sup> (equiparable a un movimiento interno) y un fonema Movimiento (un movimiento externo) en el que se define una trayectoria, velocidad y aceleración. Las transiciones entre dos Lugares se realizan linealmente (salvo que implique colisión con otra parte del cuerpo), un Movimiento puede definir una trayectoria lineal, o de mayor grado de complejidad. Visualmente, si se trata de dos ejecuciones similares debemos considerar que se trata del mismo signo, la diferencia está en la aproximación fonológica que se utilice. Dado que este es nuestro objetivo, el sintetizador mantiene la flexibilidad necesaria como para admitir ambos tipos de definición, que llevarán a la misma visualización y, por tanto, al mismo signo percibido (idealmente).

Por último, no se debe olvidar que se está definiendo la evolución temporal del signo, ya sea utilizando la aproximación de segmentos de Liddell y Johnson o las distintas aproximaciones silábicas. En todos estos casos se debe contemplar la definición de la duración de las distintas partes del signo. El sistema de síntesis debe ser capaz de describir o inferir la duración de los bloques estáticos y la duración del movimiento.

### 2.3. Representación y almacenamiento de la descripción de los signos

En los dos apartados anteriores (2.1 y 2.2) se han tratado los distintos aspectos relacionados con la fonología, el código lingüístico y la estructura temporal que son relevantes para abordar la síntesis automática de la LSE. Por lo tanto es necesario el uso de una notación o sistema de descripción que permita:

- representar los fonemas (y alófonos) de manera independiente
- describir el aspecto temporal de la ejecución de un signo
- recoger, para una misma glosa, distintas variaciones dialectales
- facilitar el acceso individual a los fonemas para realizar operaciones fonéticas (flexión e introversión).

#### 2.3.1. La base de datos relacional

El uso de una base de datos relacional es la mejor aproximación para representar con gran flexibilidad toda la información antes descrita (López *et*

---

3 Recordemos que en este trabajo, el *Lugar* de articulación se considera como composición de los *parámetros Localización y Plano*.

alii, 2006): parámetros fonológicos y evolución temporal contemplando la simultaneidad-secuencialidad de la LS. Además permite incluir fácilmente múltiples definiciones para distintos elementos permitiendo variaciones en el signado (variaciones geográficas o dialectales).

La descripción de los elementos antes descritos se recoge en la base de datos en cuatro niveles distintos de complejidad (ver ilustración 2) que se detallan a continuación:

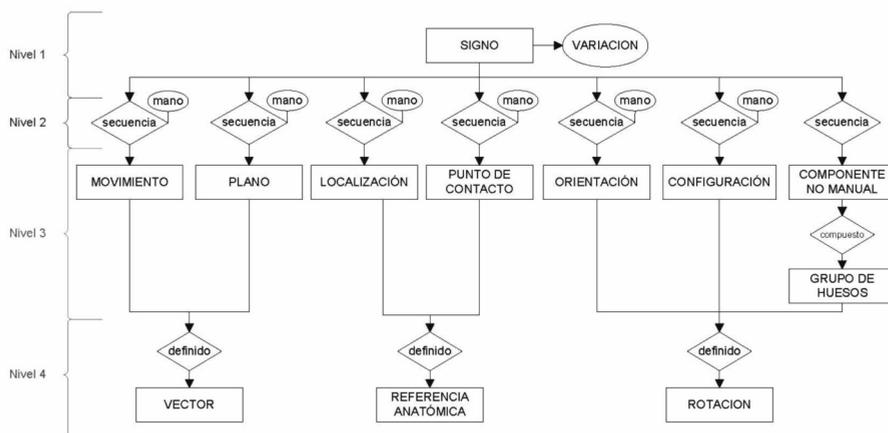


Ilustración 2. Diagrama de la base de datos relacional que permite almacenar las descripciones fonéticas de los signos

### 2.3.1.1. Nivel 1, los conceptos

Este primer nivel cumple la función de entrada en un diccionario basado en glosas. Cada una de las entradas de este diccionario fonético se acompaña de una relación con las posibles variaciones geográficas que puede sufrir un signo, ya que las variaciones gramaticales, expresivas o fonéticas deben ser definidas en el mensaje. En este nivel de la base de datos se define igualmente la duración del signo en su representación aislada, que posteriormente podrá ser modificada según la prosodia de la frase en la que se encuentre.

### 2.3.1.2. Nivel 2, la descripción fonológica y el CNM

En este segundo nivel de la base de datos se incluyen las secuencias temporales para cada uno de los parámetros fonológicos que definen el signo. En este nivel, se especifican los periodos temporales en la realización del signo

en los que cada fonema permanece estático. Por lo tanto, se interpretan los huecos como periodos de transición o movimientos internos. Dentro de este nivel también se considera el Movimiento, que define el desplazamiento de una mano siguiendo una determinada trayectoria.

El caso particular del CNM, dividido en múltiples secuencias, está contemplado en este nivel de la base de datos. Si el signo definido requiere definir algún elemento del CNM, podrá definir de manera individualizada cada uno de los subcanales que componen el CNM.

Las definiciones de los distintos parámetros fonológicos y CNM se realizan de manera independiente, por lo que se respeta la flexibilidad planteada como requisito al permitir ajustar la sincronización de los distintos parámetros y canales con total libertad.

### 2.3.1.3. Nivel 3, los fonemas y alófonos

Este tercer nivel recoge la lista de todos los fonemas y variaciones alofónicas existentes para cada parámetro fonológico: las configuraciones de la mano, las posibles orientaciones, las localizaciones frente al cuerpo, las posibles trayectorias de los movimientos, etc.

Esta lista de elementos es accesible tanto para la definición formal de los signos a partir de sus glosas, como para las posibles operaciones fonéticas que se puedan aplicar a un signo o el uso de clasificadores.

### 2.3.1.4. Nivel 4, descripción de los elementos articulatorios

Este último nivel contiene las definiciones primarias utilizadas para las unidades fonéticas. Todos los componentes de la LS se definen en base a tres elementos primarios:

- *Rotaciones*: La definición de una rotación se aplica en la *Configuración*, para definir la flexión de las articulaciones de los dedos, en la *Orientación*, para definir la flexión de la muñeca, y para el CNM, ya que permite definir la variación de los distintos elementos que lo componen (rotación de la mandíbula para definir su apertura, cejas, párpados, boca, toros, etc.).
- *Vectores*: Estos vectores asociados al espacio de signado nos permiten definir el *Plano*, la distancia horizontal al cuerpo, y el *Movimiento*, al utilizar una secuencia de vectores para definir la trayectoria.
- *Referencias anatómicas*: Los puntos anatómicos se utilizan para definir tanto la *Localización* como el *Punto de Contacto*.

### 2.3.2. Analogía con la lengua oral

Por último, y para facilitar la interpretación de la estructura lógica de la base de datos, se propone el siguiente símil que pretende describir la información que se recogería en la base de datos si se utilizara para la representación del español hablado: El nivel uno almacenaría cada concepto representable por el lenguaje. El nivel dos establecería la secuencia de unidades básicas, secuencias de alófonos, que unidos forman la palabra que representa el concepto propuesto. El nivel tres debería contener la lista de todos los alófonos que se requieren en el idioma propuesto (español). Por último, el nivel cuatro define a nivel acústico cada uno de los alófonos recogidos en el nivel tres.

## 3. Representación textual del mensaje signado

Un mensaje signado contiene distintos tipos de construcciones: secuencias de deletreo mediante el diccionario dactilológico, glosas de signos, descripciones fonológicas junto a descripciones del CNM, variaciones flexivas de signos, construcciones clasificatorias, etc. Estas construcciones que se plantean, requieren distintas aproximaciones en su descripción, debido a sus diferencias fonéticas, morfológicas y semánticas. Un signo dactilológico es un signo monosilábico y monomanual, que generalmente es estático, mientras que existen signos bimanuales complejos que requieren de una detallada descripción. En otro aspecto, algunas de estas unidades tienen una definición normativa, que puede ser establecida con anterioridad al proceso de traducción, como un signo dactilológico, o la forma neutra de un signo. Pero también existen otras construcciones que, en función del contexto del mensaje, requieren de una modificación fonética sobre alguno de los signos (el caso de una flexión, introversión o suplección sobre un signo) y por último, construcciones que no pueden ser definidas hasta que no se conoce el contenido del mensaje a transmitir, siendo el caso de las construcciones clasificatorias.

Para describir el mensaje a signar, hemos definido una notación que incluya todas las posibles construcciones que pueden aparecer en un mensaje de LSE, hemos llamado a esta notación HLSML. El objetivo de esta notación es el de simplificar la representación escrita del mensaje signado en los casos que sean factibles y permitir una descripción precisa del mensaje signado contemplando su complejidad. El uso de glosas simplifica la descripción del mensaje en gran medida; pero esta aproximación sólo sirve para elementos cuya descripción está previamente almacenada en el diccionario del sintetizador. Para los elementos que no están previamente definidos debemos

permitir, en la notación de entrada, la inclusión de su definición. Esta definición se puede obtener de otros recursos, como diccionarios temáticos.

Dada la compleja estructura multicanal que caracteriza a los mensajes signados, y por facilitar el uso de esta notación como entrada de un módulo de síntesis automática, HLSML se basa en XML (Bosak, 1996). Utilizando la estructura jerárquica de XML, podemos fácilmente representar los fenómenos multicanal que describió Huenerfauth (2009) para la síntesis de lengua de signos americana (ASL).

A continuación describimos la aproximación definida para los elementos que aparecen en un mensaje signado.

### 3.1. Glosas de signos y secuencias de deletreo

Tanto los signos del diccionario dactilológico, como los signos normativos, son signos cuya descripción fonética está definida. Por lo tanto, la descripción de estos elementos propios de un mensaje signado se puede realizar con anterioridad a cualquier proceso de traducción y la secuencia de fonemas que los definen se puede encontrar en la base de datos que describimos en la sección 2.3. Es por esto por lo que la opción elegida para indicar un signo de diccionario dentro de un mensaje signado es la utilización de glosas.

(2) *Para signar el concepto coche, el mensaje HLSML utiliza la glosa COCHE como se incluye en la siguiente sentencia:*  
`<sign value= "coche" />`

Una aproximación similar se plantea para la secuencia el deletreo de palabras mediante el diccionario dactilológico. La aproximación más sencilla en el mensaje es indicar qué palabra se quiere deletrear.

(3) *Para deletrear una palabra, el mensaje HLSML incluye la siguiente sentencia:*  
`<spell value= "elua" />`.

### 3.2. Flexión e introversión en el signo

En la LSE se observan los procesos de flexión e introversión en los que la definición fonética de un signo se modifica. Esta modificación responde a diversos fenómenos lingüísticos, en los que no vamos a entrar en este trabajo. Lo que se aplica de la flexión e introversión a la síntesis de LSE atañe al modo en que se indica este fenómeno en la descripción del mensaje y al modo que tiene el sistema de obtener la descripción correcta del signo que debe representar.

Los morfemas ligados que modifican al signo pueden ser:

- Morfemas flexivos o desinencias, que añaden elementos a su forma fonológica, como un movimiento o la repetición del movimiento original del signo (que acarrea la repetición íntegra del signo, a veces en cierta dirección). [...]
- Morfemas introflexivos, que modifican alguno de los fonemas del signo (lugar, configuración, orientación y movimiento) para representar distintos valores gramaticales (Herrero Blanco, 2009)

Como indica Herrero Blanco, un morfema flexivo añade elementos a la forma fonológica de un signo, llegando incluso a repetir completamente el signo en una determinada dirección. Esta última modificación es morfológicamente asimilable a la realización del signo pero modificando su inicio, y realizándolo en un lugar distinto. Es por lo tanto una modificación del valor inicial (primera sílaba) de los parámetros Localización y Plano.

La definición de una modificación introflexiva de un signo requiere definir qué parámetro fonológico del signo se modifica y qué fonema debe tomar dicho parámetro. Anteriormente se ha expuesto la idoneidad de utilizar glosas para la definición del mensaje. Esta aproximación es también explotada para la definición del nuevo valor de dicho parámetro. La notación HLSML permite definir el nuevo fonema utilizando dos aproximaciones: (i) La primera es utilizar el nombre del fonema que se desea utilizar. El sistema que se utiliza para nombrar las unidades de los fonemas en la base de datos sigue un sistema numérico, pero dada la flexibilidad de HLSML, resultaría trivial adaptar el sistema para que utilice los nombres que se les da a estas unidades en otras notaciones como HamNoSys (Prillwitz *et alii*, 1989), a través de su versión adaptada para ordenador SiGML (Elliot *et alii*, 2004), o SEA (Herrero Blanco, 2003). (ii) La segunda opción que se plantea es la de utilizar la glosa de un signo. En esta aproximación el sistema buscará la definición fonológica de este segundo signo y utilizará dicha definición para realizar la modificación introflexiva del primero. En esta segunda opción, cuando se trata de la modificación de la Configuración, se considera que la modificación se realiza utilizando los fonemas del segundo signo, no el clasificador correspondiente dicho signo. Por ejemplo, la configuración del signo COCHE es la de puño cerrado, pero el clasificador que se utiliza para representar al coche usa una configuración de mano abierta y todos los dedos juntos.

(4) a. Para la realización del signo MESES-TRES, se sustituye la configuración del signo MES por la configuración del numeral TRES, utilizando HLSML la definición se realiza de la siguiente manera:

<sign value= "mes">

<inflectiveModification value= "configuration">

```

    <sign value= "tres" />
  </inflectiveModification>
</sign>

```

(4) b. Para la realización del signo VER-a-mi, se sustituye la orientación del signo VER por la orientación del signo YO, utilizando HLSML la definición se realiza de la siguiente manera:

```

<sign value= "ver">
  <inflectiveModification value= "orientation">
    <sign value= "yo" />
  </inflectiveModification>
</sign>

```

### 3.3. Descripción de los signos

En el apartado 3.1 hemos descrito la aproximación utilizada en HLSML para definir un mensaje que incluye signos de diccionario. Los signos que se recogen en la base de datos pertenecen al diccionario normativo de la LSE y en algún caso a variantes dialectales. En todos los casos, se trata de signos que han sido aceptados dentro de la comunidad sorda. Por lo tanto, los neologismos o tecnicismos que puedan darse en una conversación pueden no estar recogidos dentro de un diccionario nuclear y por tanto en la base de datos. El proceso de normalización de un signo no es objeto de discusión en este trabajo, pero si resulta relevante contemplar un método por el cual el sintetizador pueda representar signos que no estén en la base de datos normativa.

En este caso, es el propio mensaje descrito en la notación HLSML el que debe recoger la descripción fonética del signo. La solución que podría plantearse es la de utilizar notaciones alfanuméricas ya existentes como SEA (Herrero Blanco, 2003) o basadas en un formato XML, como SiGML (Prillwitz *et alii*, 1989) o SWML (da Rocha y Pereira, 2004). Pero en ninguna de estas notaciones anteriores existe la posibilidad de realizar una descripción fonética en la que se describa explícitamente la duración de los distintos bloques *Hold* y *Move* de cada segmento o la duración de las sílabas definidas.

La solución que incluye HLSML es la descripción de toda la información que aparece en el modelo antes descrito. Esta descripción recoge, para cada uno de los fonemas del signo, la misma información que pudiera estar almacenada en la base de datos (ver sección 2.3). La información omitida en la descripción es completada por el sistema suponiendo que no debe haber actividad en esos fonemas (Ejemplo: si la definición fonética en el mensaje HLSML no incluye información para la mano pasiva, se considera un signo monomanual).

(5) *La descripción fonética del signo YO utilizando la notación HLSML es de la siguiente manera:*

*La definición comienza describiendo la duración que debe tener el signo, en este caso ese valor es de 500 ms. A continuación se describen los valores que se asignan a cada uno de los fonemas que intervienen en el signo: La configuración debe establecerse con el índice extendido. El código del fonema orientación es el siguiente: el plano elegido, en este caso es el horizontal (h); la dirección en ese plano, hacia atrás (b); y la dirección de la palma, en este caso hacia el exterior (e). La localización se define en el centro del pecho y el plano define la existencia de un contacto entre la mano dominante y el cuerpo. Por último, se define que la parte de la mano que debe entrar en contacto con la localización es el extremo del dedo índice. El valor de todos los fonemas de esta descripción se deben mantener estáticos durante todo el signo (desde el instante 0 hasta el 100 en una escala porcentual):*

```
<holdMoveDefinition time= "500">
  <configuration><phoneme value= "point_finger" side= "dominant"
    fraction_ini= "0" fraction_end= "100" /></configuration>
  <orientation><phoneme value= "h_b_e" side= "dominant" side= "dominant"
    fraction_ini= "0" fraction_end= "100"/></orientation>
  <location><phoneme value= "chest" side= "dominant" fraction_ini= "0"
    fraction_end= "100" /></location>
  <plane><phoneme value= "contact" side= "dominant" fraction_ini= "0"
    fraction_end= "100" /></plane>
  <contact><phoneme value= "point_end" side= "dominant" fraction_ini= "0"
    fraction_end= "100" /></contact>
</ holdMoveDefinition>
```

### 3.4. Iconicidad de las LS, conmutación referencial y expresión referida

El signante es capaz de transmitir un mensaje no sólo a través de una secuencia de signos; mediante acciones que podría considerarse que rayan en la mímica, puede transmitir información con una gran plasticidad. “Esta acción representada emplea el código paralingüístico, mediante una simulación de la acción real. (Herrero Blanco, 2009: 379–384)” Así pues estudiemos el siguiente ejemplo:

(6) *La frase en español encajar la maleta en el portamaletas se representa en LSE utilizando el signo MALETA seguido de un clasificador en el que mediante la mano izquierda se presenta la posición del portamaletas a una altura sobre la cabeza, y con la configuración de esta mano se presenta la forma de la boca del portamaletas. Con la mano derecha se simula sujetar el asa de la maleta y se realiza un movimiento que simula el esfuerzo de encajar una maleta en este habitáculo.*

El proceso automático de traducción que requiere pasar de la frase encajar la maleta en el portamaletas, a la representación antes descrita, requiere de análisis semántico y de un conocimiento de mundo de gran complejidad. Por lo tanto una traducción automática que permita obtener dicha descripción plantea una línea de investigación con mucho trabajo por delante, del que ya se han comenzado a dar algunos pasos (Huenerfauth, 2006). Pero recordemos que en este trabajo nos estamos centrando exclusivamente en la síntesis y en la definición de una notación que permita describir una escena como la anterior para que el avatar la represente, esto se ha logrado con HLSML. En los párrafos siguientes se resumen distintas teorías y aproximaciones lingüísticas sobre estas estructuras del mensaje signado y que deben ser contempladas para la descripción de un mensaje signado. La literatura se refiere a este tipo de construcciones como los *classifier predicates* (Supalla, 1978 y 2003) o tratados como construcciones clasificatorias (Herrero Blanco, 2004). Liddell (1996) recurre a la teoría de espacios mentales para la definición de estas construcciones. La descripción mímica que está haciendo el signante sólo adquiere sentido cuando se establece una relación con la realidad descrita, en lo que Liddell denomina un espacio mental fusionado. El signante utiliza sus manos para completar descripciones de objetos y situarlos en el espacio de signado. Liddell (2003) defiende las características lingüísticas de este tipo de clasificadores, dado que aunque habla de un proceso de flexibilización de los parámetros formativos sigue considerándolos como tal (morfemas).

Cogill-Koez (2000) plantea que estos clasificadores pueden responder a un conjunto de plantillas preestablecidas. En su disertación define para cada uno de los parámetros fonológicos, cómo pueden adaptarse a estas plantillas prefijadas. El parámetro Configuración parece ser el más regulado, y como se ve en el trabajo de Huenerfauth (2004) es un hecho bastante aceptado. Según esto, ciertas configuraciones se relacionan de manera constante con la representación de ciertos elementos, como pueden ser personas, animales u objetos. El parámetro que no consigue establecerse dentro de una plantilla preestablecida es el parámetro de Lugar. Esto adquiere sentido si se aplica la teoría de Liddell (1996), ya que dicha construcción respondería a la representación del espacio mental del signante. Este hecho hace que Cogill-Koez dude de la naturaleza de estos clasificadores como entidades lingüísticas.

Dentro de esta representación se incluye el conjunto de modificaciones que es capaz de incluir el signante en el discurso para emular un diálogo entre dos o más personas. Lillo-Martin (1995) defiende que el signante (mediante movimientos de los hombros y la cabeza, cambiando la dirección en la que mira y/o variando la expresión de la cara) tenderá a interpretar al personaje

activo dentro de la presentación. Aunque según rebaten Liddell y Metzger (1998), esta acción no suele estar presente a lo largo de toda una representación, por lo que no debe ser considerada como el único elemento en la definición de papeles. De manera paralela, Engberg-Pedersen (1995) añade a la variación en la expresión del signante para cambiar de papel, un cambio de referencias. Este cambio se aplica, tanto al espacio de signado donde ha colocado de manera imaginaria objetos, como a la realización de pronombres (en este caso, un pronombre de primera persona no hace referencia al narrador).

Por último, se debe resaltar la aportación que hacen Liddell y Metzger (1998) planteando el concepto de acción construida definido como la capacidad que tiene el narrador (signante) para representar la actitud, posición y acciones de los personajes sobre los que está narrando, ejecutando dichas acciones en el mundo real. En particular, la acción más frecuente se basa en la dirección en la que los ojos y la cabeza apuntan.

Independientemente de si estas construcciones responden a gestos lingüísticos o no, su morfología difiere en cierta medida de los signos de diccionario, lo cual resulta relevante para la síntesis de un mensaje signado. Si bien la descripción de determinados fonemas como la Configuración y la Orientación puede realizarse siguiendo la aproximación expuesta en el apartado anterior, la posición de las manos del signante se rige por un sistema de posicionamiento continuo, en oposición al sistema discreto definido por los parámetros Plano y Localización. De igual modo, la definición del movimiento de las manos debe permitir la representación de cualquier trayectoria, no quedando limitado a los movimientos fonéticos de la LSE.

Como se ha planteado antes, este tipo de elementos tiene una importante dependencia del contenido del mensaje y de lo que se desea representar. La gran variabilidad de representaciones que se pueden producir al contemplar estos clasificadores hace que sea imposible tener descritas todas las posibilidades. Además, la posición de las manos no se puede regir por los fonemas Plano y Localización. Por estas razones, el uso de la base de datos para la descripción de estas unidades no es posible. En un sistema automático, como el que se planteó al comienzo de este trabajo, se requiere que el análisis del contenido del mensaje a representar y la generación de la descripción del clasificador que se utilice, se realicen en el módulo de traducción, con anterioridad a la síntesis. Este requisito obliga a la notación HLSML a incluir un método para describir estos clasificadores:

(7) *HLSML utiliza la siguiente etiqueta para indicar la existencia de un predicado clasificador:*

*<classifierPredicate>*

...  
 </classifierPredicate>

*Entre estas dos etiquetas se pueden encontrar las descripciones que se presentan en los párrafos siguientes.*

El primer problema a solventar es la definición de la posición de las manos ya que el sistema fonológico basado en la separación de Localización y Plano, resulta insuficiente. Por esta razón, hemos utilizado un sistema cartesiano que representa el espacio léxico (Herrero Blanco, 2009: cap. 3.7). Este sistema toma como origen de coordenadas el centro del pecho y mediante una coordenada (X,Y,Z) se define la posición de la mano en el espacio de signado (podemos considerar que los lugares fonológicos descritos por Herrero son casos particulares en nuestro sistema de posicionamiento). En este sistema de coordenadas, la X indica el eje horizontal, de izquierda y derecha, la Y el eje vertical, arriba y abajo, y el valor de Z el eje frontal, adelante y atrás. El uso más frecuente de estas construcciones es la describir una escena, en el espacio de signado, relacionada con el discurso signado. Por lo tanto las acciones más frecuentes que se realizan son: situar objetos en el espacio signado y manipularlos (rotarlos, escalarlos y desplazarlos). Los objetos que se sitúan en el espacio de signado requieren del uso del clasificador correspondiente, que define los fonemas utilizados en la Configuración y la Orientación, y de la posición del objeto en el espacio de signado.

(8) *Supongamos que es necesario situar un gato en el espacio de signado, utilizando la mano derecha (dominate), frente al signante, a la altura del ombligo y bajo el hombro, además hay que definir el tiempo que la mano se queda en esa posición (salvo que se use un valor por defecto). Utilizando la notación HLSML se define que se desea situar un objeto en el espacio de signado. La categoría de este objeto es de un cuadrúpedo, lo que define la configuración y orientación necesarias para esta acción. La mano que realizará esta acción es la dominante y el tiempo que debe permanecer la mano en esa posición es de 200 ms. A continuación se establece que la posición en la que debe situarse la mano está en las coordenadas (1,-1,1), siguiendo el sistema de coordenadas que describimos antes:*

```
<placeObject value= "cuadruped" side= "dominant" time= "200">
  <coordinate horizontal= "1" vertical= "-1" frontal= "1"/>
</placeObject>
```

Dentro de la descripción de la escena, el signante puede representar el desplazamiento de los objetos referenciados, desplazando sus representaciones en el espacio de signado. El desplazamiento de dichos objetos emula, en cierto grado, el realizado en el mundo real. Por esta razón, la descripción de las trayectorias que deban seguir las manos del signante tiene que ser flexible

para acomodarse a la arbitrariedad de estos movimientos. La notación HLSML permite definir cualquier trayectoria en el espacio de signado a través de sus puntos intermedios; según la complejidad de la trayectoria, serán necesarios distinto número de puntos.

(9) *Supongamos que el gato del ejemplo (8) avanza en línea recta alejándose del signante. La duración total del movimiento es de un segundo, indicado en el mensaje HLSML como 1000 ms; dado que sigue tratándose de un cuadrúpedo, la categoría se mantiene y se sigue indicando que la ejecución se lleva a cabo con la mano derecha. La trayectoria que sigue la mano se especifica por una secuencia de coordenadas del espacio de signado, en este caso (1,-1,1) y (1,-1,4):*

```
<moveObject value= "cuadruped" side= "dominant" time= "1000">
  <coordinate horizontal= "1" vertical= "-1" frontal= "1"/>
  <coordinate horizontal= "1" vertical= "-1" frontal= "4"/>
</moveObject>
```

En ciertas ocasiones el signante requiere realizar acciones más complejas que pueden requerir el uso de configuraciones no asociadas a un clasificador (realizaciones muy cercanas a la mímica) u orientaciones que representen la rotación de un objeto en el espacio de signado. Para este tipo de construcciones del mensaje signado, HLSML plantea una aproximación fonética básica (similar a la expuesta en el apartado 3.3), pero adaptada a los requisitos de posicionamiento de las manos descritas en este apartado. El mensaje HLSML permite definir la Configuración y Orientación de la mano utilizando un valor fonético y el lugar de signado mediante el sistema de coordenadas.

(10) *Para transmitir el mensaje de un coche volcado, se signa COCHE, se posiciona dicho coche en el espacio de signado y a continuación se rota la mano asemejando la rotación que sufre el coche al volcarse. La posición del coche en el espacio se describe como en el ejemplo (8), en este caso la categoría del objeto es diferente y se usará una configuración distinta. El elemento <positionTransition> indica el tiempo que se tarda de pasar entre las dos posiciones estáticas. Dado que la representación del coche volcado no es la estándar, la orientación que se asume para el clasificador del coche por defecto es en el plano horizontal (h), dedos orientados al frente (f) y la palma hacia abajo (b), el clasificador indica que hay una rotación y la palma debe orientarse hacia el interior (i). Utilizando la notación HLSML, el clasificador y el efecto de rotación se representan así:*

```
<placeObject value= "coche" side= "dominant" time= "200">
  <coordinate vertical= "-1" horizontal= "1" frontal= "1"/>
</placeObject>
<positrionTransition time= "500" time= "200"/>
<staticPosition side= "dominant">
  <configuration><phoneme value= "open_hand" /></configuration>
  <orientation><phoneme value= "h_f_i" /></orientation>
```

```
<coordinate horizontal= "1" vertical= "-1" frontal= "1"/>
</staticPosition>
```

Por último, el signante utiliza la mirada para expresar concordancia con el sujeto o el objeto del predicado. Tanto el sujeto o el objeto se han situado con anterioridad en el espacio de signado. Este fenómeno debe ser contemplado, permitiendo que el avatar dirija su mirada a un punto determinado del espacio de signado. En nuestra aproximación contemplamos de manera independiente la definición de la orientación de la cabeza y la definición de la dirección de la mirada. Para ello, HLSML presenta dos alternativas: La primera es definir la coordenada en el espacio de signado como definimos antes. La segunda es indicar el punto de la anatomía del signante al que se debe dirigir la mirada:

(11) a. *Supongamos que durante la signación, se debe mirar hacia el gato que planteamos en el ejemplo (8), ya que es el objeto de la frase que se esté signando. El avatar debe orientar su mirada hacia la posición del gato, situado en las coordenadas (1,-1,1). Utilizando la notación HLSML:*

```
<eyesLookAt>
  <coordinate horizontal= "1" vertical= "-1" frontal= "1"/>
</eyesLookAt>
```

(11) b. *Supongamos que durante la signación, se debe mirar hacia el gato que planteamos en el ejemplo (8), ya que es el objeto de la oración. En esta ocasión, la mano del signante sigue en esa posición, luego el avatar debe orientar su mirada hacia la mano (que representa el gato). Utilizando la notación HLSML:*

```
<eyesLookAt>
  <element value= "right_hand"/>
</eyesLookAt>
```

### 3.5. Clasificadores afijales

El último elemento de un mensaje signado que contemplamos en HLSML es la clasificación afijal.

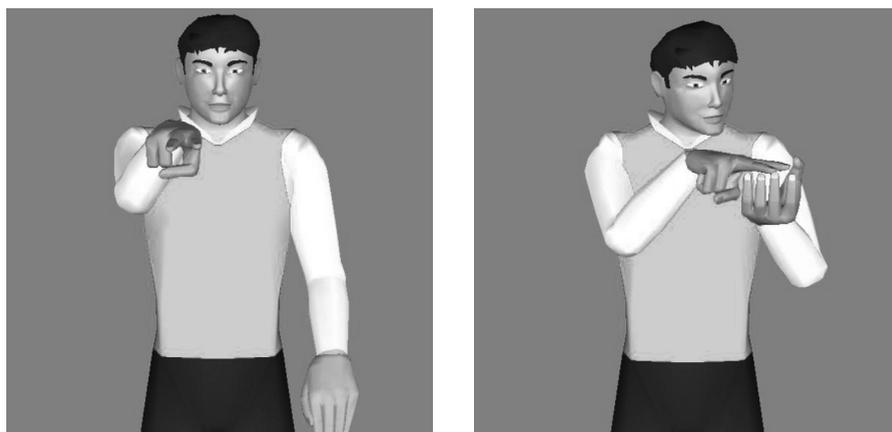
Una de las características fonológicas de las LSs es que cuentan con dos articuladores manuales, las dos manos. [...]

En estas construcciones la mano pasiva representa el objeto de la acción expresada por la mano activa, a través de un clasificador de ese objeto o, en su defecto, de su misma configuración, [...] Sólo el lugar de realización del signo y la orientación se especifican en relación a la mano pasiva. (Herrero Blanco, 2004).

Estas construcciones aúnan dos de los efectos que hemos descrito con anterioridad, la representación icónica de un objeto, mediante el uso de un

clasificador en la mano pasiva, y la modificación fonética de un signo realizado por la mano dominante. La modificación que se aplica al signo realizado por la mano dominante modifica los valores de los fonemas Localización, Plano y Orientación en función de la posición de la mano pasiva. Esta modificación añade un afijo, en el caso de signos monosilábicos, o modifica la última sílaba del signo. En ambos casos, la construcción resultante comienza en el lugar de articulación correspondiente al signo original, pero termina en la posición de la mano pasiva y orientada hacia ésta.

Esta construcción muestra la necesidad de describir realizaciones simultáneas en los distintos canales que compone la comunicación signada. Además, existe una relación de dependencia entre las realizaciones de los distintos canales (la modificación que se aplica a la mano dominante está en relación a la posición de la mano pasiva. De igual manera, si el signante dirige su mirada a la mano pasiva, nos encontramos con una modificación del canal correspondiente al CNM). Proponemos un ejemplo en la ilustración 3 sobre el efecto de aplicar un clasificador afijal sobre el signo MIRAR-DETENIDAMENTE.



**Ilustración 3.** Resultado de una clasificación afijal sobre el signo MIRAR-DETENIDAMENTE. A la izquierda, la representación del signo sin modificar. A la derecha, el signo se orienta para que el objeto de la acción sea la entidad definida por la mano izquierda (en este caso se trata de mirar una naranja)

Dado que nos encontramos frente a una construcción que toma la descripción de un signo y la modifica, HLSML utiliza la misma estructura que para las construcciones introflexivas. Pero en este caso no se indica el parámetro fonológico a modificar y el nuevo fonema, se indica que se trata de una modificación afijal (con lo que se sabe qué parámetros se modifican) y cuál

es el punto al que se debe orientar la mano dominante (por defecto se toma la palma de la mano pasiva). Una vez presentada la aproximación mediante la cual indicamos la existencia de esta modificación, debemos resolver el segundo problema antes planteado: describir la simultaneidad en la realización de distintos elementos del mensaje. Esta descripción debe estar en consonancia con el modo que se utilice para describir la secuencialidad de los elementos de un mensaje. Recordemos que la estructura de cualquier notación basada en XML, como es el caso de HLSML, es jerárquica. Por lo tanto, podemos definir características similares a todos los elementos que desciendan del mismo antecesor. La notación HLSML utiliza para la descripción de la secuencialidad y la simultaneidad dos elementos auxiliares: <sentence> y <compound> respectivamente. Todos los elementos que descienden de <sentence> se representan en secuencia y todos los que dependen de <compound> son representados de manera simultánea.

(12) *Supongamos que se desea indicar que se está observando el gato que planteamos en el ejemplo (8) y se utiliza un clasificador afijal. La mano izquierda sitúa el gato en el espacio de signado como hemos indicado en dicho ejemplo. Al mismo tiempo se realiza el signo MIRAR\_DETENIDAMENTE, sobre el que se aplica la clasificación afijal para que el final de la ejecución del signo, la mano derecha esté orientada y situada en relación con la posición hipotética del gato, definida por la mano izquierda:*

```
<compound>
  <classifierPredicate>
    <placeObject value= "cuadruped" side= "left" time= "200">
      <coordinate vertical= "-1" horizontal= "1" frontal= "1"/>
    </placeObject>
  </classifierPredicate>
  <sign value= "mirar_detenidamente">
    <inflectiveModification value= "affixal">
      <element value= "left_hand"/>
    </inflectiveModification>
  </sign>
</compound>
```

### 3.6. Simultaneidad del CNM en el discurso

Los CNM en la LSE aportan información morfológica, sintáctica y discursiva. Esta acción se realiza de manera simultánea a las acciones manuales correspondientes a la realización de otros elementos del mensaje signado (como signos de diccionario). Por lo tanto se debe contemplar este hecho en la descripción del mensaje signado. La necesidad de describir la realización simultánea de

distintos elementos dentro del mensaje signado se puso de manifiesto al tratar los clasificadores afijales. La aproximación que tomamos para esta problemática se basa igualmente en el uso del elemento <compound> de HLSML para indicar que todos los elementos que descienden de éste, deben ser realizados de manera simultánea. Sólo resta definir el elemento que nos permite describir una realización en el CNM: <nonManualSequence>

(13) *En la LSE, las construcciones interrogativas absolutas se realizan mediante la apertura de los ojos, elevación de las cejas e inclinación de la cabeza (Herrero Blanco, 2009: 93). Esta acción se presenta de manera simultánea a la realización del signo de la pregunta. En este ejemplo, utilizamos HLSML para representar la pregunta ¿Vienes?*

```
<compound>
  <sign value= "venir" />
  <nonManualSequence>
    <phoneme value="eyes_open"/>
    <phoneme value="eyebrows_up"/>
    <phoneme value="head_bent_front"/>
  </nonManualSequence>
</compound>
```

Como discutimos en el apartado 2.1.4, el CNM está subdividido en múltiples canales paralelos e independientes. En este ejemplo los valores *eyes\_open* (abrir ojos) y *eyebrows\_up* afectan a ambos ojos y ambas cejas de manera simultánea agrupándolos. Esta agrupación permite la simplificación en la definición del mensaje, ya que equivaldría a los valores *l\_eye\_open* y *r\_eye\_open* (abrir ojo izquierdo y abrir ojo derecho). Del mismo modo, una expresión facial característica, utilizada para una función lingüística, puede ser predefinida (ej. *question\_absolute*, para representar el CNM de este ejemplo).

#### 4. Herramientas de definición fonética

El objetivo de la aplicación desarrollada es proporcionar una interfaz que permita la definición y visualización de signos en LSE a partir de sus componentes articulatorios. Para ello se ha diseñado una interfaz gráfica que permite definir un signo utilizando únicamente el ratón y el método “pinchar y arrastrar” para la gestión de todos los fonemas. A continuación se describe brevemente la aplicación (ver ilustración 4), una definición más detallada puede encontrarse en (López-Colino y Colás, 2009b; López Colino, 2009). La interfaz de la aplicación se divide en dos secciones principales: en la parte de la izquierda, el usuario tiene acceso a un avatar 3D que reproduce el signo. Este avatar reproduce los signos siguiendo las descripciones realizadas por el usuario de la aplicación. El usuario puede variar la posición de la cámara, visualizar el signo

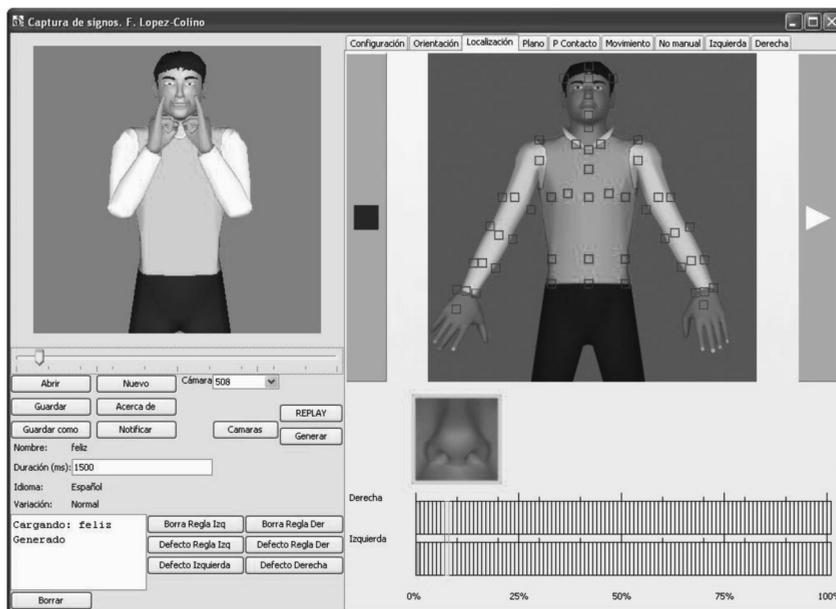


Ilustración 4. Vista general de la aplicación utilizada para la definición y visualización de signos en LSE

las veces que desee o detener la reproducción del signo en un instante concreto de su definición. La sección de la derecha de la aplicación contiene un conjunto de paneles que permiten realizar, y visualizar, la definición de los signos en función de sus fonemas. Cada uno de los seis paneles correspondientes a los componentes manuales contiene dos secuencias temporales en las que se recoge la definición del fonema correspondiente para cada una de las manos del avatar. En el caso del CNM se permite una doble definición; se puede definir mediante ‘expresiones simples’ que definen la única característica de este CNM, como puede ser la posición de cada ceja, la forma de la boca, la posición de la lengua, la rotación de la cabeza, etc. De igual modo, el CNM admite el uso de ‘expresiones complejas’ que agrupan varias de las ‘expresiones simples’, como puede ser el caso de una expresión facial que represente sorpresa, o un alzamiento de los hombros para definir duda, etc.

#### 4.1. Interactuando con la definición de los signos

La gran ventaja de esta aplicación reside en la posibilidad de modificar la descripción fonética de un signo y acceder de manera inmediata a una visualización 3D de la ejecución de la misma. Para simplificar el proceso, todos los parámetros

se definen utilizando imágenes. De esta manera, los usuarios no tienen necesidad de conocer ninguna de las notaciones existentes para la definición fonológica ni fonética de los signos. En estas secuencias, el usuario puede fácilmente definir los bloques que Liddell y Jonhson definirían como Hold para el fonema correspondiente. Como indicamos antes el espacio entre dos bloques Hold, se considera que es una transición o movimiento interno. La gestión del Movimiento externo sigue la misma aproximación, pero en este caso el segmento que define el usuario corresponde al periodo de tiempo en el signo en el que la mano está realizando ese movimiento. La ilustración 5 muestra un ejemplo de definición de dos signos monomanuales (la definición de la mano pasiva se realizaría de manera similar y por eso no incluimos un ejemplo relativo que incluya dicha mano).

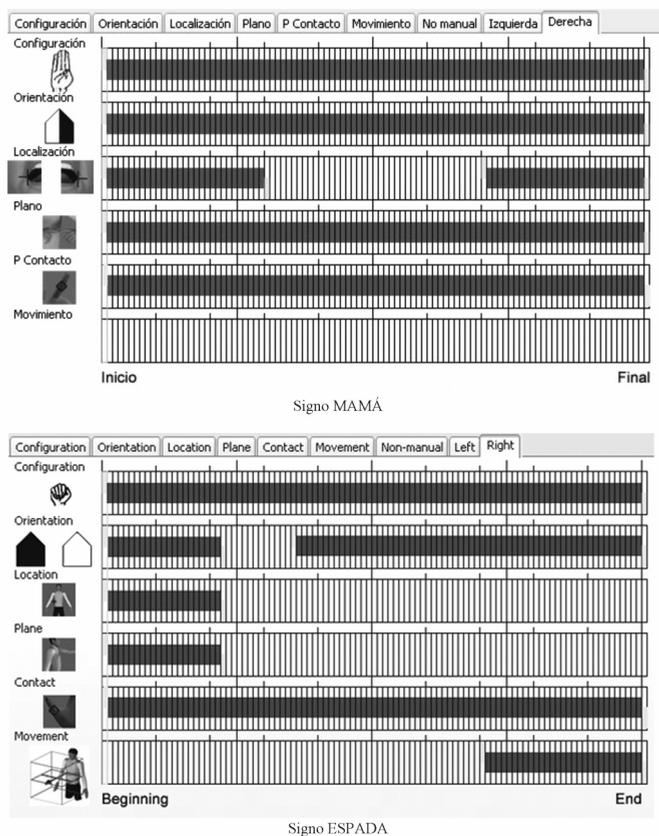


Ilustración 5. Ejemplo de dos signos monomanuales en LSE definidos con la aplicación de LSE. Las barras azules corresponden a bloques *Hold* para ese fonema, salvo para el fonema del parámetro *Movimiento*, que implica el desplazamiento de la mano

Aparte de los paneles relativos a cada parámetro, existen dos paneles que permiten visualizar todas las secuencias relativas a cada una de las manos. Este panel permite visualizar y establecer de manera sencilla la sincronización entre las secuencias de los distintos fonemas para cada una de las manos.

## 5. Evaluaciones

El último apartado que se presenta en este trabajo describe la evaluación de la inteligibilidad de los signos sintetizados. Anteriormente se ha expuesto la aproximación fonológica, el modo de representar las definiciones de los signos, la aproximación a la notación de descripción textual de las unidades que componen el mensaje signado y las distintas modificaciones que se pueden introducir en un mensaje signado. Estas definiciones y aproximaciones no tendrían gran impacto, si su aplicación generara signos sintéticos ininteligibles. Por esa razón se han planteado una serie de evaluaciones con nativos de LSE.

Estas evaluaciones se plantearon para minimizar el impacto que tiene en el resultado la presencia de un evaluador y un intérprete durante la realización de las pruebas. Por ese motivo se optó por realizar todas las evaluaciones mediante páginas web en las que los usuarios podían indicar libremente las respuestas que consideraran apropiadas.

Las evaluaciones se centraron en medir la inteligibilidad de la síntesis en los siguientes contextos y restricciones: evaluaciones con signos individuales, permitiendo tanto respuestas abiertas como cerradas en forma de test, frases sencillas, basadas principalmente en secuencias de signos, frases con predicados clasificadores y expresiones referidas. Los experimentos se propusieron a un grupo de entre siete y once individuos, con edades comprendidas entre los veinticuatro y cincuenta años. Algunos de ellos pertenecientes a la Fundación CNSE. Los resultados obtenidos en las distintas pruebas se muestran en el cuadro 1.

<b>Evaluación</b>	<b>Media de aciertos</b>	<b>Varianza de respuestas</b>
<b>Signos aislados y respuesta abierta</b>	96,4 %	6,1 %
<b>Signos aislados y respuesta cerrada</b>	98,6 %	2,7 %
<b>Frases como secuencias de signos establecidos</b>	95,0 %	5,0 %
<b>Frases con predicados clasificadores y expresiones referidas</b>	92,2 %	12,3 %

Cuadro 1. Resultado de las distintas evaluaciones sobre los contenidos sintéticos generados por el sintetizador

Los resultados obtenidos en esta evaluación superan a los que otros autores han logrado para distintas lenguas de signos (Cox *et alii*, 2002; Cox *et alii*, 2003; Kennaway *et alii*, 2007).

## 6. Conclusión

En este trabajo se ha descrito la aproximación lingüística que se ha seguido para la implementación de un sintetizador de LSE. Los aspectos más relevantes de este trabajo se han centrado en la fundamentación lingüística, el modo de describir un mensaje en LSE para que el sintetizador pueda representarlo y destacando los puntos más relevantes de la herramienta que permite la definición de los signos y los resultados obtenidos de la síntesis de LSE.

La aproximación fonológica que se ha utilizado en este trabajo aún distintas aproximaciones publicadas en los últimos años. Esta fusión tiene como objetivo permitir la gestión automática, de la manera más sencilla posible, y dar lugar a la síntesis de contenidos, con la mayor calidad posible. La calidad de la síntesis de signos, en cuanto a su inteligibilidad, ha sido medida en evaluaciones llevadas a cabo por signantes de LSE y han proporcionado una tasa de reconocimiento de signos superior al noventa y cinco por ciento de aciertos. Estos resultados no significan el final de la investigación, se debe seguir desarrollando y evaluando el sistema de síntesis junto con el desarrollo del sistema de traducción. Además, es muy relevante realizar estas evaluaciones con un número elevado de personas. En todas las evaluaciones citadas en este trabajo, se describe que el número de usuarios que realizaron la evaluación ronda la decena. Con la aproximación descrita para el protocolo de evaluación, el número de personas que pueden participar es muy elevado, lo que aumentará la significancia de los resultados.

La descripción textual de los mensajes signados contempla tanto la existencia de signos establecidos o de diccionario, secuencias deletreadas, modificaciones fonéticas y construcciones clasificatorias. En particular, se ha establecido una aproximación que permite la descripción de la iconicidad inherente al mensaje signado. En la evaluación de estas construcciones, los signantes reconocieron correctamente el noventa y dos por ciento de las escenas representadas.

Por último, y como ya se anotó en la introducción, se debe resaltar que en ciertos casos la información que se debe definir de manera explícita excede a las necesidades de cualquier persona con conocimientos de LSE, o no sigue una aproximación puramente lingüística, pero recordemos que estamos tratando con máquinas, sin ningún tipo de intuición o experiencia. Y que el objetivo no es realizar una formalización lingüística de la LSE, sino generar contenidos sintéticos en LSE que puedan ser entendidos por los signantes de esta lengua.

## 7. Agradecimientos

Los autores desean expresar su agradecimiento a la Fundación CNSE por su apoyo, confianza e interés en este trabajo, proponiendo mejoras e instando a continuar con la labor realizada sobre la síntesis de LSE. Por último, agradecer al programa de becas FPU-UAM por la financiación recibida.

## Referencias bibliográficas

- Baldassarri, S. y F.J. Royo Santas (2007): “Traductor de Español a LSE basado en reglas gramaticales y morfológicas”. *Actas de INTERACION*.
- Battison, R. (1974): “Phonology in American Sign Language: 3-D and digit-vision”. *Actas de la California Linguistic Association Conference*.
- Corina, D. P. (1996): “Sign Linguistics Phonetics, Phonology and Morpho-syntax”, *Lingua*, 98, págs. 73–102.
- Bosak, J. (1996): “Extensible Markup Language”. En T. Bray; J. Paoli; C. Sperberg-McQueen; E. Maler y F. Yergueau (eds.) W3 Consortium: <http://www.w3c.org/TR/REC-xml/> (19-7-2010).
- Brentari, D. (1996): “Trilled movement: Phonetic realization and formal representation”, *Lingua*, 98, págs. 43–71.
- Cogill-Koez, D. (2000): “A model of signed language ‘classifier predicates’ as templated visual representation”, *Sign Language and Linguistics*, 3, págs. 209–236.
- Cox, S., M. Lincoln, J. Tryggvason, M. Nakisa, M. Wells, M. Tutt y S. Abbott (2002): “Tessa, a system to aid communication with deaf people”, *Actas de 5th international ACM conference on Assistive Technologies*, ACM, págs. 205–212.
- Cox, S., M. Lincoln, M. Nakisa, M. Wells, M. Tutt y S. Abbott (2003): “The Development and Evaluation of a Speech-to-Sign Translation System to Assist Transactions”, *International Journal of Human-Computer Interaction*, 16, págs. 141–161.
- Elliot, R., J. Glauert, V. Jennings y R. Kennaway (2004): “An Overview of the SiGML Notation and SiGML Signing Software System”. *Actas de Language Resources and Evaluation Conference*, Association pour le Traitement Automatique des Langues, págs. 98–104.
- Engberg-Pedersen, E. (1995): “Point of view expressed through shifters”. En K. Emmorey y J. Reilly (eds.) *Language, Gesture, and Space*. págs. 133–154.
- Fernández Villalibre, F. (1851): *Diccionario usual de mímica y dactilología: útil a los maestros de sordo-mudos, a sus padres y a todas las personas que tengan que entrar en comunicación con ellos*. Imprenta del Colegio de Sordo-mudos y Ciegos.
- Fundación CNSE. (2008): *Diccionario Normativo de la Lengua de Signos Española*. Madrid, Fundación CNSE.
- Gobierno de España, (2007): “LEY 27/2007, de 23 de octubre, por la que se reconocen las lenguas de signos españolas y se regulan los medios de apoyo a la comunicación oral de las personas sordas, con discapacidad auditiva y sordociegas”. *Boletín Oficial del Estado*.

- Herrero Blanco, A. (2003): *Escritura Alfabética de Lengua de Signos Española. Once lecciones*. Alicante, Publicaciones de la Universidad de Alicante.
- Herrero Blanco, A. (2004): “Una aproximación morfológica a las construcciones clasificatorias en la lengua de signos española”, *ELUA*, 18, págs. 151–167.
- Herrero Blanco, A. (2009): *Gramática Didáctica de la Lengua de Signos Española (LSE)*. Madrid, SM.
- Huenerfauth, M. (2004): “Spatial Representation of Classifier Predicates for Machine Translation into American Sign Language”. *Actas de Language Resources and Evaluation Conference*, Association pour le Traitement Automatique des Langues.
- Huenerfauth, M. (2006): *Generating American Sign Language classifier predicates for English-to-ASL machine translation*. Computer and Information Science. University of Pennsylvania. Tesis doctoral.
- Huenerfauth, M. (2009): “A Linguistically Motivated Model for Speed and Pausing in Animations of American Sign Language”, *Journal of ACM Transactions on Accessible Computing*, 2, págs. 1–31.
- van der Hulst, H. y A. Mills (1996): “Issues in sign linguistic: Phonetics, phonology and morpho-syntax”, *Lingua*, 98, págs. 3–7.
- Ibáñez, E., A. Huerta, R. San-Segundo, L. F. D’Haro, F. Fernández y R. Barra (2006): “Prototipo de traducción de voz a lengua de signos española”. *Actas de las IV Jornadas en Tecnología del Habla*, págs. 117–122.
- Kennaway, R., J. Glauert y I. Zwitterlood (2007): “Providing signed content on the Internet by synthesized animation”, *Journal of ACM Transactions on Computer-Human Interaction*, 14, págs. 1–29.
- Liddell, S. K. y R. E. Johnson (1989): “American Sign Language: The Phonological Base”, *Sign Language Studies*, 64, págs. 195–278.
- Liddell, S. K. (1996): “Spatial representations in discourse: Comparing spoken and signed language”, *Lingua*, 98, págs. 145–167.
- Liddell, S. K. (2003): “Sources of Meaning in ASL Classifier Predicates”. En K. Emmorey (ed.): *Perspectives on Classifier Constructions in Sign Languages*. Psychology Press, págs. 199–220.
- Lillo-Martin, D. (1995): “The Point of View Predicate in American Sign Languages”. En K. Emmorey y J. Reilly (eds.), *Language, Gesture, and Space*. págs. 155–170.
- López, F., J. Tejedor, D. Bolaños y J. Colás (2006): “Intérprete de Lenguaje de Signos en Español Multidispositivo”. *Actas de IADIS-CIAWI*, págs. 293–296.
- López Colino, F. J. (2009): *Sintetizador Paramétrico Multidispositivo de Lengua de Signos Española [Spanish Sign Language Multidevice Parametric Synthesizer]*. Escuela Politécnica Superior. Universidad Autónoma de Madrid. Tesis doctoral.
- López-Colino, F. y J. Colás (2009a): “Nueva herramienta interactiva para la enseñanza de vocabulario de LSE”. *Actas del III Congreso Nacional de Lengua de Signos Española*.
- López-Colino, F. y J. Colás (2009b): “HLSML, una notación de alto nivel para la descripción de mensajes sintéticos en LSE”. *Actas del III Congreso Nacional de Lengua de Signos Española*.

- Muñoz Baell, I. M. (1999): *¿Cómo se articula la lengua de signos española?* Confederación Nacional de Sordos de España.
- Muñoz Baell, I. M. (2000): “La querología aplicada a la enseñanza de la LSE”, *Revista Española de Lingüística de las Lenguas de Signos*, 2, págs. 1–18.
- Prillwitz, S., R. Leven, H. Zienert, T. Hanke y J. Herming (1989): *HamNoSys. Version 2.0; Hamburg Notation System for Sign Languages. An introductory guide*. Signum-Verlag.
- da Rocha, A. C. y G. Pereira (2004): “Supporting Deaf Sign Languages in Written Form on the Web”. *Actas de Language Resources and Evaluation Conference*, Association pour le Traitement Automatique des Langues, págs. 26–28.
- Rodríguez González, M.A. (1992): “Lenguaje de Signos”. Fundación CNSE.
- San-Segundo, R., J. M. Montero, J. Macías-Guarasa, R. Córdoba, J. Ferreiros y J. M. Pardo (2008a): “Proposing a speech to gesture translation architecture for Spanish deaf people”, *Journal of Visual Languages and Computing*, 19, Academic Press, Inc., págs. 523–538.
- San-Segundo, R., R. Barra, R. Córdoba, L. F. D’Haro, F. Fernández, J. Ferreiros, J. M. Lucas, J. Macas-Guarasa, J. M. Montero y J. M. Pardo (2008b): “Speech to sign language translation system for Spanish”, *Speech Communication*, 50, Elsevier Science Publishers, págs. 1009–1020.
- Sandler, W. (1989): *Phonological Representation of the Sign: Linearity and Nonlinearity in American sign Language*. Foris Publications Holland.
- Stokoe, W. C. (1960): “Sign Language Structure: An Outline of the Visual Communication Systems of the American Deaf”, *Studies in linguistics: Occasional papers*, 8, págs. 1–78.
- Supalla, T. (1978): “Morphology of verbs of motion and location”. *Actas del Second National Symposium on Sign Language Research and Teaching*, págs. 27–45.
- Supalla, T. (2003): “Revisiting Visual Analogy in ASL Classifier Predicates”. En K. Emmorey (ed.): *Perspectives on Classifier Constructions in Sign Languages*. Psychology Press, págs. 249–257.