

Lexicografía de las lenguas de signos y el estado de la lexicografía en la LSC

Gemma Barberà* y Eulàlia Ribera**

*Universitat Pompeu Fabra y **Universitat Autònoma de Barcelona

1. INTRODUCCIÓN

Actualmente el reconocimiento de las lenguas de signos (LS) como lenguas naturales es un hecho incuestionable, pero hace relativamente poco que eso es así. De hecho, no fue hasta la mitad del siglo XX que algunos lingüistas publicaron las primeras obras en que defendían el reconocimiento de las LS basándose en principios lingüísticos (Tervoort 1953, Stokoe 1960). Los fundamentos de estas primeras obras supusieron un cambio de perspectiva que posteriormente chocó con numerosas dificultades para ser aceptadas y generalizadas entre los especialistas. La fuerte tendencia oralista, no solo de los estados, que se han resistido a implementar políticas sociales y educativas basadas en este reconocimiento, sino también la de muchos profesionales o las mismas familias- en la mayoría de los casos, padres oyentes con hijos sordos- hizo que la investigación lingüística de las LS se estancara y estuviera muy atrasada en comparación con los estudios que se llevaron a cabo en las lenguas orales (LO).

Esta situación afectó también los estudios en el campo lexicográfico. La tendencia oralista hizo desestimar el primer trabajo en lexicología del signo de la ASL (lengua de signos americana). Dorothy Casterline, Carl Groneberg y William Stokoe publicaron en 1965 el primer diccionario en LS basado en principios lingüísticos. Las entradas se presentaban a través de un sistema de transcripción desarrollado para la descripción de los signos, que los describían en relación a la localización, la configuración y el movimiento de las manos, es decir, a través de los parámetros fonológicos de la modalidad signada. Como consecuencia del contexto social, cultural y político del momento, así como consecuencia de su formato, que lo hacía poco accesible a las personas no lingüistas por el número limitado de ilustraciones y por las anotaciones de las entradas, esta vía de explotación no tuvo éxito y no se aprovechó para la configuración de otros diccionarios. La mayoría de diccionarios en LS que se publicaron durante éste período y hasta prácticamente los años 80 consistían en listados de signos utilizados como equivalentes de palabras de una LO,

ordenados alfabéticamente conforme con los principios lingüísticos de aquella lengua (es decir, por orden alfabético). Además, normalmente no había ninguna referencia en estas publicaciones sobre el significado del signo, porque se sobreentendía que era el mismo que el de la traducción de la LO que acompañaba cada signo. Y es que el principal objetivo del diccionario consistía en dar a los padres oyentes y a los profesores un libro de referencia que les facilitara la comunicación con los niños sordos, y por lo tanto, confeccionaban listas de palabras y se identificaba un signo con aquellas palabras.

Este sistema no describía el componente léxico característico y propio de las LS, es decir, omitía la matriz semántica, morfológica, sintáctica y pragmática propia de cada signo, con lo cual dejaba de lado gran parte de la información que el hablante tiene en la mente como parte de su conocimiento sobre un signo.

Por ejemplo, en el caso de la elaboración del diccionario de la BSL (lengua de signos británica), Brien explica en la introducción que al principio de su confección, antes de que el primer responsable de la edición, Allan Hayhurst -fallecido en el año 1981- leyera el trabajo de Stokoe hacia el 1975 e introdujera modificaciones importantes en la perspectiva y elaboración del diccionario, cuando se consideraba que no había ninguna correspondencia entre la LO y el signo, o bien se eliminaba el signo o bien se creaba un signo nuevo a partir del Gestuno¹

Por otra parte, se utilizaban fotografías para ilustrar los signos y se añadían dibujos para indicar los movimientos involucrados. Cada fotografía iba acompañada de una descripción escrita de cómo se realizaba, juntamente con las palabras inglesas más comunes asociadas con el signo.

En definitiva, no fue hasta los años 80 que empezaron a verse publicaciones de diccionarios de LS organizados de acuerdo con los principios similares a los defendidos por Stokoe y sus colegas veinte años atrás. Las nuevas tecnologías han ayudado mucho, y gracias a ellas actualmente en LS podemos encontrar diccionarios editados y recursos lexicográficos multimedia, aunque la investigación y la variedad de materiales lexicográficos que se han hecho en esta modalidad queda muy lejos de la que podemos encontrar en la modalidad oral.

A continuación veremos las dificultades con las que chocan los lexicógrafos cuando quieren elaborar cualquier material lexicográfico de una lengua. Empezaremos exponiendo

¹ Compilación de signos internacionales publicado por la Federación Mundial de Sordos (1975) para el contacto entre signantes de diferentes LS (Brien 1992)

las dificultades comunes en todas las lenguas naturales; veremos las características propias de la modalidad visual, y posteriormente hablaremos de las características de las comunidades sordas. En el segundo apartado trataremos los materiales lexicográficos que se han editado en LS y específicamente en LSC. Finalmente, en el tercer apartado expondremos una propuesta de base de datos (BD de ahora en adelante), con una muestra de la información que esta BD podría recoger para confeccionar materiales lingüísticos tan importantes como un diccionario normativo descriptivo de la LSC.

1.1 Características específicas de las lenguas de signos

Describir el componente léxico de una lengua no es fácil. De entrada el lexicógrafo tiene que tener en cuenta los siguientes aspectos:

- La selección de las unidades léxicas. En una lengua natural viva coexisten diferentes variantes. No pasa desapercibido a ningún hablante nativo que dentro del dominio de una lengua natural hay variantes geográficas, sociales o de registro muy acentuadas que pueden afectar la ejecución de un signo, su sintaxis o la semántica. Dentro de cualquier proyecto lexicográfico se tendrá que definir qué variantes se recogen y cómo se identificarán estas formas dentro de la BD, además de decidir qué unidades se introducirán y cuales no.
- La representación abstracta que figurará como lema. Una unidad léxica de una lengua puede adoptar diferentes formas en función del contexto sintáctico, la categoría morfológica, la semántica u otros aspectos. Estas variaciones no modifican el significado del signo, y el lexicógrafo tendrá que encontrar una forma que las represente a todas. Esta representación abstracta recibe el nombre de “lema”.
- Las entradas y las subentradas. Este apartado implica el tratamiento de la homonimia y la polisemia. Cuando dos signos tienen la misma forma pero un significado totalmente independiente y diferente el uno del otro, se dice que son palabras homónimas. En cambio, si un signo tiene diferentes significados relacionados entre ellos decimos que se trata de un signo polisémico. El lexicógrafo tendrá que decidir a qué signos les da el tratamiento de homónimos y a cuales el de polisémicos.

- La microestructura. Las unidades léxicas de una lengua tienen muchísima información asociada. Esta información es la que permite que la unidad pueda ser insertada en una combinación de palabras, o bien que esto no se pueda hacer, ya sea porqué no pertenece a la categoría o al registro adecuado, etc. Dependiendo de los objetivos y de los destinatarios del material lexicográfico que se tenga que elaborar, el lexicógrafo tendrá que escoger la información que recogerá.

Además de estos aspectos comunes que hay que tener en cuenta en cualquier trabajo lexicográfico, sea de la modalidad que sea, también hay que tener en cuenta las diferencias que encontramos en las LS y las LO, que son las siguientes:

- La forma de la señal que emiten. Las LO emiten una señal acústica, mientras que las LS emiten una señal óptica. Eso significa que los canales de emisión y recepción de la lengua son diferentes en una y otra modalidad.
- Los parámetros fonológicos y su interpretación fonética. Como utilizan órganos articulatorios diferentes, la interpretación fonológica y los parámetros fonológicos son diferentes. En el caso de la modalidad oral, la interpretación fonética se basa y se materializa en los órganos articulatorios vocales, mientras que en el caso de la modalidad visual, se utilizan las manos, la cara, la cabeza y el torso.
- La simultaneidad vs la secuencialidad. Aunque ambas modalidades tienen componentes fonológicos que se realizan simultáneamente y componentes que se realizan secuencialmente, la modalidad visual tiene un porcentaje de simultaneidad mayor que el que encontramos en las LO.
- El uso del espacio. La modalidad visual permite la expresión en tres dimensiones, y eso comporta un uso más complejo y preciso del espacio. Además, la ubicación variable de los signos respecto al cuerpo del signante puede comportar cambios de significado. Por ejemplo, el signo OPERAR se puede articular en distintas partes del cuerpo según cual sea el órgano que ha sufrido la operación.
- La iconicidad. El nivel de iconicidad es mayor en las LS que en las LO. De hecho, la encontramos también a un nivel subléxico, ya que los parámetros formativos pueden asociarse a un significado determinado. Un ejemplo muy claro es la asociación entre la localización cerca de la frente y ciertos movimientos léxicos relacionados con los procesos mentales. Así, signos como PENSAR, IMAGINAR,

IDEA o REFLEXIONAR se articulan cerca de la frente o bien inician allí el movimiento.

Cuando el objetivo es elaborar un material lexicográfico en la modalidad visual, hay que tener en cuenta estas diferencias y buscar una manera de recoger y presentar los datos de forma adecuada a la modalidad. Algunas dificultades con las que se han encontrado los lexicógrafos que han empezado a trabajar con LS son las siguientes:

- La segmentación de las unidades léxicas. Cuando un enunciado entero se realiza mediante un solo signo, a veces cuesta saber cuantas unidades léxicas tiene aquel signo.
- La ausencia de código escrito. En las LS no hay código escrito. Esta es una dificultad añadida, ya que la escritura es fuente y recurso principal para la elaboración de este tipo de materiales en las LO. El material escrito ayuda a segmentar las palabras, y a tratar informáticamente los datos, de forma que, aunque a veces una unidad léxica puede consistir en un grupo de palabras- lo que da no pocas dificultades a los lexicógrafos- la linealidad y la organización de la escritura ofrece una facilidad que en las LS no encontramos.
- La representación de las unidades léxicas. Habitualmente se hace mediante la glosa, que es la simplificación del significado del signo a una palabra de la LO más cercana. Es necesario buscar nuevas formas de representación más adecuadas a la modalidad visual, como la imagen. Un dibujo o una fotografía puede representar el signo más fielmente que una glosa, por ejemplo.
- La descripción de las unidades léxicas. En la modalidad oral las definiciones se ofrecen en formato texto. Las LS no tienen código escrito, con lo cual, no se puede ofrecer la información en este formato si no es haciendo uso de la LO más cercana. Será necesario utilizar la escritura para hacer posible el procesamiento de los datos por parte de la BD, pero es evidente que hay que buscar nuevas formas de presentación al usuario más adecuadas a la modalidad visual. Una posible solución es hacer las descripciones en formato vídeo.
- La búsqueda de las unidades léxicas. Actualmente el sistema de búsqueda más extendido es el alfabético. Este sistema es el más utilizado también entre las LO indo-europeas. Pero en el caso de la modalidad visual, este sistema la vuelve a

hacer dependiente de la glosa, y en consecuencia, de la LO más cercana. Es necesario buscar un sistema alternativo más adecuado a la modalidad visual, como por ejemplo, el sistema que presentaremos en este artículo, basado en los parámetros fonológicos de los signos.

La juventud de los trabajos lexicográficos en LS hace que muchos aspectos requieran seguir siendo explorados. Se arrastra el voluminoso bagaje de las LO, con su sistema de organización y presentación, y eso hace que resulte muy difícil encontrar sistemas más adecuados para la compilación lexicográfica de la modalidad visual. La introducción de las nuevas tecnologías ha significado un salto muy importante en ese tipo de trabajos porque, tal como veremos en los apartados siguientes, permiten obtener unos criterios visuales para recoger y organizar los datos.

1.2 Particularidades de la comunidad sorda

Entendemos por *Comunidad Sorda* aquella “comunidad de habla signante que incluye un conjunto de individuos que comparten una lengua y unos valores sociolingüísticos comunes, como reglas de uso y actitudes lingüísticas, así como una ideología y una actitud hacia estos usos lingüísticos” (Gras 2006: 5). Aún así, la edad de la primera exposición de la lengua, la lengua materna de casa, la variedad social y regional y la escolarización son elementos que contribuyen a dar una gran variedad de color y heterogeneidad a esta comunidad.

Un aspecto común de las comunidades sordas es que la gran mayoría de sus integrantes entraron en contacto con la lengua durante la escolarización, puesto que gran parte de los niños sordos proceden de familias oyentes (se calcula que entre el 90-95%). Según los estudios estadísticos de Gras 2006, a nivel de todo el Estado Español el 73% de los informantes que sirvieron para elaborar las estadísticas de su trabajo entraron en contacto con la LS en la etapa infantil, mientras que sólo el 9,6% lo hicieron en casa y por lo tanto eran nativos de la LS. Además cabe destacar que estamos ante una lengua que estuvo prohibida durante muchos años y relegada, por lo tanto, al uso familiar. Las distancias territoriales y las políticas educativas aplicadas en distintas épocas también han hecho que la lengua evolucionara de modo muy distinto en cada región y que nos encontremos con una variedad regional muy poco homogénea.

En el momento de iniciar cualquier estudio lingüístico de esta lengua hay que tener en cuenta los aspectos que acabamos de mencionar. Conocedores de este hecho y conscientes de que la lengua es un mosaico representativo de toda una región, los equipos

de investigación vinculados a la LS han intentado llevar a cabo distintas medidas. Un caso muy claro es la solución efectiva a la que han llegado desde el Instituto de lengua de signos alemana (DGS) y la comunicación de los sordos, de la Universidad de Hamburgo². Este equipo trabaja desde hace 15 años en un proyecto de creación de corpus que ha permitido desarrollar un diccionario electrónico basado en este corpus, una colección de vocabulario básico y un diccionario de gramática. Conscientes de la realidad heterogénea de la lengua, han creado un grupo de expertos formado por unas 20 personas sordas que trabajan como asesores de la lengua. En caso de dudas sobre el estatus léxico de un signo, cuestiones de cambio en la lengua, diferentes significados o información regional de los signos, acuden a este grupo de expertos, que ha sido previamente formado en cuestiones lingüísticas y que son representantes directos de la comunidad sorda de toda Alemania. Así, cuestiones relacionadas con la selección de lemas y con la buena formación de construcciones gramaticales son constantemente validadas por parte de este grupo (Prillwitz *et al.* 2008). A veces también actúan de contacto para obtener datos de otras personas sordas, en el caso de que sea necesario. Así y todo, al tratarse de un grupo de expertos reducido han establecido un sistema de votación público que permite que la comunidad lingüística alemana en general participe directamente en la construcción del proyecto. De este modo, las personas que lo deseen pueden votar y aportar variantes regionales que en un primer momento no aparecían en el corpus de datos.

Cualquier iniciativa de creación de material, de estudios lexicográficos o de proyectos de corpus de las LS tiene que trabajar desde esta perspectiva, haciendo que la comunidad lingüística participe activamente en el proceso de realización de los proyectos de la lengua, obteniendo así una muestra de la lengua lo más representativa posible.

2. MATERIAL LEXICOGRÁFICO EN LENGUA DE SIGNOS

Una base de datos (BD) lexicográfica es una recopilación de lemas de una lengua natural. Las BD lexicográficas creadas con la ayuda de la informática amplían enormemente las posibilidades de almacenar información sobre el léxico, y la agilidad en el momento de hacer búsquedas permite que pueda utilizarse no sólo para los usos habituales de los diccionarios, sino también para hacer búsquedas lingüísticas sobre aspectos concretos de la lengua estudiada, como por ejemplo sobre variación sociolingüística, fonología u otros aspectos sintácticos o semánticos. Además, el avance en las nuevas tecnologías ha beneficiado mucho los materiales en LS, ya que ha permitido visualizar la lengua e integrar el aspecto visual en la presentación de las entradas.

² <http://www.sign-lang.uni-hamburg.de/>

Dentro de los recursos lexicográficos es importante hacer la distinción entre lo que entendemos por *diccionario* y por *base de datos lexicográfica* (BD), porque a veces la frontera entre ambos cuesta de distinguir. Los diccionarios son el producto final que llega al usuario para buscar información sobre una lengua natural, mientras que las BD son aquellas recopilaciones de todo el material lexicográfico que hace un lexicógrafo para permitir que cuando se elabore un diccionario ofrezca las mejores garantías de búsqueda para el usuario. Ello implica la inclusión de un número de campos mucho más elevado de los que después se verán en el producto final, como por ejemplo, la vinculación con el corpus y los informantes, la transcripción con sistemas de anotación conocidos solo por el círculo de investigadores, o bien la inclusión de campos con etiquetas semánticas que permitan la búsqueda por esta vía. De hecho, los recursos multimedia permiten almacenar mucha más información que los diccionarios impresos, y esto hace que el dominio lexicográfico sea mucho más amplio.

La mayoría de diccionarios publicados desde 1965 no han sido sobre los signos utilizados en las LS, sino sobre signos utilizados como suplementos de las palabras de una LO, ordenados de acuerdo con los principios de la LO. Normalmente en estas publicaciones no hay ninguna referencia al significado de los signos, sino que se da por supuesto que éste es el mismo que el de la traducción de la LO que le acompaña. La motivación de esta dependencia hacia la LO, como ya hemos comentado antes, procede del hecho que en ese momento el principal objetivo del diccionario era dar a los padres oyentes y a los profesores un libro de referencia que les facilitara la comunicación con los niños sordos.

A pesar de las divergencias en la organización de la información de los distintos diccionarios, los diccionarios impresos tienen en común que la información que aparece sobre el signo es básicamente sobre la manera como se ejecuta y aparece también la glosa o palabra equivalente en la correspondiente LO.

2.1 Diccionarios impresos de lenguas de signos

La variedad de diccionarios en LS no es tan espectacular como en el caso de las LO, pero también encontramos diferencias en la organización de los signos. En este apartado presentamos brevemente dos diccionarios impresos de LS que son importantes por la presentación de las entradas léxicas.

2.1.1 *Dictionary of British Sign Language / English* (1992) de David Brien y Mary Brennan

En esta obra las entradas se han organizado desde una perspectiva semasiológica y se ha utilizado un sistema visual de clasificación utilizando las partes del signo transcritas a partir de un sistema de transcripción propio ampliamente conocido entre los investigadores del grupo. Según explica Brien en la introducción del diccionario, los signos que se recogen se agrupan a partir de los parámetros fonológicos. Esto quiere decir que encontramos distintos grupos de símbolos clasificados según:

- (i) la configuración de la mano
- (ii) la localización del signo
- (iii) la orientación de la mano y de los dedos
- (iv) el movimiento de las manos

Cada uno de estos grandes grupos se divide en subgrupos que ayudan en la búsqueda de los signos, aunque para acceder fácilmente es necesario conocer el sistema de transcripción descrito en las páginas precedentes de la obra, y decidir primero si el signo que se quiere buscar es de configuración simple o compleja (es decir, si se hace con una sola mano o con las dos), y a partir de entonces escoger los subgrupos de signos según sus características fonológicas. Las entradas constan de la descripción de la ejecución del signo y de la traducción en inglés.

2.1.2 *La langue des signes, dictionnaire bilingue LSF / français (1997) de M. Girod et al.*

Esta obra organiza las entradas onomasiológicamente, esto es, por temas y subtemas que se encuentran listados en un índice. Este es el criterio utilizado en la mayoría de diccionarios de otras LS. Según las explicaciones del editor en la introducción, se ha hecho así porque el diccionario iba dirigido a estudiantes y consideraban importante que pudieran ver diferencias fonológicas sutiles entre signos del mismo campo semántico. Tampoco aparece una definición del significado del signo, sino que mediante dibujos se describe la ejecución y se pone el equivalente glosado en francés. Por otro lado, también contiene un índice de clasificación de los signos según la configuración, pero para encontrar un signo primero hay que ver cuál es la configuración principal del signo mediante la traducción; en segundo lugar, ver la localización (si es en la cabeza, en el cuerpo o delante del torso), y finalmente mirar el listado alfabético de las traducciones, donde aparece la numeración de la página y el número de viñeta dentro de la página.

2.2 Diccionarios impresos en lengua de signos catalana

Aquí presentamos los diccionarios impresos en LSC que existen actualmente.

2.2.1 *Lenguaje de Signos Manuales (1987) de Jorge Perelló y Juan Frigola*

En su segunda edición, del 1998, esta obra pionera consta de 4.500 signos, e incluye signos de las distintas comarcas catalanas. Está organizado temáticamente y cada entrada consta del dibujo del signo, más la glosa en castellano y catalán, y un código de ejecución del signo, cuya interpretación podemos encontrar en los primeros capítulos. Otra característica de este diccionario es que tiene capítulos dedicados a ciertos aspectos gramaticales, como el artículo, los sustantivos, los verbos, los adverbios, los adjetivos, los pronombres, las preposiciones, las conjunciones, las interjecciones y la construcción de frases. En el prólogo se especifica que el objetivo no es ser un libro de gramática, sino un libro de refuerzo para los aprendices de la LS, y por esta razón ofrece algunas directrices prácticas y muy generales de algunos aspectos de esta lengua.

2.2.2 *Diccionari Temàtic Bàsic de la llengua de signes catalana de Ponent (1996) de Ignasi Martín y Marisa Alvarado*

Este diccionario, con 3900 entradas en su segunda edición, está organizado temáticamente, de modo muy parecido al de LSF comentado anteriormente, pero sin índice por configuración u otros parámetros fonológicos del signo. Las entradas están presentadas a partir de unos dibujos sencillos pero claros, donde el movimiento aparece en un segundo icono del dibujo. Su objetivo es proporcionar recursos y documentación ordenada para complementar el aprendizaje de la LS, por lo que contiene un índice alfabético de términos en lengua catalana y otro en lengua castellana. La segunda edición de esta obra se hizo en el 2004 y se publicó bajo el nombre de *Diccionari Temàtic de Llenguatge de Signes Català*.

2.3 Recursos multimedia en lenguas de signos

Uno de los desarrollos más importantes de los recursos multimedia aplicados a la LS es el increíble aumento de las capacidades de almacenamiento de los ordenadores, así como el desarrollo de la tecnología del vídeo digital, hecho que ha permitido relacionar un vídeo con cada entrada. Sin embargo, esto también ha comportado que los sistemas de transcripción de las LS no sean tan relevantes para la identificación de los signos, ya que el vídeo permite visualizarlos en 3D. En este sentido, no significa que podamos prescindir de ellos, puesto que son necesarios por ejemplo para los sistemas de traducción, sino que el uso del vídeo ha provocado que los sistemas de transcripción y anotación sean menos relevantes en la presentación de las entradas (Johnston 2001).

Actualmente, el hecho de que puedan utilizarse canales multimedia hace que tanto los diccionarios como las BD lexicográficas se adapten perfectamente a la modalidad visual de las LS y que los criterios lingüísticos aplicables puedan ser mucho más cercanos a esta modalidad. En los últimos años han surgido varios grupos de investigación dedicados al trabajo lexicográfico en distintas LS que trabajan con BD. Algunos han presentado productos finales en formato CD-ROM y DVD, y otros se encuentran en la red telemática, beneficiándose así de la actualización inmediata que permite esta vía.

2.3.1. Bases de datos de lenguas de signos

2.3.1.1 BD del Auslan (lengua de signos australiana)

La base de datos de la lengua de signos australiana (Auslan) es un ejemplo de trabajo lexicográfico que se está haciendo desde hace tiempo para una LS concreta. Incluye información fonológica, gramatical, semántica, así como las glosas y las definiciones de los signos. La información visual se presenta con gráficos, vídeos y transcripciones (Johnston 2001). El fruto final de esta BD es el material en línea que podemos consultar en <http://www.auslan.org.au/>. Aquí podemos hacer búsquedas a partir de la traducción en LO (en este caso, en inglés) y aparece el signo, con las variantes introducidas, la definición, la categoría gramatical y la zona de uso del signo. Además, la página incluye unos apartados donde los usuarios pueden proponer signos que no aparezcan y dar *feedback* sobre los signos ya introducidos.

2.3.1.2 BD de la ÖGS (lengua de signos austríaca)

Recientemente se ha hecho pública en la web la BD de la ÖGS. Desarrollada en la Universidad de Klagenfurt, se puede consultar en línea en <http://ledasila.uni-klu.ac.at/>. Esta BD tiene dos particularidades. Por un lado, es una plataforma abierta donde es posible introducir nuevas categorías o valores, siempre y cuando se tenga el estatus de administrador. Ha sido diseñada de modo que no haya decisiones fijas desde el inicio del proyecto que impidan la introducción de nuevas categorías a medida que la investigación avanza (Krammer *et al.* 2001). Además, no es específica para una LS en particular y puede servir como plataforma para documentar otras LS, a través de la aplicación web LedaSila (*Lexical Database for Sign Languages*). Actualmente la página web puede consultarse en inglés y alemán, pero el equipo está trabajando para incluir la ventana de ayuda en ÖGS y en sistema de signos internacional. Las búsquedas dentro de la BD pueden hacerse a partir de las siguientes opciones: semántica, región, tipo de signo, lugar de articulación, configuración y ámbito. Hasta la fecha, LedaSila solo ha sido utilizada para

la descripción de la ÖGS y contiene 14.000 vídeos de signos, incluidas las variantes regionales (Krammer *et al.* 2008).

2.3.1.3 *BD de la NGT (lengua de signos holandesa)*

La BD de la NGT es uno de los proyectos que se desarrollan dentro del Centro para la lengua de signos holandesa (NGc), fundado en el año 1992. Desde el 2006 este centro se ha convertido en el instituto nacional de lexicografía de la NGT y ha desarrollado proyectos sobre materiales de aprendizaje de la lengua, cursos educativos, una base de datos y diccionarios en DVD y en línea. La BD en línea

(<http://www.gebarententrum.nl/gebarenwoordenboek.asp>) contiene actualmente unos 14.000 signos, y ha sido la herramienta de uso durante la recogida de signos para el proyecto de planificación de la lengua.

2.3.1.4 *BD de la DSGS (lengua de signos suizo germana)*

Esta BD multimedia ha servido como base para posteriores diccionarios y materiales de enseñanza, y también como herramienta para la investigación lingüística. Utiliza vídeos de *demo*, fotogramas del vídeo, dibujos, anotación fonética en HamNoSys y una pequeña parte en SignoEscritura para la representación gráfica de los signos. Además incluye información de variantes generacionales y de la etimología que se conoce (Boyes Braem 2001).

2.3.1.5 *SignPhon*

En este último subapartado mencionamos una base de datos con objetivos específicos. Se trata de SignPhon, una BD fonética y fonológica diseñada para cualquier LS. Como cualquier otra BD, SignPhon fue diseñada para encontrar patrones, regularidades distribucionales y frecuencias de un dominio concreto, con el objetivo de utilizar esta información para construir una teoría. En este caso, el objetivo era claramente realizar estudios fonéticos y fonológicos de las LS y utilizar los datos del SignPhon para comprobar ciertas hipótesis, o bien desarrollar otras. En todos los inicios de las construcciones de BD no se tiene un modelo comprensivo del dominio. Al empezar a codificar los primeros datos es necesario saber para qué aspecto será útil la BD, esto es, haber determinado ya la categorización relevante de los datos. Y por lo tanto el sistema de codificación ideal es, de hecho, la mejor teoría del dominio. Nos encontramos delante de lo que Crasborn *et al.* 2001 llaman “la paradoja de la base de datos”, y es que parece que cualquier sistema de codificación será una aproximación del primer ideal que se tenga.

2.3.2 Aplicaciones de las bases de datos

La aplicación más usual de cualquier BD es la creación de un diccionario de la lengua. Pero, como ya hemos comentado anteriormente, este diccionario sólo puede incluir una parte de todo el conocimiento de la lengua contenido en la BD.

2.3.2.1 *Diccionario de la DSL (lengua de signos danesa)*

Esta obra contiene unos 2.000 signos de la DSL, con ejemplos y referencias a sinónimos y homónimos. Actualmente está accesible desde www.tegnsprog.dk y podemos hacer búsquedas de la DSL a partir de las características manuales de los signos o bien a partir de la palabra en danés oral. En esta obra los criterios que se han seguido para la selección de las entradas (lematización) son fonológicos y semánticos. Así, dos signos que difieren en significado se convierten en dos entradas distintas y pasa lo mismo con signos que tienen algún parámetro que les diferencie. El problema aparece cuando se tiene que establecer la frontera entre entradas que difieran en rasgos fonológicos y variantes del mismo signo. En este material se tratan como variantes los signos que se diferencian por no más de un parámetro. Los signos que varían en dos parámetros son tratados como sinónimos y tienen entradas diferentes (Troelsgard y Kristoffersen 2008).

2.3.2.2 *Diccionario de la VGT (lengua de signos flamenca)*

Este material lo podemos encontrar en línea en <http://gebaren.ugent.be/>. Desde este enlace podemos hacer búsquedas a partir de la palabra en flamenco y, al mismo tiempo que visualizamos el vídeo con el signo, podemos ver el equivalente en SignoEscritura. De todos modos, desde <http://gebaren.ugent.be/visueelzoeken.php> podemos buscar signos por configuración y punto de contacto.

2.3.2.3 *El proyecto bilingüe SignBase*

Este proyecto tuvo dos aplicaciones muy concretas. SignBase es una obra que presenta la información en LO y LS, y contiene información morfé mica que incluye la unidad más pequeña de significado de cada signo (es decir, tiene en cuenta el léxico productivo de las LS, como los clasificadores). Además, las entradas se presentan con fotogramas, vídeos e incluyen la definición del término (Schermer *et al.* 2001). Este proyecto dio a luz dos aplicaciones: el Diccionario BSL/inglés de la cultura y la comunidad sorda, y el CD-ROM de la naturaleza y el medio en NGT/holandés (material de educación bilingüe). Por un lado, en el Diccionario BSL se puede buscar a partir de la palabra oral, o por el signo (a partir de la configuración), y también a través de información gramatical, esto es, por categorías léxicas. Por el otro, el CD-ROM de la naturaleza y el medio en NGT es, en cambio, una aplicación más específica que presenta conceptos y explicaciones de los signos. No incluye una descripción gramatical explícita, pero sí que la presenta indirectamente con los

ejemplos, ya que el signo BEBER, por ejemplo, varía de configuración según si el sujeto es un animal u otro.

2.3.2.4 *Diccionario de la FinSL (lengua de signes finlandesa)*

Este diccionario, como ha ocurrido en otros casos, fue el resultado del trabajo lexicográfico realizado en el desarrollo de la BD de la lengua de signes finlandesa (Savolainen 2001).

2.3.3 Aplicaciones lexicográficas y proyectos de corpus

En los últimos años la creencia que los trabajos lexicográficos tienen que basarse en proyectos de corpus se ha extendido ampliamente. Además, la selección de las entradas de un corpus requiere la construcción paralela de un léxico. Por esta razón, algunos de los trabajos lexicográficos que se hacen en distintas LS van unidos a la recopilación de datos y posterior anotación, puesto que ambos trabajos se enriquecen entre ellos. La ventaja de disponer de un corpus es que las decisiones para la identificación y selección de los signos pueden atribuirse a los datos. La desventaja es que un corpus es sólo un ejemplo limitado del uso de la lengua, que a veces provoca que se excluyan otros signos apropiados (Konrad *et al.* en prensa).

2.3.3.1 *Diccionarios de términos especializados*

En la Universidad de Hamburgo se han desarrollado unos diccionarios de terminología especializada de temática diversa, que va desde la terminología computacional, la psicología, la carpintería, la economía doméstica hasta el trabajo social y la salud. Estas recopilaciones de términos técnicos pueden consultarse libremente en la web:

<http://www.sign-lang.uni-hamburg.de/Projectseng/SLDicts.html>

Una característica importante de la LS es que los signos manuales interactúan con la vocalización para expresar el significado léxico del elemento. Con vocalización nos referimos al “componente hablado”, que es el movimiento de la boca que forma parte del signo y que tiene su origen en el gesto de la pronunciación de la palabra en LO (Quer *et al.* 2005). Este componente hablado es de gran importancia en los ámbitos de vocabularios especializados, ya que para las LS la situación es especial. Generalmente, podemos encontrar pocos sordos (o ninguno) en los ámbitos más especializados, y hay sólo unos contextos determinados y reducidos donde los sujetos sordos utilicen los signos especializados. A menudo, pues, la realidad es que estos pocos profesionales sordos utilizan la LO escrita para la comunicación y la documentación de estos términos. Por este

motivo, en la terminología especializada en signos encontramos una gran influencia de la LO que aparece reflejada en el componente hablado del signo.

Llegados a este punto es importante distinguir la introducción de las entradas a partir de *tokens* y *types*. Los *types* se refieren a la forma de citación de los términos. Los *tokens* son la variación o la diferente realización del signo en un contexto determinado (König *et al.* 2008). En cuanto a los términos especializados, la LS a menudo utiliza el mismo signo manual con un componente hablado distinto. En LSC, por ejemplo, el signo manual MENTIRA puede acompañarse de la oralización de la palabra “virtual”, cambiando así el significado del término. Nos encontramos delante de dos instancias del mismo signo, ya que la variante del significado “virtual” es un *token* del mismo *type*. Identificar estas dos formas con dos lexemas diferenciados sería engañoso, porque se trata de la misma forma manual combinada con la palabra de la LO.

2.3.3.2 *El iLex, la herramienta que enlaza corpus y BD*

El iLex³ es una herramienta destinada al trabajo lexicográfico y de corpus, que permite trabajar paralelamente en la construcción de un corpus y de un léxico relacionados. Con el precedente del GlossLexer, una BD léxica, y el SyncWriter, los mismos autores de la Universidad de Hamburgo han creado el iLex (nombre que proviene de “léxico integrado”). Se trata de una herramienta que modela los procesos de transcripción de los corpus que garantiza el enlace entre *token* y *type* (Hanke y Storz 2008). Las transcripciones consisten en *tokens*, es decir signos, que están referidos a sus tipos respectivos, hecho que tiene una relación directa con el proceso de lematización. Este hecho permite crear corpus con muchos anotadores detrás y garantiza la coherencia intra e inter anotadores. El iLex no sólo se utiliza para trabajos lexicográficos y de análisis del discurso, sino también para generar secuencias de transcripciones de la BD a través de un avatar signante.

2.3.3.3 *Otros proyectos de corpus hechos con herramientas específicas*

El MediaTagger es una herramienta de anotación multimodal. Permite observar las ocurrencias de un signo concreto dentro del contexto. Ha sido utilizado para el estudio del ISN (Idioma de Señas Nicaragüense). El objetivo era analizar y comparar enunciados emitidos por signantes de las dos cohortes (grupos de personas que entraron en la escuela en momentos diferentes). La herramienta les ha servido para la anotación y posterior análisis entre las descripciones que hacían los dos grupos (Senghas 2001).

³ <http://www.sign-lang.uni-hamburg.de/ilex/>

El SignStream⁴ forma parte del Proyecto de investigación lingüística en ASL (lengua de signos americana) de la Universidad de Boston. Se trata de una aplicación que proporciona un entorno único donde se puede ver, analizar y hacer búsquedas de los datos anotados, muy útil para la investigación lingüística en LS y para el componente gestual de la LO. Permite hacer búsquedas complejas mediante unas ventanas que muestran los campos y valores disponibles en el archivo seleccionado. Esta aplicación ha permitido la creación de un corpus anotado en ASL (Neidle 2001).

El BTS (Berkeley Transcription System) es una herramienta que transcribe los elementos a nivel morfémico y que se aplicó a un proyecto sobre la adquisición de las LS. El sistema se desarrolló utilizando datos de interacciones adultos-niños en ASL y NGT (Slobin *et al.* 2001).

En último lugar, el ELAN⁵ es una herramienta de anotación de código abierto que funciona con distintas plataformas y que ha sido utilizada en el corpus ECHO⁶, de acceso público y multilingüe, que tiene la voluntad de establecer estándares de anotación y facilitar el intercambio de datos.

Finalmente, un último punto que nos gustaría destacar es el protagonismo que toma la parte gestual de la lengua cuando trabajamos con corpus anotado. A menudo en un discurso signado encontramos ítems propiamente lingüísticos y otros que podemos catalogar de gesto-mímica, muy frecuentes en datos espontáneos. Cuando trabajamos con una BD que está relacionada a un corpus, nos encontramos con la dificultad de tener que diferenciar de algún modo todo lo que es lingüístico de lo que no lo es. Evidentemente, el lexicógrafo trabaja con elementos lingüísticos, pero a veces es difícil obviar una parte de la lengua difícil de catalogar, pero que por reducida que sea, existe. Johnston 2008 propone transcribir estos signos a partir de la traducción que haríamos. Esto es, escribir “g: *que-bobo-soy*”, en lugar de “g: *picar-la-palma-en-la-frente*”. La forma del gesto no tiene que quedar plasmada en la tira de las glosas del signo, ya que ésta puede apreciarse en el vídeo relacionado o bien en la tira dedicada a los parámetros fonológicos. Evidentemente, aunque esta parte de la lengua no estaría incluida en la BD, sí que se tendría constancia de ella en el corpus.

⁴ <http://www.bu.edu/asllrp/signstream/index.html>

⁵ <http://www.lat-mpi.eu/tools/elan/>

⁶ <http://www.let.kun.nl/sign-lang/echo/>

2.3.4 Recursos multimedia de la Lsc

2.3.4.1 *Dilscat, el diccionari bàsic de la llengua de signes catalana*

Un ejemplo claro de las posibilidades técnicas de las nuevas tecnologías aplicadas a la LSC que tenemos de momento es el *Dilscat, el diccionari bàsic de la llengua de signes catalana*, editado por Illescat el 2004. Este diccionario, que contiene 1.500 vídeos, se presenta en formato CD-ROM y esto permite apreciar las características intrínsecas propias de la modalidad de la lengua. Este diccionario permite dos tipos de búsquedas: temática y alfabética (a partir de la glosa), y se presenta en catalán, castellano, inglés y portugués.

2.3.4.2 *Mira què dic*

*Mira què dic*⁷ es un diccionario básico en LSC y disponible en internet realizado por Montse Quijo, Montserrat Puigdurí y Pilar Viana. Ha sido elaborado por el centro CEE Les Aigües, de Mataró, con el objetivo de dotar de una herramienta lingüística los profesores y padres de alumnos con discapacidad y mejorar la comunicación de personas con dificultades. Contiene un número muy limitado de palabras (unas 500), las justas para poderse comunicar con este tipo de alumnos, y propone dos maneras de efectuar la búsqueda: mediante un índice alfabético de la glosa, y mediante un índice temático. Este material contiene vídeos que muestran cómo se hace el signo, y también dibujos que identifican los signos, extraídos del *Diccionari Temàtic en Llengua de Signes Catalana*.

2.3.4.3 *Terminología en LSC del ámbito de lengua*

Desde el Departament d'Ensenyament de la Generalitat de Catalunya, un grupo de docentes vinculados a programas de enseñanza bilingüe creó en el 2006 un glosario⁸ consultable en línea de terminología específica de signos relacionados con el área de lengua, con la finalidad de estandarizar los signos para potenciar la enseñanza y el aprendizaje de la lengua catalana y castellana a través de la LSC. Este glosario, coordinado por Marta Vinardell y Marga Cirera, se creó a partir de unos talleres con distintos perfiles de profesionales que trabajan a partir de un léxico ya analizado de términos relacionados con la lengua y el lenguaje, la comunicación, la lingüística, las partes de la oración y del texto y las tipologías textuales. El resultado fue un CD-ROM que se repartió entre el profesorado que trabaja con alumnado sordo desde la vertiente educativa bilingüe (LSC y lengua catalana / castellana).

⁷ <http://www.xtec.net/dnee/act/miraquedic/>

⁸ http://www.xtec.net/dnee/lsc/terminologia_llengua/index.htm

3. PROPUESTA DE BASE DE DATOS PARA LA LSC

La propuesta de BD que presentamos a continuación fue presentada como trabajo de investigación dentro del curso de doctorado interuniversitario de *Ciencia Cognitiva y Lenguaje*, y fue dirigido y tutorizado por Josep Quer y M. Teresa Espinal respectivamente (Ribera 2007). El proyecto consiste en el diseño de una BD que permita almacenar y consultar información lingüística sobre las unidades léxicas de la LSC y tiene tres objetivos básicos:

- (i) Crear una herramienta de consulta sobre las unidades léxicas de la LSC para cualquier usuario.
- (ii) Crear una herramienta de investigación lingüística.
- (iii) Crear una BD para la elaboración de materiales lexicográficos.

El formato digital permite almacenar muchos más datos que el formato impreso, y también permite organizar los tres objetivos dentro de una BD adaptada a la modalidad visual. Para hacerla posible hemos tenido en cuenta las aportaciones de las investigaciones existentes sobre LS y sobre BD en LS, que nos han ayudado a establecer criterios de selección y organización de la información. A continuación veremos como está organizada la BD. Los estudios sobre la fonología del signo nos han servido para organizar el sistema de búsqueda de las unidades léxicas de acuerdo con las características de la modalidad visual. En el apartado de la morfología veremos que también existen diferencias entre las LS y las LO que tendremos que tener en cuenta dentro de la BD. Las dificultades que encontramos en la semántica y la pragmática son las propias de cualquier lengua natural, por lo que en estos apartados trataremos los criterios que hemos establecido para el diseño de esta BD en concreto.

Tal como hemos dicho anteriormente, las diferencias de modalidad hacen necesaria la búsqueda de nuevos sistemas de organización de las unidades léxicas. En los sistemas utilizados hasta ahora se ha intentado adaptar la modalidad visual a los sistemas de búsqueda existentes, que permiten una organización alfabética o temática. Pero tal como hemos dicho anteriormente, ese tipo de búsqueda hace que la modalidad visual sea una modalidad dependiente de la glosa.

En este trabajo proponemos un nuevo tipo de organización para la LSC basada en los parámetros fonológicos de la modalidad visual. En otras LS se ha empezado, desde hace ya unos años, la investigación en esa dirección porque las nuevas tecnologías han facilitado esta nueva perspectiva. Consiste en permitir la búsqueda de las unidades léxicas a través de los parámetros fonológicos de la LSC, a parte de la tradicional búsqueda a

través de la glosa, más cómoda quizá para otros usuarios no familiarizados con la LS. Estos parámetros son: la configuración (forma de las manos y los dedos mientras se ejecuta el signo), la localización (el lugar donde se articula), el movimiento (acción de las manos) y la orientación (la dirección de la palma de la mano).

Este sistema de búsqueda está pensado para aplicarse en BD digitales. Tal como podemos ver en la Figura 1. Consiste en un sistema de menús que conducen a submenús, que poco a poco, a través de la selección de parámetros, conducen al usuario hasta el signo que busca. La búsqueda mas detallada empieza con el menú de la configuración. En esta primera pantalla aparecen todas las configuraciones relevantes identificadas en LSC. El usuario tendrá que apretar la imagen de la configuración que considere que tiene el signo que busca. Entonces aparecerá un menú en el que el usuario tendrá que decidir si el signo que busca se realiza con las dos manos o solo con una. Si es con las dos manos, antes de seguir, tendrá que seleccionar el tipo de movimiento de las dos manos: alterno o simultáneo.

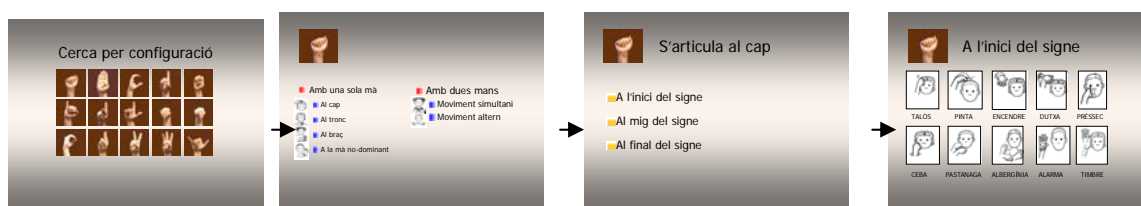


Figura 1. Sistema de menús

Una vez escogida la opción anterior, aparecerá la localización. El usuario tendrá que decir si el signo se articula en la cabeza, el tronco, el brazo o la mano no dominante (siguiendo la división del espacio sígnico que establecen Sandler y Lillo-Martin (2006).

Cuando esté dentro del área mayor del cuerpo donde se articula el signo, el usuario podrá decir si aquella configuración aparece al principio, en el medio o al final del signo. Este paso es importante para los signos compuestos.

Finalmente le aparecerá en pantalla la imagen de todos los signos que cumplen los requisitos solicitados. Como podrían ser muchos, también hay que buscar un orden visual para estos signos. Para ordenarlos, hemos utilizado el concepto de *setting* (Sandler y Lillo-Martin 2006), que hace referencia a la posición concreta de la articulación del signo referente al área mayor donde se encuentra. El orden puede ser el siguiente: dentro de una área mayor del cuerpo, primero los signos situados en la zona alta, después los situados en la zona media y finalmente los de la zona baja. Si dos signos coinciden en una

de estas zonas, primero se pueden poner los que están en el mismo lado que la mano que signa (posición ipsilateral), y luego los que se articulan al otro lado (contralateral). Finalmente, si dos signos coinciden en estos puntos, primero se listan los que tienen contacto, luego los que se encuentran en una distancia mediana (proximal), y por último los que están en la zona más alejada (distal).

Este sistema permite un nivel de concreción del signo bastante alto. Además, en cualquier de los niveles de selección se puede hacer un acceso directo al listado con los elementos seleccionados hasta el momento, lo que puede ser de utilidad en caso de que el usuario desconozca mas características distintivas del signo que busca. A parte de esta forma de búsqueda que se inicia con la configuración del signo pero que después incluye los otros parámetros fonológicos de los signos en LSC, una BD en formato informático permite listar todos los signos que cumplen un requisito concreto como un movimiento, una localización o una orientación determinada. Esta opción se consigue poniendo en la página principal iconos que permitan la búsqueda directa como:

- Búsqueda por movimiento: aparecerá un submenú con los movimientos más relevantes en LSC.
- Búsqueda por localización: aparecerá un submenú con las áreas mayores del cuerpo que después se pueden detallar en localizaciones más pequeñas (*setting*).
- Búsqueda por orientación: no todos los autores la consideran un parámetro fonológico independiente, ya que algunos la incluyen dentro de la configuración, pero se puede añadir como posibilidad de búsqueda, listando también las orientaciones más relevantes de la LSC.

Para hacer posible este sistema de búsqueda, tendremos que determinar, en primer lugar, cuales son las configuraciones, los movimientos, las localizaciones y las orientaciones distintivas en LSC y añadir esta información detallada en la entrada del signo. Una vez encontrado el signo que buscamos, podremos acceder a la información que tiene asociada seleccionándola con el ratón.

Uno de los aspectos más importantes que tiene que decidir el lexicógrafo es la información que quiere incluir en el trabajo lexicográfico que desea elaborar. Cada unidad léxica de una lengua contiene numerosa información. Esta información permite que pueda combinarse o no con otras unidades de la lengua, de forma que el hablante pueda expresar conceptos e intenciones mediante la unión de estas unidades léxicas en estructuras bien formadas. El lexicógrafo tiene que escogerla según sus objetivos y según los usuarios a los que vaya dirigido el material lexicográfico.

Por otra parte, una vez escogida la información que se recogerá en la entrada léxica, tiene que organizarla. El formato informático permite crear un formulario de entrada para cada signo y diferentes accesos para cada uno de los aspectos recogidos. Mediante el sistema que proponemos, que se puede ver en la Figura 2, el usuario podrá acceder rápidamente a lo que desea saber, apretando directamente con el ratón la imagen que contiene lo que le interesa. De esta forma se evitará acumular toda la información en una sola pantalla, lo que podría hacer la consulta poco ágil.



Figura 2. Formulario de entrada del signo

A continuación detallaremos algunos de los apartados que aparecen en la Figura 2. Por razones de extensión hemos escogido algunos de los muchos aspectos que podrían recogerse en un trabajo lexicográfico.

3.1 Información morfosintáctica

La morfología es aquella parte de la lingüística que describe la estructura interna de las palabras y los procesos de creación de nuevas palabras en una lengua. La sintaxis, en cambio, es aquella área de la lingüística que estudia la estructura de las oraciones.

3.1.1 Aspectos morfológicos de los signos

Los signos pueden estar formados por un solo morfema o bien por la unión de distintos morfemas. En las LO la combinación de morfemas para crear un lexema se realiza mediante la unión de radicales y afijos en secuencias lineales. En las LS, en cambio, hay dos formas de combinar los morfemas (Sandler y Lillo-Martin 2006):

- a) Mediante adjunción de afijos.
- b) Añadiendo material fonológico a sus componentes.

En estudios que se han hecho en distintas LS se ha observado que es mucho más productiva la segunda opción que la primera, que es más propia de las LO. La simultaneidad en LS posibilita que estructuras monosilábicas sean a la vez bimorfémicas y esta característica afecta tanto la morfología flexiva como la derivativa.

La morfología flexiva es aquella que modifica la forma de la palabra para adaptarla al contexto sintáctico donde aparece, sin que este cambio formal implique un cambio en el significado léxico. La morfología flexiva que se acostumbra a recoger en diccionarios suele ser la irregular, es decir, la de las palabras que no siguen la derivación habitual de la lengua. En una BD lexicográfica, como no hay problemas de espacio, se pueden incluir en un apartado específico tanto la morfología flexiva regular como la irregular. Por otra parte, la morfología derivativa forma parte de los procedimientos que utilizan las lenguas para crear nuevas palabras y, en consecuencia, las formas procedentes de este tipo de derivación se recogerán como lemas o entradas independientes.

Uno de los fenómenos que encontramos en LSC, igual como en otras LS como la ASL (Sandler y Lillo-Martin 2006) son pares de signos con distinta categoría, nombre / verbo que se distinguen por el movimiento. Se trata de verbos que representan una actividad, mientras que los nombres concretos están semánticamente relacionados con la actividad que denota el verbo. Por ejemplo, en LSC existen parejas como DUCHARSE / DUCHA o bien AVIÓN / VOLAR-EN-AVIÓN que se distinguen solo por el movimiento general de la ejecución del signo. En el caso de DUCHARSE la raíz se reduplica y se ejecuta más rápidamente que en el caso de DUCHA, que tiene un movimiento único más corto y tenso. En el ejemplo de VOLAR-EN-AVIÓN la trayectoria es más larga y el movimiento no es tenso, mientras que en AVIÓN el movimiento es corto y tenso.

En estos ejemplos el signo derivado se construye mediante la modificación de algún aspecto del signo de origen y el resultado es un signo con un significado distinto, de otra categoría léxica, que podría tener dentro de la BD su propia entrada. Una descripción detallada de estos aspectos fonológicos que producen cambios de significado, además, podría ser valiosa para la investigación, mediante consultas de selección dentro de la BD lexicográfica. Por este motivo, además de entrar estas unidades léxicas como dos signos independientes, se pueden describir diferencias fonológicas tales como si el movimiento es corto o largo respecto el signo de origen, lento o rápido, o si la raíz se reduplica.

CONSTRUCCIONES CON CLASIFICADORES

Una de las formas de derivación morfológica más productivas en LS son las construcciones con clasificadores. También encontramos este tipo de construcciones en LO pero no es un fenómeno tan habitual como lo es en LS. Una construcción con clasificadores consiste en una unión de un morfema raíz con un morfema clasificador. Esto hace que la unidad léxica adquiera otra dimensión semántica. En el caso de las LO este morfema se añade a la raíz de forma secuencial, mientras que en las LS se produce simultáneamente. En ambas modalidades el resultado es una estructura morfológica compuesta, que a menudo acaba lexicalizándose y creando una nueva palabra.

El uso de clasificadores está sometido a restricciones. No todos los signos pueden utilizar este tipo de construcciones. Por ejemplo, un signo como PENSAR no tiene construcciones de clasificadores asociadas, pero en cambio un signo como COMER sí que tiene, ya que adopta una configuración u otra dependiendo de lo que se coma (un bocadillo, espaguetis, una pizza, pipas, etc.)

Como no se pueden entrar dentro de la BD todas las formas que puede adoptar el verbo COMER, BEBER o COGER, porque la lista sería interminable (aunque podemos poner ejemplos), tendremos que buscar la manera de dar al usuario esta información sin saturar la entrada. Una forma de hacerlo es hacer constar si un signo tiene o no asociadas construcciones con clasificadores, en caso que así sea, de qué tipo.

3.2 Aspectos sintácticos de los signos

Son los aspectos que determinan por un lado la ubicación de la unidad léxica cuando se inserta dentro de una estructura sintáctica, y por otro lado, con qué elementos lingüísticos pueden combinarse.

3.2.1 La categorización y la subcategorización

La categorización sintáctica define cada unidad léxica respecto la clase distribucional y funcional de la que forma parte. En catalán por ejemplo, son categorías sintácticas los nombres, los verbos, los adjetivos, los adverbios, las preposiciones y las conjunciones (Cabré 1994). El hecho de pertenecer a una categoría significa que la unidad léxica pertenece necesariamente a subcategorías de orden morfológico que están determinadas para cada clase distribucional. Por ejemplo, y de acuerdo con la clasificación tradicional de Padden (1988) los verbos en LS se dividen en:

(i) **Verbos simples o invariables:** aquellos que no presentan concordancia y que por lo tanto no modifican su forma en función de los argumentos del sujeto y objeto. Verbos como PENSAR, RECORDAR, etc. son invariables.

(ii) **Verbos direccionales o de concordancia gramatical:** aquellos que concuerdan con el sujeto y objeto modificando parámetros fonológicos. Verbos como AYUDAR, DAR, ENGAÑAR, OBLIGAR, etc. son direccionales.

(iii) **Verbos espaciales o de concordancia locativa:** aquellos que modifican la forma según los argumentos locativos del predicado o de la localización de sus argumentos animados. Verbos como PONER, TRANSPORTAR, etc. son espaciales.

Por otra parte, los verbos pueden ser transitivos o intransitivos, por ejemplo. Dependiendo de si pertenecen a una subcategoría u otra tendrán que aparecer en la oración con uno o más de un complemento. Esta información también se puede recoger dentro de la entrada del signo, ya que podría orientar al usuario sobre cómo tiene que construir morfológicamente y sintácticamente el signo.

3.2.2 Selección categorial

Por otra parte, la subcategorización especifica las combinaciones posibles de la unidad léxica desde dos puntos de vista: el de la selección categorial (las unidades léxicas presentan restricciones respecto a las categorías con las que pueden combinarse), y de la selección semántica (también tienen restricciones respecto del papel temático de cada pieza con las que se pueden combinar).

Un ejemplo de selección categorial es el hecho que en LSC se pueda decir PAN COMO pero no AZUL COMO. No podemos decir AZUL COMO porque el verbo COMER selecciona un nombre como complemento, no un adjetivo. Un ejemplo de selección semántica es que digamos en LSC, PERSONA PIENSA pero no COCHE PIENSA. El verbo PENSAR selecciona un complemento animado que conlleve la acción.

Este tipo de información va mucho más allá de lo que suelen recoger los diccionarios normativos y por eso lo citamos brevemente. No obstante, es interesante saber que se suelen recoger en los diccionarios denominados combinatorios.

3.3 Información semántica

Es aquella que hace referencia al significado del signo⁹. Es fundamental poder definirlo. En los materiales existentes a menudo se ha utilizado la glosa como única definición del signo. La glosa puede ser utilizada en algunas ocasiones, pero nunca puede sustituir la definición en LS, puesto que podría rasgar su significado. De entrada la glosa es una palabra de otra lengua, y los significados no siempre son equivalentes. También podemos encontrarnos con que una única glosa pueda definir más de un signo que en LS tienen diferentes matices, o bien que distintas glosas puedan utilizarse para un solo signo.

Por eso la mejor opción es definir los signos en la propia lengua, mediante otros signos, y teniendo en cuenta, entre otras cosas, los siguientes aspectos:

- a) Hacer las descripciones con signos de uso común y de alta frecuencia entre los signantes con el objetivo que sean lo más comprensibles posible para los signantes de la comunidad.
- b) También es importante hacer una extracción del significado de todos los ejemplos que se encuentren en el corpus, y ordenar las excepciones siguiendo unos criterios definidos. Un signo puede utilizarse en muchos contextos y tener un matiz diferente en cada uno de estos contextos. No podemos recogerlos todos porque el listado de significados sería demasiado largo y el usuario acabaría perdiéndose. Es necesario, pues, hacer una síntesis de todos los significados que encontremos. También hace falta establecer un orden de las acepciones sintetizadas. Un criterio de orden podría ser la frecuencia con que un signo se utiliza con aquel significado, o bien la posición del signo dentro de la comunidad, es decir, listar primero el significado común a un colectivo más amplio de signantes dentro de la comunidad. Otro criterio que también se utiliza a menudo es el lógico: primero se lista el significado básico de la palabra y después los significados relacionados. En todo caso, tendrá que haber una cierta flexibilidad para intentar encontrar el criterio que se adapta mejor a cada entrada.
- c) Se tienen que evitar las definiciones circulares. Una definición circular es aquella donde la definición de una entrada A se basa en el concepto de una entrada B, y la definición de la entrada B se basa en el concepto de la entrada A. A veces las definiciones se elaboran a través de sinónimos. Este puede ser un recurso útil pero tenemos que ir con cuidado porque se podría caer fácilmente en una definición circular.
- d) Una técnica que se utiliza para definir las unidades léxicas de una lengua es usando unidades de la misma categoría léxica, ya que de esta forma se está facilitando simultáneamente información sobre su ubicación dentro de la estructura sintáctica.

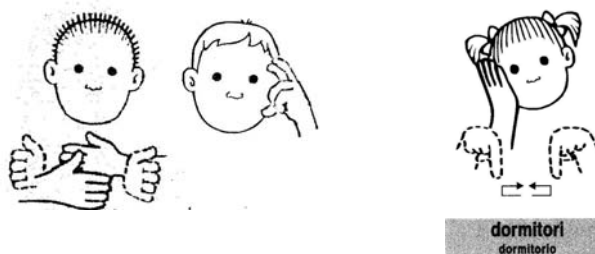
⁹ Existen algunos aspectos del área de la semántica que también podrían recogerse en un material lexicográfico, como la estructura eventiva y la estructura argumental de la unidad léxica, la selección semántica o los sinónimos y antónimos. Por razones de extensión no los trataremos en este curso.

3.4 Información pragmática

La información pragmática es aquella relacionada con el uso de la pieza léxica en un contexto comunicativo determinado. Las piezas léxicas forman parte de una red asociativa basada en elementos de la experiencia personal y colectiva. Eso significa que los hablantes interiorizan, por una parte, en qué situaciones se utilizan las palabras con más o menos frecuencia, como son las personas que las utilizan, qué reacciones provocan cuando se utilizan en un acto comunicativo, etc.

En LSC, como en cualquier lengua natural, podemos encontrar diferentes tipos de variantes:

- a) Variantes diatópicas: unidades léxicas que se utilizan en una zona geográfica concreta o bien que presentan diferencias de uso, de ejecución o de interpretación según la zona del dominio lingüístico. Este es el caso del signo DORMITORIO en LSC, que en la zona de Barcelona se realiza de una forma y en la zona de Lleida, de otra¹⁰:



- b) Variantes diastráticas: unidades léxicas que se utilizan en un sector específico de la comunidad lingüística, como por ejemplo signos como COLOR-ROSA o bien LETRA-U, que en LSC tienen distintas configuraciones según la generación:¹¹

COLOR ROSA



¹⁰ El primer dibujo es del material facilitado por LLESIG (Barcelona) para el aprendizaje de la LS, hecho por Frank Vidiella (curso 2006-2007). El segundo dibujo es del DTLSC (Martín y Alvarado 2004), que contiene signos de la zona de Lleida.

¹¹ Dibujos del DTLSC (Martín y Alvarado 2004)

Signo actual

Signo que todavía hacen algunas personas mayores (enseñando las uñas)

LETRA U



Signo actual



Signo propio de algunas personas mayores

- c) Variantes diafásicas: unidades léxicas que se utilizan en algunos registros concretos o signos que presentan diferencias según el registro donde aparecen. Un ejemplo de variante diafásica es el signo PRESERVATIVO, que presenta tres variantes según la situación:



Muy formal



Normal



Muy informal

- d) Variantes diatécnicas: unidades léxicas propias de una materia especializada. El signo DENTRO, por ejemplo, significa MATERIA dentro del campo de la educación.



Definición 1: Situación en el interior de un espacio limitado

Definición 2: (En educación) Contenido, materia, aspecto conceptual de una asignatura, curso, etc.

dins / dintre
dentro

Como las BD permiten vincular las unidades léxicas, las variantes pueden vincularse de forma que dentro de la información de un signo se pueda acceder a las distintas formas existentes en todo el dominio lingüístico.

En este apartado hemos intentado mostrar ejemplos del tipo de información que se podría incluir dentro de una BD lexicográfica. Tal como hemos dicho anteriormente, la información que se incluya dependerá de los objetivos y del tipo de destinatarios a los cuales vaya dirigido.

4 CONCLUSIONES

Las diferencias de modalidad de las LS y las LO hacen que su componente léxico no pueda describirse mediante los mismos parámetros. Ambas modalidades comparten los principios lingüísticos pero difieren en los parámetros que los configuran y en la forma como se transmite y se recibe la competencia gramatical que los hablantes tienen interiorizada. Estas divergencias han afectado los estudios sobre la lexicografía del signo, muy atrasada respecto a los de la modalidad oral, y también a los sistemas de almacenamiento de las unidades léxicas de las LS, ya que las formas tradicionales de compilación de datos se adaptan muy poco o nada a la modalidad visual.

Las diferencias de modalidad hacen que sea imprescindible buscar otras formas de descripción y almacenamiento de las unidades léxicas de las LS. Otras formas que vayan más allá de la simple transcripción del significado a una glosa en la LO más cercana. Tal como hemos visto, las nuevas tecnologías nos permiten ofrecer a los signantes una BD lexicográfica mucho más completa, con más información sobre los signos, como qué significan, con que otras unidades pueden combinarse, en qué contextos aparecen o se usan, entre muchos otros.

En este artículo hemos descrito el estado de la cuestión en las LS en general. Nos hemos centrado también en el estado de la lexicografía en LSC y hemos esbozado las características que podría tener una BD que estuviera adaptada a las características intrínsecas de las LS. Esperamos que este tipo de organización nueva de BD, basada en los parámetros fonológicos de la modalidad visual, pueda servir a futuros investigadores en la creación y aplicación de BD digitales.

5 REFERENCIAS

Boyes Braem, P. 2001. A multimedia bilingual database for the lexicon of Swiss German Sign Language". *Sign Language and Linguistics* 4:1/2 : 133- 143

Brien, D. 1992. *Dictionary of British Sign Language / English*. London-Boston

Cabré, M.T. 1994. *A l'entorn de la paraula (I); Lexicologia General*. València

Crasborn, O. *et al.* 2001. A phonological database for sign languages. A *Sign Language and Linguistics* 4:1/2: 215-228

Girod *et al.* 1997. *DICTIONAIRE LSF / FRANÇAIS*. Paris

- Gras, V. 2006. *La Comunidad Sorda como comunidad lingüística: panorama sociolingüístico de las lenguas de signos en España*. Tesis doctoral. Universitat de Barcelona. Facultat de Filologia. Departament de Lingüística. Barcelona
- Hanke, T. i Storz, J. 2008. iLex – A database tool integrating sign language corpus linguistics and sign language lexicography. A *Proceedings of the Sixth International Conference on Language Resources and Evaluation*, May 26-June 1, 2008, Marrakech, Morocco. European Language Resources Association
- Johnston, T. 2001. The lexical database of Auslan (Australian Sign Language). A *Sign Language and Linguistics* 4:1/2: 145-169
- Johnston, T. 2008. Corpus linguistics & signed languages: no lemmata, no corpus. A *Proceedings of the Sixth International Conference on Language Resources and Evaluation*, May 26-June 1, 2008, Marrakech, Morocco. European Language Resources Association
- König *et al.* 2008. Corpus-based Sign Dictionaries of Technical Terms – Dictionary Projects at the IDGS in Hamburg. A *Proceedings of the Sixth International Conference on Language Resources and Evaluation*, May 26-June 1, 2008, Marrakech, Morocco. European Language Resources Association
- Konrad *et al.* (en premsa) What's in a sign? Theoretical lessons from practical sign language lexicography. A *Signs of the time. Selected papers from TISLR 8*, ed. J. Quer. Hamburg: Signum Verlag
- Krammer, K. *et al.* 2001. The Klagenfurt database for sign language lexicons. *Sign Language and Linguistics* 4:1/2 : 191-201
- Krammer, K. *et al.* 2008. The Klagenfurt lexicon database for sign languages as a web application: LedaSila, a free sign language database for international use. A *Proceedings of the Sixth International Conference on Language Resources and Evaluation*, May 26-June 1, 2008, Marrakech, Morocco. European Language Resources Association
- Martín, I. i Alvarado, M. 2004. *Diccionari Temàtic de Llenguatge de Signes Català*. Lleida.
- Neidle, C. 2001. SignStream: A database tool for research on visual-gestural language. A *Sign Language & Linguistics* 4:1/2: 203–214
- Padden, C. 1988 [1983]. Interaction of Morphology and Syntax in American Sign Language. Dins de *Outstanding Dissertations in Linguistics*. New York: Garland (originalment publicada com a: PhD dissertation. San Diego: Universitat de California.)
- Perelló, J. i J. Frigola. 1998. *Lenguaje de signos manuales*. Madrid
- Prillwitz, S. *et al.* 2008. "DGS corpus project – Development of a corpus based electronic dictionary German Sign Language / German". A *Proceedings of the Sixth International Conference on Language Resources and Evaluation*, May 26-June 1, 2008, Marrakech, Morocco. European Language Resources Association
- Quer, J. *et al.* 2005. *Gramàtica bàsica LSC*. CD-ROM. Barcelona: Federació de persones sordes de Catalunya
- Ribera, E. 2007. Projecte de base de dades lexicogràfica per a la LSC (llengua de signes catalana). Treball de recerca dins el programa de doctorat Ciència cognitiva i llenguatge. Director: Josep Quer. Tutora: Maria Teresa Espinal
- Sandler, W. i D. Lillo-Martin 2006. *Sign Language and Linguistic Universals*. Cambridge.
- Savolainen, L. 2001. The database system used in the Finnish Sign language Dictionary Project. *Sign Language and Linguistics* 4:1/2 : 125-132
- Schermer, T. *et al.* 2001. Developing linguistic specifications for a sign language database: The development of Signbase. A *Sign Language & Linguistics* 4:1/2: 253-274.

Senghas, A. 2001. Spatial and temporal coding of Nicaraguan Sign Language inMediaTagger. A *Sign Language & Linguistics* 4:1/2: 229–240

Slobin, D. *et al.* 2001. Sign language transcription at the level of meaning components: The Berkeley Transcription System (BTS). A *Sign Language & Linguistics* 4:1/2: 63–104

Stokoe, W. 1960. Sign Language structure: an outline of the visual communications systems of the American deaf. Dins de *Studies in Linguistics: Occasional Papers*. Buffalo: University of Buffalo

Stokoe, W. *et al.* 1965. *A Dictionary of American Sign Language on Linguistic Principles*. Washington DC: Gallaudet College Press.[2a edició 1976 per Linstok Press]

Troelsgård, T. i J.H. Kristoffersen. 2008. An electronic dictionary of Danish Sign Language. Sign Languages: spinning and unraveling the past, present and future. TISLR9, forty five papers and three posters from the 9th. Theoretical Issues in Sign Language Research Conference, Florianopolis, Brazil, December 2006. (2008) R. M. de Quadros (ed.). Editora Arara Azul. Petrópolis/RJ. Brazil. <http://www.editora-arara-azul.com.br/EstudosSurdos.php>.

Recursos web (última actualització el juliol del 2008)

<http://gebaren.ugent.be/>

<http://ledasila.uni-klu.ac.at/>

<http://www.auslan.org.au/>

<http://www.gebarententrum.nl/gebarenwoordenboek.asp>

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

<http://www.sign-lang.uni-hamburg.de/Projectseng/SLDicts.html>

<http://www.xtec.net/dnee/act/miraquedic>

http://www.xtec.net/dnee/lsc/terminologia_llengua/index.htm

www.tegnsprog.dk