

Signbank: una herramienta para el estudio de las lenguas de señas como L2

Miroslava CRUZ-ALDRETE (Autora para correspondencia)

Centro Interdisciplinario de Investigación en Humanidades

Universidad Autónoma del Estado de Morelos, UAEM

miroslava.cruza@uaem.edu.mx

<https://orcid.org/0000-0001-8110-4300>

Resumen: El estudio de las lenguas de señas (LS) como segundas lenguas es un nuevo campo de conocimiento, cuyo desarrollo está íntimamente relacionado con la lingüística de corpus (LC). Si bien hemos sido testigos del auge en la LC con respecto al análisis de las lenguas orales minoritarias, en el caso de aquellas de modalidad visogestual pareciera que su crecimiento va a la zaga. Esto se debe a varios factores que involucra la discusión sobre las características de las LS, desde su reconocimiento como una lengua natural (objeto de estudio de la lingüística, de otras disciplinas del ámbito de las humanidades y de las ciencias cognitivas), hasta la complejidad de su representación, al tratarse de sistemas lingüísticos simultáneos, secuenciales, espaciales y temporales. En la actualidad observamos un creciente interés por el aprendizaje de estas lenguas como una segunda lengua (L2), tanto por personas oyentes, como por personas sordas usuarias de una lengua de señas que se interesan por el aprendizaje de otra LS. Esta situación conduce a los estudiosos de las LS a volcarse en los datos primarios que se tienen sobre este tipo de idiomas, con el fin de discutir su estructura y su variación sociolingüística. Teniendo en cuenta todas estas cuestiones, la intención del presente artículo es presentar el desarrollo de la aplicación *Signbank* (<https://signbank.cls.ru.nl>) como una herramienta que puede contribuir a la investigación en el campo del estudio de las LS como L2.

Palabras clave: lengua de signos, lingüística de corpus, documentación, modalidad visogestual

Català:

Signbank: una eina per a l'estudi de les llengües de signes com a L2

Resum: L'estudi de les llengües de signes (LS) com a segones llengües és un nou camp de coneixement, el desenvolupament del qual està íntimament relacionat amb la lingüística de corpus (LC). Si bé hem estat testimonis de l'auge a la LC pel que fa a l'anàlisi de les llengües orals minoritàries, en el cas d'aquelles de modalitat visogestual sembla que el creixement va al darrere. Això és degut a diversos factors que involucra la discussió sobre les característiques de les LS, des del seu reconeixement com a llengua natural (objecte d'estudi de la lingüística, d'altres disciplines de l'àmbit de les humanitats i de les ciències cognitives), fins a la complexitat de la seva representació, en tractar-se de sistemes lingüístics simultanis, seqüencials, espacials i temporals. Actualment, observem un interès creixent per l'aprenentatge d'aquestes llengües com una segona llengua (L2), tant per persones oyents com per persones sordes usuàries d'una llengua de signes que s'interessen per l'aprenentatge d'una altra LS. Aquesta situació condueix els estudiosos de les LS a abocar-se a les dades primàries que es tenen sobre aquest tipus d'idiomes, per tal de discutir-ne l'estructura i la variació sociolingüística. Tenint en compte totes aquestes qüestions, la intenció del present article és presentar el desenvolupament de l'aplicació *Signbank* (<https://signbank.cls.ru.nl>) com una eina que pot contribuir a la recerca al camp de l'estudi de les LS com a L2.

Paraules clau: llengua de signes, lingüística de corpus, documentació, modalitat visogestual

English:

Signbank, a tool for the study of sign languages as L2

Abstract: The study of sign languages (SL) as second languages is a new field of knowledge. Its development is closely related to corpus linguistics (CL). Although we have witnessed a boom in CL with respect to the analysis of minority oral languages, in the case of visogestual languages it seems that their growth is lagging behind. This is due to several factors that involve the discussion on the characteristics of SLs, from its recognition as a

natural language (object of study of linguistics other disciplines in the humanities and cognitive sciences), to the complexity of their representation, since they are simultaneous, sequential, spatial and temporal linguistic systems. At present we observe a boom in the learning of these languages as a second language (L2), both by hearing people, as well as by deaf people who use a sign language and are interested in learning another SL. This situation leads SL scholars to turn to the primary data available on this type of language in order to discuss its structure and sociolinguistic variation. Taking into account all these issues, the intention of the present article is to present the development of the Signbank (<https://signbank.cls.ru.nl>) application as a tool that can contribute to research in the field of the study of SL as L2.

Keywords: sign language, corpus linguistics, documentation, visogestual modality

Introducción

Es lugar común hablar sobre la juventud de la investigación en las lenguas de señas (LS), aun cuando podamos identificar, entre los siglos XVII y XVIII, algunos estudios pioneros que hacen referencia al uso de las manos para comunicarse con las personas sordas, e incluso identifican algunas características de este tipo de sistemas lingüísticos. Un ejemplo de ello lo podemos encontrar en obras cuyo tema central era la enseñanza de las personas sordas, en dichos textos los autores expresaban cómo a través de las señas se podía propiciar el aprendizaje de la lengua dominante, tal es el caso de Bonet (1620) y Hervás y Panduro (1795). Con todo, está claro que su reconocimiento como un campo de estudio de la lingüística se da alrededor de la segunda mitad del siglo XX, a partir de los estudios de Stokoe sobre la lengua de señas americana (1960, 1976, 1978, 2001).

Desde una perspectiva histórica podemos identificar las diversas etapas que han sido un hito en la definición de la LS como un objeto de estudio de la lingüística, de las ciencias cognitivas y de la mirada socioantropológica (Stokoe, 2001; Vera Villaverde, 2020)¹. El primer periodo del estudio de las lenguas de modalidad visogestual nos revela el interés principal por demostrar que las LS son lenguas naturales, así como por buscar similitudes entre estos sistemas lingüísticos con respecto a la organización de las lenguas orales (LO). En una etapa posterior, se aprecia una discusión centrada en las características que hacen únicas a las lenguas de modalidad visogestual, cuya realización involucra el uso de los articuladores activos (las manos), el reconocimiento del uso de los gestos realizados con la cara y el movimiento del cuerpo o de la cabeza, cuyo conjunto se ha denominado componente de rasgos no manuales (RNM). Todos estos elementos se organizan de manera simultánea, secuencial y espacial. Es la optimización del *espacio señante*² una de las principales diferencias de las LS, pues este al igual que el uso de los RNM, son parte fundamental del sistema gramatical de este tipo de lenguas a diferencia de las LO (Crasborn, 2006; Barberà Altimira, 2015; Brentari, 2019).

Entrado el siglo XXI, cobra relevancia la preocupación por el peligro de extinción en que se encuentran algunas de las LS (Nonaka, 2004), debido al prejuicio en torno a su uso, al considerar que sería mejor para las personas sordas aprender la LO dominante y, en consecuencia, negar el derecho de las personas sordas a adquirir su lengua natural, la LS. Un ejemplo de ello lo encontramos en el reconocimiento de las LS como idiomas nacionales. Se trata de un suceso

¹ Algunos autores a esta nueva rama de la lingüística la denominan *signolingüística*, entre ellos, Minguet Soto (2001) y Herrero Blanco (2002).

² Se llama espacio señante al espacio frente al cuerpo del usuario de la LS; este tiene diversos valores fonológicos, morfosintácticos, semánticos, y pragmáticos (Liddell, 2003).

reciente a nivel mundial, como se puede observar en el caso de algunos países de América Latina, por ejemplo, México: es en el 2005 cuando la Lengua de Señas Mexicana (LSM) es considerada como una lengua nacional y parte del patrimonio lingüístico y cultural de este país, a la par del español y de los idiomas de los pueblos originarios. Esto no es cosa menor, debido a su impacto en el desarrollo de políticas públicas para las comunidades lingüísticas minoritarias, entre ellas las comunidades sordas (CS).

En este mismo sentido, hay un llamado de atención sobre las lenguas de señas emergentes, rurales o indígenas, de las que poco se conoce y que también se encuentran en vulnerabilidad ante la presencia de las LS dominantes (Zeshan y de Vos, 2012). Se observa el desplazamiento de estas LS minoritarias ante el uso de las LS dominantes que, por ejemplo, se suelen emplear en los espacios educativos. De igual modo resulta de interés actual entender los efectos del desplazamiento o del contacto, ya sea entre lenguas con la misma o con diferente modalidad (Cruz-Aldrete, Serrano Morales y Medrano, 2020).

Como se puede observar, el estudio de las LS ha tenido un camino sinuoso que hoy nos enfrenta a nuevos retos. Es indudable que los hallazgos en la investigación de las LS, a la par de abonar en nuestra comprensión del lenguaje, también han contribuido en la lucha de los derechos lingüísticos y culturales de los miembros de la CS. Hemos sido testigos de cómo hoy en día hay una mayor visibilidad de las LS en los distintos espacios de la vida pública, ya sea en los medios de comunicación masiva o en los centros educativos. El acceso a la información, a la cultura y recreación, a la justicia y a la salud por parte de los sordos señantes ha puesto de relevancia la figura de los intérpretes de LS (ISL), lo cual nos conduce a otro tema de estudio, que es el aprendizaje de la LS como una L2.

Si bien nos parece notorio el aumento del número de personas oyentes que por razones laborales desean aprender la LS como una segunda lengua (L2), también se ha observado que hay un interés en adquirir otra LS/L2 por parte de las personas sordas que ya tienen una LS como L1. Este interés responde en gran medida a la nueva era digital en la que se vive, la cual ha impactado en la dinámica de interacción de los sordos señantes con respecto a la apropiación de los espacios virtuales, el uso de las redes sociales y la posibilidad que hoy tienen de acceder a distintos dispositivos tecnológicos, entre ellos los teléfonos inteligentes, que les permiten tener una comunicación cara a cara en tiempo real. El desarrollo tecnológico ha abierto la posibilidad a sordos y oyentes de conocer miembros de distintas comunidades con lenguas y culturas diferentes. Esta situación ha marcado la pauta para el inicio de un nuevo ámbito en el estudio de las LS.

La investigación de las LS como segundas lenguas (L2) es un campo reciente en el cual se abordan temas de distinta naturaleza y del cual se desprende una primera discusión: la enseñanza de la LS como L2 para personas oyentes *versus* personas sordas señantes. Es decir, existe un complejo universo en la enseñanza de la LS como L2, en el cual encontramos aprendices oyentes de una L2 con diferente modalidad (M2-L2) y aprendices sordos de una L2 de la misma modalidad (M1-L2). Incluso habría que agregar aprendices sordos que no tienen la LS como primera lengua, que

interactúan empleando la lengua oral dominante y que se interesan por aprender la LS como L2 (Chen Pichler y Koulidobrova, 2016; Schönström, 2021).

La enseñanza de la LS como L2 constituye un desafío para los docentes y especialistas que requieren no solo de conocer la gramática de la LS, sino de contar con una diversidad de recursos y materiales didácticos para su enseñanza. Por consiguiente, es necesario tratar el tema de la documentación lingüística de las LS, la propuesta de una gramática didáctica y la creación de un corpus lingüístico robusto del cual surja una base de datos léxica para el desarrollo de un diccionario que constituya una fuente de ejemplos sobre las diferentes estructuras gramaticales de la LS que se enseña.

Los puntos que se han enumerado para la enseñanza de la LS como L2 nos conducen a la discusión sobre la falta de corpus de LS. De acuerdo con Fenlon *et al.* (2015) los corpus modernos de LS son de reciente aparición, lo cual se explica por la ausencia de un sistema de transcripción ampliamente reconocido y favorable para los datos de una lengua de señas, así como por la falta de una tecnología adecuada (Crasborn *et al.*, 2008). No obstante, mencionan que en los últimos años se ha propiciado la aparición de una serie de proyectos de corpus de LS, gracias a los avances tecnológicos y al creciente reconocimiento entre los investigadores de LS sobre el papel que desempeña el contar con un conjunto de datos empíricos más amplio para comprobar o descartar las hipótesis sobre la estructura y el uso de la LS, o para verificar las afirmaciones que hasta el momento se hacían de estas lenguas (basadas en descripciones provenientes de un número limitado de datos).

Es en este contexto en el que ha nacido el recurso llamado *Signbank* (<https://signbank.cls.ru.nl>). Como se verá más adelante, *Signbank* fue originalmente construido para apoyar al diccionario *Auslan Signbank* (<http://www.auslan.org.au>), un diccionario en línea de la Lengua de Señas Australiana (Auslan), que tiene como antecedente el trabajo de documentación y análisis de Johnston (2001). Esta aplicación provee de un marco para el desarrollo de una base de datos léxica de una LS determinada, a la cual se le asocian videos que ilustran la realización de estas señas. El proyecto original *Signbank* ha sido la base de una serie de nuevos diccionarios de lengua de señas y de corpus de otras LS, entre ellas la lengua de señas británica (BSL), la lengua de señas de los Países Bajos (NGT), la lengua de señas finlandesa (FinSL), la lengua de señas flamenca (VGT, por sus siglas en holandés), la lengua de señas americana (ASL), y la lengua de señas brasileña (LIBRAS).

La intención de este artículo es discutir la pertinencia de la aplicación web *Signbank* (<https://signbank.cls.ru.nl>) como un punto de referencia para el desarrollo de una base de datos léxica, que pueda ser empleado para el estudio de una LS como L2. Para ilustrar su organización y uso, tomo como ejemplo *ASL Signbank*, lo cual responde a varios motivos: 1) la ASL es una de las LS más estudiadas; 2) por su empleo extendido fuera de los Estados Unidos de Norteamérica (p. ej. en Canadá, en la frontera con México y en regiones de África Occidental y del Sudeste Asiático); 3) por las relaciones de parentesco entre la ASL, LSM (lengua de señas mexicana) y LIBRAS (lengua

de señas brasileña); 4) y al interés por aprender ASL por razones laborales y educativas, debido al contacto entre ASL y LSM, por la proximidad geográfica de la frontera norte de México con Estados Unidos.

1. La lingüística de corpus de LS

Un tema del cual poco se ha discutido en la lingüística de la LS es la falta de corpus sobre estas lenguas, como apuntan Johnston y Schembri (2010). Quizá esta declaración puede causar sorpresa para quien no esté familiarizado con el estudio de este tipo de lenguas, pero, desafortunadamente, para aquellos que se han aproximado al análisis de las LS dicha afirmación no es de extrañar.

Si bien es invaluable el trabajo realizado por los investigadores de distintas lenguas de señas desde la década de los sesenta del siglo pasado, el cual ha allanado el camino para el desarrollo de la lingüística de la LS es importante plantear una reflexión crítica sobre las condiciones en las que se desarrollaron estos primeros estudios. Así, conviene tener en cuenta, el tipo de equipo con el cual se registraba la producción de los colaboradores sordos y, en consecuencia, los medios para archivar estos datos, y también hay que atender al hecho de que varias de las generalizaciones sobre la estructura de las LS, en muchos casos, se derivó de los datos obtenidos de un pequeño grupo de señantes.

Es claro que hay un mayor interés por la labor de la documentación lingüística, incluso desde una etapa previa a la recolección de los datos primarios, con el fin de mejorar la descripción y el análisis gramatical de la LS (Haviland y Farfán, 2007). Sin embargo, pareciera que solo hemos pasado del registro o la representación de las señas en un solo plano (el empleo de fotografías o dibujos para ilustrar la forma y uso de las señas), a contar con su imagen de manera tridimensional (videos), sin atender al sustento teórico que implica la metodología para la creación de corpus, es decir, criterios de selección de participantes, el papel de colaboradores sordos dentro del proceso de investigación, las formas de elicitación, obtención de datos de contextos naturales o de la vida cotidiana, la participación de niños sordos hasta adultos mayores, etc., al igual que la inconsistencia en la adopción de convenciones para la anotación y traducción de los datos en un archivo digital.

Así, la lectura de los trabajos de Johnston (2001), Johnston y Schembri (2010), y de Cassidy *et al.* (2018), entre otros, nos revela que no será hasta los primeros años del siglo XXI cuando en la literatura lingüística sobre las LS se aborde la noción de *corpus* en el sentido contemporáneo de *corpus lingüístico*. Estos autores consideran que en el estudio de las LS se había mantenido el uso del término de *corpus* en el sentido más tradicional, es decir, el investigador basaba su descripción o análisis lingüístico a partir de un conjunto de datos extraídos de un corpus que solía consistir en una colección de videos (por lo general de difícil acceso o archivados de manera inadecuada), el cual además podía o no tener una transcripción. Incluso, un corpus podría componerse solo de la transcripción de los datos de LS. Por tanto, será con el paso de la primera década de este nuevo siglo cuando los interesados en el estudio de las LS empleen el término de *corpus lingüístico* para hacer referencia a una recopilación de ejemplos reales de la LS en uso. No se obvia el

hecho que también se pueda obtener una colección de señas, con miras a ser lo más representativas posible de la lengua y sus usuarios, y con metadatos asociados, bajo un formato accesible y legible por un dispositivo electrónico.³

Por otra parte, habría que añadir que la falta de corpus lingüísticos modernos también responde a la complejidad en el registro de estas lenguas, en dos sentidos. Uno de ellos vinculado con el empleo de los recursos técnicos, tecnológicos y del desarrollo de *software* específicos; el otro, relacionado con la creación, la adopción y el uso de herramientas para la transcripción de las LS, que hiciera posible la lectura (amable) de los datos primarios para quienes acceden a ellos.

De acuerdo con Fenlon *et al.* (2015), todavía a finales del siglo XX y principios del XXI eran pocos los textos de LS que podían someterse a un etiquetado con el fin de estudiar la frecuencia de aparición de signos o construcciones gramaticales específicas. Para estos autores la ausencia de este tipo de textos pone en evidencia, en primera instancia, la falta de un sistema de escritura de LS aceptado de manera amplia⁴, así como la ausencia de un sistema de notación especializado y estandarizado que pudiera utilizarse de manera amigable para la transcripción⁵:

Coincido con los autores con respecto a la problemática de la anotación y transcripción de la LS, como se puede observar en el uso de las glosas, en particular al estudiar y referirse a la estructura oracional. Aun cuando el empleo de las glosas es una práctica común en los estudios de LS, esto no está exento de desventajas, pues existe una serie de elementos que de manera simultánea intervienen, por ejemplo, la referencia deíctica a partir del uso de la mirada o de la disposición del cuerpo, que proporcionan información morfosintáctica cuya notación se vuelve compleja a partir del uso de la glosa manual⁶.

Es innegable que el desarrollo tecnológico y el avance de la informática brindó una solución al problema de la transcripción. Esta se tradujo en una mayor capacidad de captura y de almacenamiento de grandes cantidades de video digital y en el acceso a un *software* de anotación de video alineado con el tiempo. Asimismo, desde la perspectiva de Fenlon y sus colaboradores, esto dio lugar a la aparición de enfoques basados en corpus para la lingüística de las LS.

Entre los resultados del uso de las nuevas tecnologías, encontramos los programas de *software* de anotación multimedia. Uno de ellos, ampliamente utilizado entre los investigadores de LS y LO, es ELAN (<https://archive.mpi.nl/tla/elan>), el cual ha sido desarrollado por Max Plank Institute for Psycholinguistics, como se muestra en la figura 1. Esta herramienta permite que las transcripciones puedan alinearse directamente con un archivo

³ Con esto no queremos decir que es mejor realizar cualquier trabajo de investigación con corpus de datos extensos. Si bien celebramos que haya corpus de datos extensos en las LS como en las LO, somos de la idea que al aproximarnos al estudio de un fenómeno de la LS también es posible emplear un corpus propio más manejable, el cual responda a los objetivos de la investigación.

⁴ Signwriting (<https://www.signwriting.org/>) es un sistema de escritura cuyo uso por miembros de distintas CS de varios países ha ganado terreno de manera paulatina.

⁵ Existen varios tipos, entre ellos el Sistema de Notación de Hamburgo, pero su uso requiere de un software de fuentes especial (Prillwitz y Zeinert, 1989)

⁶ Cabe mencionar que para la transcripción de la estructura fonética/fonológica de las señas, se ha utilizado la notación propuesta por Stokoe (1976), Liddell y Johnson (1989), en la cual a partir de una serie de símbolos se representan los parámetros que intervienen en la realización de cualquier seña (p.ej. configuración manual, lugar, movimiento, rasgos no manuales).

multimedia de un segmento de cualquier discurso en LS. Asimismo, las anotaciones superpuestas pueden almacenarse en distintas gradas para identificar el nivel de análisis correspondiente, por tanto, esto ha brindado la oportunidad de incluir la referencia sobre la participación de los articuladores no manuales (gestos, mirada, movimiento del cuerpo, cabeza, etc.) y del uso del espacio señante.

Figura 1
Corpus LSM (Guillermo Hernández Santana 2022)

The screenshot displays the ELAN 5.6-AVFX software interface. At the top, there is a menu bar with options: Archivo, Editar, Comentario, Línea, Tipo, Buscar, Visualizar, Opciones, Ventana, Ayuda. Below the menu bar is a toolbar with various icons for file operations and playback. The main workspace is divided into several tracks. The top track is labeled 'SEG' and shows a video segment. Below it is the 'TRAD' track, which contains the Spanish translation 'Y SU HIJO' and 'ELEFANTE'. The 'LEX' track shows a list of words and their corresponding sign language units. The 'MORF' track shows morphological information. The 'RNM' track shows non-manual features. The 'Notas' track shows notes. The 'LEX-cp' track shows cross-partitioned lexicon information. The interface also includes a timeline at the bottom with time markers from 00:00:04.000 to 00:00:14.000.

Otra ventaja en el uso de ELAN es que permite que el material de video digital de origen pueda mantenerse como el dato primario, y no solo la mera transcripción o la glosa, cuya lectura no siempre resulta transparente para el análisis de las producciones en LS. En este sentido, conviene tener en cuenta que en la actualidad se promueve un enfoque sistemático de la glosa dentro de los corpus de LS, conocido como glosa de identificación ID (Johnston y Schembri, 2010), el cual ha hecho posible la identificación de unidades y de tipos lingüísticos convencionales mediante identificadores de señas únicas. Esta glosa ID se utiliza para representar la seña en su forma de cita junto con todas sus variantes fonológicas y morfológicas, a partir del proceso de lematización (Cormier *et. al.*, 2012). Para ilustrar el empleo de la glosa ID, tomamos del sitio de ASL signbank (<https://aslsignbank.haskins.yale.edu/signs/search/>), la palabra *banana* que presenta dos formas de realización, como se muestra en la figura 2; no obstante, ambas producciones aparecen identificadas con la misma etiqueta glosa ID BANANA, como se muestra a continuación.

Figura 2
Glosa ID BANANA

Annotation ID	Gloss	Lemma ID	gloss	Translations	Handedness	Dominant Handshape	Nondominant Handshape	Location - major
	BANANAfull	BANANA		banana, fruit	Asymmetrical/DifferentHandshape	-	1	-
	BANANAix	BANANA		banana, fruit	Asymmetrical/SameHandshape	-	N/A	-

Asimismo, autores como Fenlon *et al.* (2015) y Quer y Steinbanch (2019) consideran que otros de los retos en la creación de un corpus es la representatividad de los datos primarios y el procedimiento para su obtención. Observan que la problemática no radica en que las producciones realizadas provengan de un número pequeño de colaboradores sordos, sino en el hecho de que la anotación sobre el perfil sociolingüístico de los participantes a menudo descansa bajo la etiqueta de señantes nativos y, por ende, lleguen a obviarse otras características que deben ser descritas de manera puntual, las cuales coadyuvarían, por ejemplo, en posibles explicaciones sobre la variación de la LS.

La heterogeneidad lingüística de los miembros de la CS, es decir, su mayor o menor competencia en la LS, se debe a varias razones, entre ellas la adquisición tardía de la LS, debido a que la gran mayoría de las personas sordas proviene de hogares oyentes, donde el aprendizaje de la LS no es la primera opción para sus hijos sordos. Así, aun cuando en el desarrollo de una investigación se pretenda contar con un mayor número de sordos señantes nativos para la recolección del corpus, existe una alta probabilidad de que en ella participe un número reducido de personas sordas provenientes de familias sordas. Por tanto, resulta pertinente conocer si los colaboradores sordos pertenecen o no a una familia sorda, la adquisición y exposición temprana a la LS, los años de escolaridad y las características del sistema educativo donde llevó a cabo sus estudios, así como si mantiene o no contacto con la comunidad sorda (Lucas *et al.*, 2001). Conocer estos datos contribuyen a diseñar y explicar los resultados de las tareas de elicitación, de juicios de gramaticalidad y de obtención de datos de discursos espontáneos.

En definitiva, defendemos la necesidad de poder contar con un corpus lingüístico moderno para el estudio de la LS, y no solo como una primera lengua, es más, resaltamos su valor dentro de la investigación de la LS como L2. Así, por ejemplo, contar con un corpus permitirá revisar el uso de los rasgos no manuales (gramaticalizados) en la

producción de una LS como L1 o como L2; y, no menos importante, abrirá la posibilidad de crear un diccionario, al ofrecer como soporte una base de datos léxica (Johnston, 2001). No obstante, reiteramos la juventud de la lingüística de corpus de LS. Cabe destacar que el primer proyecto de corpus de LS con las características que se han mencionado se inició en Australia en 2004 con un archivo digital de grabaciones de señantes sordos nativos o casi nativos de la Lengua de Señas Australiana (Auslan).

2. Signbank

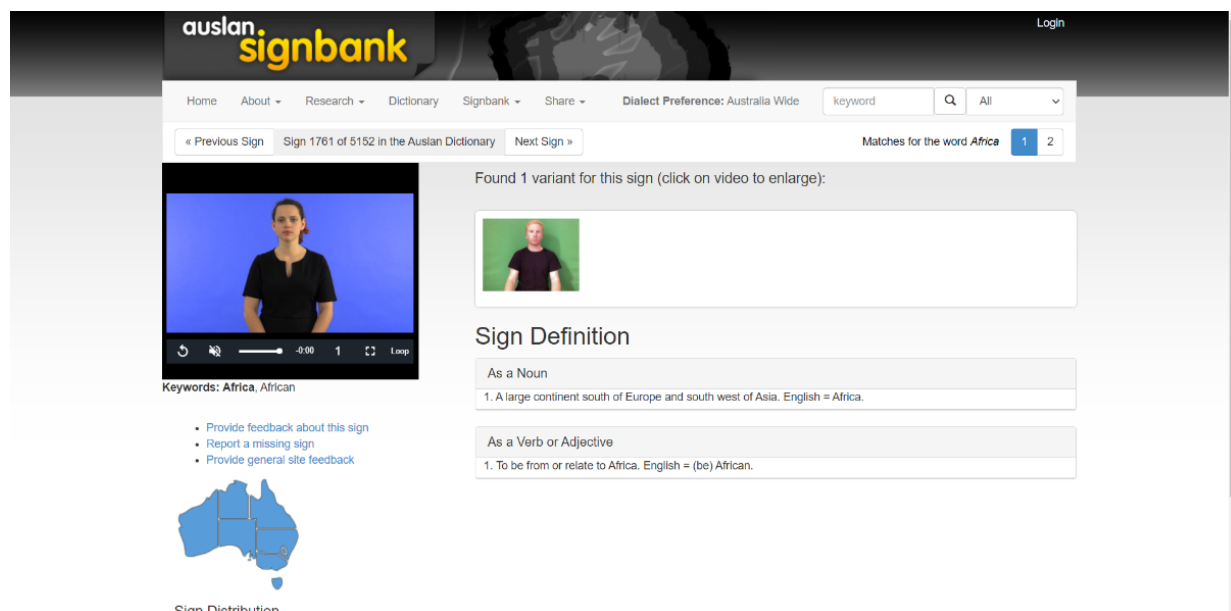
Un factor importante en el crecimiento de la lingüística de corpus de LS ha sido la formación de grupos de trabajo multidisciplinarios que han tenido como objetivo empatar el avance en el análisis gramatical de la LS con la creación de un *software* específico para tener una base de datos cuya organización y diseño posibilite la incorporación de las características fundamentales de las lenguas de modalidad visogestual: el uso de los articuladores activos (manos), la cabeza, el cuerpo, los gestos faciales y el espacio señante.

Un ejemplo del desarrollo de este joven campo disciplinar lo encontramos en la aplicación web llamada *Signbank* (<https://signbank.cls.ru.nl>). Este recurso fue originalmente construido para apoyar al diccionario *Auslan Signbank* (<http://www.auslan.org.au>) un diccionario en línea de Auslan, el cual tiene como antecedente el trabajo de documentación y análisis sobre Auslan, realizado por Johnston (2001), quien elaboró una base de datos léxica empleando la aplicación *Filemaker pro* (<https://es.wikipedia.org/wiki/FileMaker>). Esta base de datos ha servido como un recurso para informar del estudio de Auslan, así como de las convenciones de anotación de videograbaciones que forman parte del corpus de esta LS.

De manera general, podemos decir que *Signbank* provee de un marco para el desarrollo de una base de datos léxica de una LS determinada, a la cual se le asocian videos que ilustran la realización de estas señas. Como hemos destacado al inicio del artículo, el proyecto original *Signbank* ha sido la base de una serie de nuevos diccionarios de lengua de señas y de corpus de otras LS.

Los distintos sitios web de *Signbank*⁷ sirven de apoyo a diversos usuarios interesados en la LS. En algunos casos, se presenta un diccionario público de LS con un servicio de búsqueda que permite a los usuarios encontrar videos de señas que coincidan con palabras clave. Los usuarios registrados también pueden informar a los editores del banco de señas sobre las señas que faltan o errores en las entradas de señas existentes. De este modo, *Signbank* actúa como un recurso para la comunidad sorda que ayuda a construir un diccionario compartido de la lengua. En la figura 3 se ilustra una entrada del diccionario perteneciente a Auslan *signbank* (<https://auslan.org.au/dictionary/words/Africa-1.html>), en la cual se observa la articulación de la seña ÁFRICA, su definición, e incluso una variante de este signo.

⁷ BSL *Signbank* (<http://bsl.signbank.ucl.ac.uk/>); NGT *Signbank* (<https://signbank.cls.ru.nl/datasets/NGT/>); ASL *Signbank* (<https://asl.signbank.yale.edu/>); FinSL *Signbank* (<https://signbank.csc.fi/>); *Signbank da Libras* (<http://signbank.libras.ufsc.br/>).

Figura 3
ÁFRICA

Otra clase de usuarios son los investigadores, tanto aquellos interesados en utilizar *Signbank* como recurso léxico en sus investigaciones, como quienes quieren contribuir a enriquecer los modelos léxicos del diccionario. Estos investigadores tienen acceso a un conjunto de datos mucho más rico que el acceso público y pueden utilizar una función de búsqueda avanzada en los datos. Algunos investigadores proporcionan información detallada sobre los signos a los editores y pueden ayudar en la construcción de entradas de señas. Otro grupo de usuarios son los editores de *Signbank*, que pueden crear nuevas señas y subir videos asociados a estas. En conjunto, se forma una rica comunidad de usuarios de la lengua de signos que colaboran en la construcción de un recurso de apoyo a su lengua.

Signbank, disponible en <https://auslan.org.au/>, es una aplicación escrita en el lenguaje Python (<https://www.python.org/>) y utiliza la plataforma Django (<https://www.djangoproject.com/>) para la gestión de páginas WEB. Django es una infraestructura de alto nivel en Python que permite el desarrollo rápido y transparente mediante un diseño pragmático creada por expertos que facilitan el diseño para centrarse en la implementación de *Signbank*.

El núcleo de la base de datos léxica *Signbank* es un modelo léxico con el que pueden darse varios niveles de descripción de la lengua. La estructura inicial, desarrollada en FileMaker, se actualizó para permitir un número arbitrario de palabras clave asociadas a las señas.

El modelo de datos de *Signbank* contiene las siguientes entidades: glosa, relación, video, definición, traducción, palabra clave, dialecto y lenguaje. La entidad central del modelo es la seña, representada por la glosa, cada entrada léxica se asocia a una glosa de texto y cada una tiene un identificador de glosa ID (ID Gloss) que se utiliza para referirse a la entrada. Para una mayor descripción de la seña se pueden asignar etiquetas (*tags*) para acceder a tipos particulares

de vocabulario (salud, educación, deportes, etc.), pero también para describir características fonológicas o morfológicas. Sirva para ilustrar este diseño la imagen tomada de Global Signbank (<https://signbank.cls.ru.nl/>).

Figura 4
Configuraciones manuales

Handshape	Name	Selection	Configuration	Selection 2	Configuration 2	Quantity	Extended	Spreading	Aperture
	1	I	extended	-	-	one	-	N/A	N/A
	1_curved	I	curved	-	-	one	-	N/A	N/A
	3	IMR	extended	-	-	three	-	Spread	N/A
	3_curved	IMR	curved	-	-	three	-	Spread	N/A

En la figura 4 podemos observar que el acceso a los datos fue a partir de las características de la configuración manual (CM), definido a partir del bloque de dedos seleccionados, por ejemplo, el dedo índice, identificado con el número 1, o la CM en la cual interviene el dedo índice, el medio y el anular, etiquetado con el número 3. A ambas CM se les agrega un rasgo más para indicar que las articulaciones se encuentran extendidas o por el contrario presentan una ligera flexión (curva).

Por otra parte, cabe mencionar que la información que se ofrece implica que se ha pensado en el público que accede a la consulta de *Signbank*, es decir, este recurso permite dos maneras de consulta: una vista detallada dirigida a los investigadores de lenguas de señas y una vista más restringida dedicada al público en general.

Signbank es una aplicación de código abierto que permite que otros grupos de investigación implementen el banco para otras LS. Si bien esta condición ha beneficiado el desarrollo de distintos proyectos de corpus, también ha provocado algunos inconvenientes o limitantes, pues al ser grupos de trabajo independientes crearon sus propias normas. Para atender a esta problemática se creó la organización Github (<https://github.com/signbank>), con el objetivo de unificar las diferentes plataformas, proporcionando recursos informáticos para la elaboración de las bases de datos.

El *Signbank* de Finlandia (FinSL Signbank <https://signbank.csc.fi>) introdujo el concepto de *data set* (conjunto de datos) y posteriormente lo adoptó el *Signbank* de Países Bajos (<https://signbank.cls.ru.nl/datasets/NGT>). *Data set*

es un concepto útil en el manejo de grandes cantidades de datos, que se organizan generalmente de forma tabulada, de manera que en cada columna se especifica una variable y en los renglones los ejemplos de estas. Esta organización de los datos permite, entre otras cosas, que se puedan documentar varias lenguas con la misma herramienta. Los *data sets* permiten también que pueda controlarse si los datos son públicos o privados, y pueden asociarse con una lengua de señas y con múltiples lenguas escritas a fin de seleccionar la traducción equivalente a la lengua del *data set*. Esto es un paso importante hacia la constitución de un *Signbank Global* que pueda manejar un amplio rango de lenguas de señas en una única base de datos léxica.

NGT y posteriormente FinSL implementaron facilidades para exportar la información léxica a ELAN, como un vocabulario externo controlado (ECV *external controlled vocabulary*). Esto permite abreviar los procesos de anotación de los videos y hacerlo de manera más confiable ya que pueden referirse las entradas de *Signbank* a las convenciones de ELAN. Por último, es pertinente mencionar que NGT y ASL *Signbank* añadieron otras funciones que permiten enriquecer el modelo de la morfología en comparación con el original.

En el siguiente apartado describo de manera breve las características de la aplicación *Signbank* en el proyecto *ASL Signbank* y su impacto en el futuro de la investigación sobre la LS como L2.

3. La plataforma ASL Signbank

3.1. Descripción de ASL Signbank

ASL Signbank es un vocabulario en ASL cuyas señas se encuentran ligadas a identificadores de glosas (ID). Este proyecto se inicia en el 2007. Su objetivo es constituirse en una herramienta de anotación de videos de la ASL siguiendo las convenciones de ELAN y SLAASH (Sign Language Annotation, Archiving and Sharing)⁸. El acceso a la plataforma es a través de la siguiente dirección: <http://aslsignbank.haskins.yale.edu>

Veamos la sección titulada Señas/Signos. Esta se compone de dos apartados: Búsqueda (*search*), y Mostrar todas las señas. Si optamos por el primer criterio, encontramos que se presenta una plantilla para especificar las condiciones de la búsqueda como se muestra en la figura 5:

⁸ En la página principal podemos encontrar un subapartado denominado *Manual*, el cual remite a videos en YouTube en donde se explica el propósito, la estructura y la manera de acceder y contribuir al acervo.

Figura 5

Plantilla de búsqueda [Señas/Signos] ASL signbank

La plantilla que se observa en la figura 5 se compone de las siguientes gradas:

- 1) *Glosa en inglés, Lema en inglés, Traducciones al inglés.*
- 2) *Búsqueda por propiedades básicas de la lengua:* lengua de señas, dialecto regional, campo semántico, etc.
- 3) *Búsqueda por morfología,* glosa que contenga un morfema determinado, o que se parte de él o por un tipo de morfema.
- 4) *Búsqueda por fonología:* mano activa, locación mayor 1 y 2, mano dominante según los dedos seleccionados o los que se flexionan. Cambios por abducción o por flexión. Configuración de la mano no dominante. Trayectoria.
- 5) *Búsqueda por sintaxis y semántica.* Clase de palabra, categoría léxica 2, notas sobre la categoría léxica, historia de la derivación, tipo de iconicidad.
- 6) *Búsqueda por relación.* Por la glosa de señas relacionadas y tipo de relación (homónimo, sinónimo, variante, antónimo, hipónimo, hiperónimo, todas o ver también otra entrada). Búsqueda por glosa de una seña extranjera.
- 7) *Búsqueda por estado de la publicación o notas.*

En el caso de la opción “mostrar todas las señas” (*Show all signs*), se despliegan todas las señas con un video adjunto y datos sobre la glosa, el lema, la traducción y otros datos sobre la fonología. Cabe destacar que existen dos celdas de búsqueda: una por *glosa* y otra por la *traducción*. En muchos casos se puede buscar por la letra inicial, final o intermedia.

3.2. Aplicaciones de ASL Signbank

ASL *Signbank* es una herramienta para la documentación lingüística que abre el camino a múltiples aplicaciones. Si bien se compone de casi tres mil señas, este no debe de ser considerado como un diccionario, aunque esta base léxica sí pudiera dar lugar (en un futuro) a un diccionario en línea. Asimismo, aunque esta aplicación por sí sola no es un medio para aprender una LS, considero que ASL *Signbank* puede contribuir a la reflexión sobre los componentes, fonéticos y morfosintácticos, de la ASL como L2 si se tiene conocimiento de una primera LS

También, al incorporar variantes de una misma seña, se puede obtener una imagen más completa de los patrones de uso de la LS. Es decir, hay señas que hacen referencia a la misma entidad pero su uso responde a determinados contextos (cf. figura 2). Habría que agregar que en las lenguas de signos se ha reconocido variantes de algunas señas cuyo uso responde a regiones geográficas, género, o escolaridad. Por tanto, nos parece relevante el poder contar con el registro de esas variantes y las notas que acompañan al empleo de estas señas.

De igual manera, en el caso de los usuarios de una LS emparentada con la ASL, es posible una investigación sobre la producción, significado y uso, dado que se pueden encontrar señas que tienen los mismos parámetros articulatorios y tienen el mismo significado. Un ejemplo que ilustra este fenómeno es la seña empleada para referirse a ‘un poco’, cuya articulación (configuración manual) en LSM se ilustra en la siguiente figura.

Figura 6

Glosa ID LSM UN POCO



La búsqueda de dicha seña en las diferentes LS emparentadas, LSM, LIBRAS y ASL, a partir del uso de *Signbank*, nos revela que se trata del mismo signo para las tres lenguas como se muestra en la figura 7 para ilustrar su producción en LIBRAS (glosa ID POUCO)⁹ POUCO, y en la figura 8 la seña empleada en ASL (glosa ID A-LITTLE-BIT).

⁹Se puede consultar aquí: <https://signbank.libras.ufsc.br/#/busca-por-palavra/>

Figura 7
Glosa ID LIBRAS POUCO

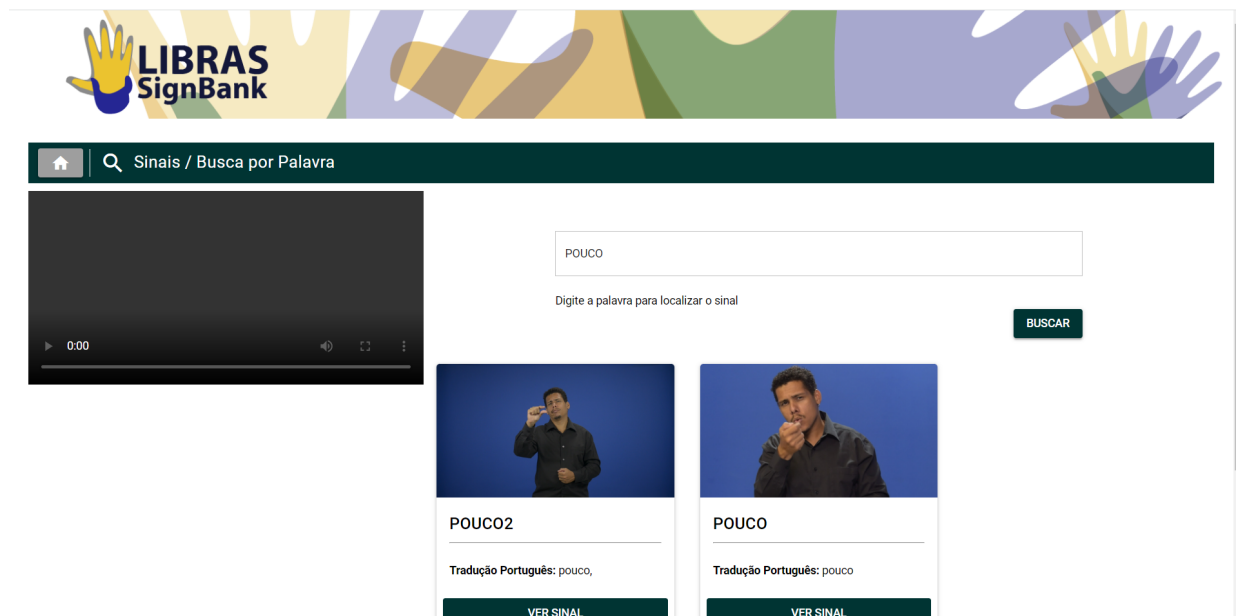
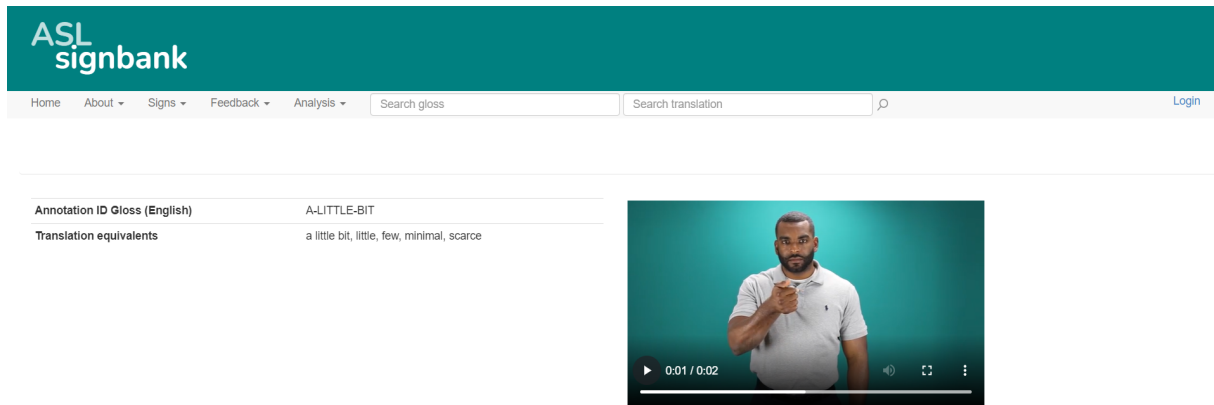


Figura 8
Glosa ID ASL A-LITTLE-BIT



[Provide feedback on the site](#)

Como se puede observar, en las tres LS el signo empleado para referirse a ‘un poco’ se realiza con los mismos parámetros articulatorios mayores: configuración manual, movimiento y ubicación. Este dato abona al conocimiento de las familias de lenguas de señas y apunta a la necesidad de impulsar la lingüística de corpus en el estudio de otras LS, así como en el uso de este tipo de aplicaciones.

3. Reflexiones finales

Uno de los puntos cuestionables de ASL *Signbank* es la decisión de no incorporar en los videos donde se ilustra la realización de las señas el componente de rasgos no manuales (RNM), entre ellos los gestos realizados con la cara. Si bien se reconoce su importancia, los creadores consideran que la inestabilidad de los gestos o movimientos realizados con la cabeza, es decir, la gran variabilidad de estos al producir una seña con los articuladores activos hace que en este momento de desarrollo de esta aplicación no sea pertinente su inclusión. Esto constituye una limitante para el estudio de las LS como L2, pues los RNM dentro del sistema gramatical de la LS son fundamentales en varios aspectos, por ejemplo, para la distinción de algunos pares mínimos; en el uso de aumentativos o diminutivos; así como para distintos aspectos relacionados con la pragmática del discurso, que requiere de ser utilizados. Por otra parte, nos parece que el no poder acceder a las señas de manera visual y solo sea posible a partir del texto escrito, constituye una limitante para el usuario. Aun cuando se ofrece un índice compuesto por una serie de imágenes correspondientes a las configuraciones manuales documentadas, este no es un motor de búsqueda, y requiere, por tanto, que el usuario tenga un conocimiento de las convenciones empleadas para facilitar la consulta de las entradas léxicas. El uso de la imagen en lugar del texto ya se ha implementado en LIBRAS (<https://signbank.libras.ufsc.br/#/sinais/configuracao-de-mao>), así que es probable que un futuro en ASL *Signbank* se incorpore este recurso.

Asimismo, nos parece oportuno discutir que para el aprendizaje de la ASL como L2 también sería necesario incorporar las definiciones, como en AUSLAN, pero no solo en formato texto, sino también en video. En el caso de nuestro país, se ha reconocido que una gran parte de la comunidad sorda no ha dominado la lengua escrita, y tiene dificultades en la recuperación de la información del texto escrito, esto significa que para muchas personas sordas el hecho de que no haya definición de la seña en LS es una dificultad para entender el significado de las señas. La situación se vuelve más compleja para los sordos usuarios de la LSM (L1) que quieren aprender ASL como L2, si el texto se encuentra en inglés (L3). A parte de que muchas glosas no se corresponden de forma exacta de una lengua a otra.

Por último, resaltamos el hecho de que la aplicación *Signbank* es un ejemplo del desarrollo de la lingüística de corpus en la LS. La posibilidad de contar con un corpus robusto y representativo, con la colaboración de sordos nativos o competentes en LS, es fundamental para una mayor comprensión de la estructura de estas lenguas. Asimismo, al constituir un registro del estado de la lengua, sus datos primarios permitirán identificar el cambio lingüístico de esta.

Al proveer de una base de datos léxica, *Signbank* ofrece un insumo que contribuye al estudio de las LS como L2, a partir de la identificación de los rasgos que los aprendices –sordos y oyentes– de una LS como segunda lengua emplean en su dominio paulatino de esta lengua. El avance en este nuevo campo de investigación impactará en la educación de las personas sordas, al igual que en la formación y capacitación de profesores e intérpretes de lengua de señas.

Referencias

- Barberà Altimira, Gemma (2015). *The meaning of space in Sign Language: reference, specificity and structure in Catalan Sign Language Discourse*. De Gruyter Mouton & Ishara Press.
- Bonet, Juan Pablo (1620). *Reducción de las letras y arte para enseñar a hablar a los mudos*. Jacobo de Orellana Garrido y Lorenzo Gascón Portero (eds.). La edición facsimilar completa puede bajarse en la página del Instituto Cervantes bajo <https://www.cervantesvirtual.com/nd/ark:/59851/bmchx1b1>
- Brentari, Diane (2019). *Sign Language Phonology*. Cambridge University Press.
- Chen Pichler, Deborah; Koulidobrova, Elena (2016). Acquisition of sign language as a second language. En Marc Marschark y Patricia Elizabeth Spencer (eds), *The Oxford handbook of deaf studies in language* (pp. 218-230). Oxford University Press.
- Cassidy, Steve; Crasborn, Onno; Nieminen, Henri; Stoop, Wessel; Hulsbosch, Micha; Even, Susan; Komen, Erwin; Johnston, Trevor (2018). Signbank: Software to Support Web Based Dictionaries of Sign Language. En *Proceedings of the Eleventh International Conference on Language Resources and Evaluation*. European Language Resources Association.
- Cormier, Kearsy; Jordan, Fenlon; Trevor Johnston; Ramas Rentelis; Adam Schembri; Kate Rowley; Robert Adam; Bencie Woll. (2012). From corpus to Lexical Database to Online dictionary. Issues in Annotation of the BSL Corpus and the development of BSL Signbank. En Crasborn Onno, Eleni Efthimiou, Evita. Fotinea, Thomas Hanke, Jette Kristoffersen y Joana Mesch (eds), *Proceedings of the 5th Workshop on the Representation and Processing of Sign Languages: Interactions between Corpus and Lexicon [Workshop Part of 8th International Conference on Language Resources and Evaluation, Lrec 2012, Istanbul, Turkey]* (pp. 7-12). ELRA. http://www.lrec-conf.org/proceedings/lrec2012/workshops/24.Proceedings_SignLanguage.pdf
- Crasborn, Onno (2006). Non-Manual Structures in Sign Language. En *Encyclopedia of Language & Linguistics* (2a. ed., pp. 668-672), vol.8. Elsevier. <https://hdl.handle.net/2066/42508>
- Crasborn, Onno; van der Kooij, Els; Waters, Daffyd; Woll, Bencie; Mesch, Joana. (2008). Frequency distribution and spreading behavior of different types of mouth actions in three sign languages. *Sign Language and Linguistics*, 11(1), 45-67. <https://doi.org/10.1075/sll.11.1.04cra>
- Cruz-Aldrete, Miroslava; Serrano Morales, Julio César; Medrano, Víctor Hugo (2020). “Lengua de Señas Mexicana y American Sign Language: efectos del contacto lingüístico en lenguas visogestuales”. En María Ángeles Soler Arechalde y Julio César Serrano Morales (eds.), *Contacto lingüístico y contexto social: estudios de variación y cambio* (pp. 379-400). UNAM.
- Fenlon, Jordan; Schembri, Adam; Johnston, Trevor; Cormier, Kearsy (2015). Documentary and corpus Approaches to sign language research. En Eleni Orfanidou, Bencie Woll y Gary Morgan (eds.), *Research Methods in Sign Language Studies: A Practical Guide* (pp. 156-172). John Wiley & Sons. <https://doi.org/10.1002/9781118346013.ch10>
- Haviland, John B.; Flores Farfán, José Antonio (coords.) (2007). *Bases de la documentación lingüística*. Instituto Nacional de Lenguas Indígenas.
- Herrero Blanco, Ángel (2002). La investigación lingüística de las lenguas de signos. *Lynx Panorama de Estudios Lingüísticos* 1, 9-47.
- Hervás y Panduro, Lorenzo. (1795) *Escuela Española de Sordomudos o Arte para enseñarles a escribir y hablar el idioma español, dividida en dos tomos*. Madrid: Imprenta Real. La edición facsimilar completa puede bajarse en la página del Instituto Cervantes, bajo <http://www.cervantesvirtual.com/FichaObra.html?Ref=10702&portal=20>

- Johnston, Trevor (2001). The lexical database of Auslan (Australian Sign Language). *Sign Language and Linguistics*, 4(1-2), 145-170. <https://doi.org/10.1075/sll.4.12.11joh>
- Johnston, Trevor; Schembri, Adam (2010). Variation, lexicalization and grammaticalization in signed languages. *Langage et societé*, 131(1), 19-35. <https://doi.org/10.3917/ls.131.0019>
- Liddell, Scott K. (2003). *Grammar, Gesture, and meaning in American Sign Language*. Cambridge University Press.
- Liddell, Scott K.; Johnson, Robert. E. (1989). ASL: The phonological base. *Sign Language Studies*, 64, 195-277.
- Lucas, Ceil; Bayley, Robert; Valli, Clayton (2001). *Sociolinguistic Variation in American Sign Language* (vol. VII). Gallaudet University Press.
- Minguet Soto, Amparo (2001). *Introducci3n la lingüística a la LSE*. FESORD.
- Nonaka, Angela M. (2004). The forgotten endangered languages: Lessons on the importance of remembering from Thailand's Ban Khor Sign Language. *Language in Society*, 33, 737-767. <http://dx.doi.org/10.1017/S004740450404504X>
- Prillwitz, Siegmund; Zeinert, Heiko (1989). Hamburg Notation System for Sign Language. Development of a sign writing computer application. *Proceedings 3rd. European congress on Sign Language Research*, 355 -379.
- Quer, Josep; Steinbach, Markus (2019). Handling Sign Language Data: The Impact of Modality. *Front. Psychol*, 10, 483. <https://doi.org/10.3389/fpsyg.2019.00483>
- Schönström, Krister. (2021) Sign languages and second language acquisition research: An introduction. *Journal of the European Second Language Association*, 5(1) 30-43.
- Stokoe, William C. (1960). Sign Language Structure. An outline of the visual communication systems of the American Deaf. *Studies in Linguistics, Occasional Papers*, 8. University of Buffalo.
- Stokoe, William C. (1976). *A Dictionary of American Sign Language*. Linstok Press.
- Stokoe, William C. (1978). *Sign Language Structure: The First Linguistic Analysis of American Sign Language*. Linstok Press.
- Stokoe, William C. (2001). *Language in hand: Why sign came before speech*. Gallaudet University Press. (Versi3n en espa3ol. *El lenguaje en las manos*. Fondo de Cultura Econ3mica, 2004)
- Vera Villaverde, Francisco (2020). Las lenguas signadas como objeto de estudio lingüístico. *Revista de Estudios de Lenguas de Signos*, 2, 76-82. <http://hdl.handle.net/11181/6108>
- Zeshan, Ulrike; de Vos, Connie (2012). *Sign Languages in Village Communities. Anthropological and Linguistic Insights*. De Gruyter.