# STK LSF: A Motion Capture Dataset in LSF for SignToKids

**Clément Reverdy, Sylvie Gibet, Thibaut Le Naour**
IRISA, Université Bretagne Sud
Vannes, France
{clement.reverdy, sylvie.gibet}@irisa.fr

**Abstract**

This article presents a new bilingual dataset in written French and French Sign Language (LSF), called *STK LSF*. This corpus is currently being produced as part of the SignToKids project. The aim of this corpus is to provide digital educational tools for deaf children, thereby facilitating the joint learning of LSF and written French. More broadly, it is intended to support future studies on the automatic processing of signed languages. To define this corpus, we focused on several grammatical phenomena typical to LSF, as well as in tales usually studied by hearing children in the second cycle in France. The corpus data represent approximately 1 hour of recording, carried out with a motion capture system (MoCap) offering a spatial precision of less than 1 mm and a temporal precision of 240 Hz. This high level of precision guarantees the quality of the data collected, which will be used both to build pedagogical scenarios in French and LSF, including signing avatar videos, and for automatic translation of text into LSF.

**Keywords:** French Sign Language, LSF, corpus, motion capture, grammatical utterances

## 1. Introduction

The aim of the *SignToKids* project is to build digital pedagogical tools for the joint learning of French sign language (LSF) and written French for deaf children. Of course, this work cannot cover all the educational needs to be made available to the Deaf, but it does provide an initial response to this very ambitious issue.

While there are a number of studies presenting the various aspects of French sign language (LSF) grammar (Cuxac, 2000; Millet, 2019), there is currently no educational book or digital application enabling schoolchildren to learn both the grammar of LSF and that of written French, which seems necessary for access to the various subjects taught in schools. Nor is there any specific bilingual LSF/French corpus built specifically for this purpose.

Through this project, it therefore proved necessary to define a specific bilingual corpus adapted to the demands of LSF and French teachers for deaf children in primary and secondary schools, corresponding to cycles II and III for hearing children. Our objectives are to: 1) make it easy to work out the grammatical structures common to and specific to each language; 2) help the child understand how to express the same concept in both languages; 3) correspond to the expectations of the cycle's curricula (`https://eduscol.education.fr/127/langue-des-signes-francaise`).

As part of our digital tools, rather than using videos in LSF, we have chosen to use virtual signing characters, which we call signing avatars. In addition, in order to obtain a high degree of precision in the movements produced, and to ensure the quality of the pedagogical exercises to be built, we opted for the use of motion capture data (MoCap). Furthermore, as the project is ambitious, both linguistically and technically, we decided to build the dataset in four phases, the first to adjust the corpus construction methodology and capture protocol, then to provide data with an increasing level of complexity.

This article describes the first two parts of the corpus, called LSF-STK1 corpus. It includes a set of phrases in written French and LSF, as well as the corresponding MoCap data.

## 2. Related work

Signed languages are visual and gestural languages. Consequently, the two main techniques for effectively capturing the sign language gestures are video and motion capture (MoCap). The corresponding two types of data do not entail the same costs or quality, either in terms of data acquisition or post-processing, and give rise to different analysis and processing possibilities.

### 2.1. Video corpora

Video recording devices (RGB or RGBD) are inexpensive, easy to set up and not very intrusive for the people being recorded. In addition, new tools (Kartynnik et al., 2019; Cao et al., 2017) have recently been developed to infer human postures (poses of the skeleton) and facial expressions from 2D images, making it possible to obtain gesture-type information. In addition, recent advances in computer vision (e.g. SMPL-X(Pavlakos et al., 2019)) make it possible to infer 3D meshes, blendshapes, skeletons and their animation parameters (joint orientation, blendshape coefficients, etc.) from human video recordings.



*Proceedings of the 11th Workshop on the Representation and Processing of Sign Languages*, pages 315–322
25 May 2024. 

Figure 1: A few postures of our signing avatar.

Video data is often the basic material for linguistic analysis and automatic processing of sign languages (e.g. automatic recognition of signed sentences, automatic translation of a sign language into the corresponding written language) using machine learning algorithms (e.g. deep neural networks).

Among the large-scale projects that have emerged over the last decade, several initiatives are worth mentioning. However, as the subject of this paper is an LSF mocap dataset, only video corpora dealing with LSF or Belgian French sign language (LSFB) are listed here. For other SLs, the reader is invited to refer to the Sign Language Dataset Compendium (Kopf et al., 2022), which provides a list of most existing video corpora up to 2022 and their main characteristics.

There are two main corpora of LSF. The Mediapiskel (Bull et al., 2020) is a ~27h corpus performed by more than 100 signers with a vocabulary size larger than 17k. It is signed by deaf journalists for a TV journal. The data are annotated with aligned written French subtitles. The CREAGEST (Balvet et al., 2010) is a corpus of ~500h signed by ~250 signers (adults and children), recorded in studio conditions, and elicited by various tasks. Within this corpus, only ~1h is annotated[1].

The LSFB dataset (Fink et al., 2021) is a corpus in Belgian French sign language. It contains ~90h of videos, performed by 100 signers. ~25h data are recorded in studio conditions with a vocabulary of ~7k words. Elicitation was carried out by asking the signers to perform various tasks, leading to spontaneous discourse. These data are annotated with glosses and written French translations.

## 2.2. MoCap corpora

MoCap systems make it possible to record human gestures with a degree of precision, consistency and robustness not yet possible with video devices. Indeed, MoCap technologies have spatial accuracy in the millimeter range and frequency accuracy in the hundreds of Hertz range (typically 60 to 200 Hz for SL (Lefebvre-Albaret et al., 2013)). They are also less prone to occlusion problems than monoview devices. The limitations of this technology are characterized by: i) the need for data postprocessing to reconstruct trajectories and produce skeletal pose sequences, which requires a considerable human investment (around one working day per minute of recording), and ii) the complexity of setting up this device, which limits its use to laboratory environments. This makes recordings less flexible than those obtained with lighter, less intrusive devices such as video. As a result, creating large corpora of MoCap data is still too costly for the time being, both in terms of equipment and human labour.

Among the MoCap corpora collected over the last decade, several have been produced with the aim of analyzing SL data and performing data-driven synthesis.

CUNY ASL (Lu and Huenerfauth, 2010, 2014) is an American Sign Language (ASL) dataset performed by 8 signers with a total duration of ~3h30. Elicitation was made by a native ASL speaker sitting behind the camera who engaged a conversation with the recorded signer. The body and fingers movements (using cybergloves ®), as well as gaze direction were tracked but facial expressions were not recorded. The data were annotated with glosses and spatial references.

(Jedlička et al., 2020) have recorded a ~30min full-body (body, face and fingers) MoCap dataset, performed by one expert Czech Sign Language (CSE) signer who was asked to sign weather forecasts. Data were annotated with glosses.

HRI JSL full-body dataset (Brock and Nakadai, 2018) has been performed by one signer. About 10k signed utterance in Japanese Sign Language (JSL) were recorded. The signer, a Child Of Deaf Adults (CODA) was asked to sign predefined sentences. Data was annotated with sign-based glosses.

[1] according to https://www.sign-lang.uni-hamburg.de/lr/compendium/corpus/creagest.html dated January 23, 2023

In LSF, several corpora have been recorded. MOCAP1 (Braffort, 2016) has been created for motion and linguistic analysis. Limbs motion and significant LSF facial movements were tracked, but fingers motion was not recorded. LSF-rosetta full-body corpus (Bertin-Lemée et al., 2022) aims to produce LSF from *AZee* specification (Nunnari et al., 2018), with a signing avatar. ~3h has been recorded by one signer. Elicitation was carried out by asking the signer to perform 4 tasks (translation to LSF, description of images, repetition of LSF video clips, production of >1200 isolated signs). Data was annotated with glosses, phonological components and *AZee* descriptions.

Several full-body high resolution LSF datasets have been recorded since 2009 at IRISA [2], for LSF analysis and data-driven synthesis. For these datasets, skeletons were reconstructed from about 40 markers for the body, 40 smaller markers for the face, and 20 even smaller markers for each hand. Data-driven animations with a signing avatar were produced, after retargeting, rigging and skinning. LSF-SignCom (Duarte and Gibet, 2010) is a MoCap dataset of ~1.5h signed by one deaf signer. Based on a dialogue between two deaf signers, the movements of one of the two protagonists were recorded, with the second giving the cues. The corpus contains recipes and short stories on cooking themes (making salads, galettes, cocktails). Data was annotated on multiple tiers using glosses and phonological components. LSF-ANIMAL (Naert et al., 2020) is a full-body LSF dataset containing ~1h of data recorded on two deaf LSF professors fluent in written French. Elicitation was carried out with 3 main tasks: 1) isolated signs, 2) utterances illustrating grammatical mechanisms (pointing gestures, classifier predicates) and 3) Continuous signing (26 free descriptions of animals). Data was manually annotated with glosses on 3 channels (right hand, left hand and both). Phonological components (hand configuration, placement) were also annotated using automatic segmentation methods (Naert et al., 2018).

Several works have shown that data-based methods are capable of producing realistic LSF animations by concatenative synthesis. Thanks to a scripted language based on two coupled databases (motion and semantic annotations), new LSF utterances that respect LSF grammatical rules could be synthesized by editing and concatenating recorded motion, and used to animate a signing avatar (Gibet et al., 2011). Following the same approach, the *Sign3D* dataset has been recorded at Mocaplab. It contains utterances describing places of interest and events taking place in a city (Lefebvre-Albaret et al., 2013). Another recent project (Bertin-Lemée et al., 2022) follows a similar path.

The choice of MoCap therefore appears to be an appropriate solution for the recording of our corpus, both for the naturalness of the movements recorded, and for the precision of these movements, which meets the grammatical requirements of sign languages.

## 3. Corpus Design

The design of our corpus is based on the following approach: using a few a priori grammatical objectives, we construct a set of sentences which we then record (video, MoCap). But, rather than recording our corpus all at once, we have chosen to record four sub-corpuses over a longer period - STK1.1, STK1.2, STK2.1, STK2.2 - in order to draw technical, pedagogical and linguistic lessons as we go along.

This paper refers to the first two sub-corpora that have been already recorded. Our first bilingual LSF/French sub-corpus, STK1.1, is based on several grammatical objectives that LSF teachers consider useful for learning both LSF and French. Of course, for each grammatical target, not all the processes used in French or LSF are covered exhaustively. For the creation of this sub-corpus, we constructed a set of sentences, guided by the use of teaching resources defined by deaf teachers in LSF (Centre Gabriel Deshayes, Auray) or from Millet's descriptive grammar (Millet, 2019). Our second sub-corpus, called STK1.2, describes three tales in LSF.

### 3.1. Motivations for our Grammatical Targets in STK1.1

**Clausal aspects.** We first looked at the clausal form of sentences, i.e. negative, assertive and interrogative sentences.

The construction of negation is difficult for deaf children to learn, as it is expressed very differently in LSF and French. In French, negation is carried by the structures "ne ... pas", "ne ... plus", "ne ... jamais", etc. that surround the verb. In LSF, a distinction is made between negative sentences in which: 1) negation is marked by a specific lexical sign, generally placed at the end of the sentence, and those in which 2) negation is integrated into the verb. In addition, the expression of negation in LSF is generally accompanied by a facial mimic. The advantage of recording two logical forms of sentences in our data, positive and negative, enables us to construct simple exercises in which the negative form is requested from the positive form, in French or in LSF, and vice-versa.

Interrogative sentences are also used in exercises (for example question-and-answer exercises).

[2](Institut de Recherche en Informatique et Systèmes Aléatoires)

In this case, the question is direct, i.e. it is represented by a sign from the LSF lexicon. These signs can also be used in synthesis editing processes in rhetorical questions (false questions in LSF), which are not interrogative, but which serve to link propositions together. Here too, facial expressions are crucial, as they can be the unique mark of the interrogative clause as opposed to the assertive clause.

**Indicating verbs.** In STK1.1, we were interested in directional verbs that represent in French the syntactic structure Subject - Verb - Complement (direct or indirect object complements), with the possibility of representing subjects and recipients by pronouns.These verbs unfold along a trajectory in the signing space. This enables them to distribute syntactically the roles of the actants of the sentence (actants being the agent, the object, or the recipient). In this way, they move from one locus (spatial referent) to another. For example, the verb [TO GIVE] can be flexed along different trajectories from an agent locus to a recipient locus, and these actants can be pronouns positioned in the signing space. For example, the French sentence *Je te donne* (I give you) can be translated into LSF by a movement of the hand from a neutral zone near the torso (person 1) to a zone in front of the signer (person 2), while the sentence *Tu me donnes* is translated by an inverted trajectory. Some indicating verbs can also be flexed along the object. In this case, the configuration of the hand representing the object is modified. For example, the French sentences *Je te donne un verre* or *Je te donne un livre* are translated into LSF in the same way, except that the hand configuration changes to represent either the glass or the book.

The second STK1.2 sub-corpus is not associated with specific grammatical targets. It includes those of the STK1.1 corpus and other grammatical mechanisms typical to sign languages.

### 3.2. Corpus Content

**STK1.1** In the negation category, we have identified the signs [NON] (no), [JAMAIS] (never), [RIEN] (nothing), [Y-A-PAS] (nothing). Some signs, such as [NON], combine exclusively with verbs, and others with nouns, such as [Y-A-PAS]. For example, in the French sentences "Il ne boit pas" (He does'nt drink), "Elle ne boit jamais" (She never drinks), or "Je ne comprends rien" (I don't understand anything), negation is expressed in French by the words NE... PAS, NE... JAMAIS, NE... RIEN which surround the verb, whereas in LSF, this negation is expressed by a word at the end of the sentence ([NON], [JAMAIS], [Y-A-PAS]).

In the second negation category, we considered sentences with modalities, including: [NE-PAS-POUVOIR] (can't), [NE-PAS-VOULOIR] (won't), [NE-PAS-CROIRE] (don't believe), [NE-PAS-SAVOIR] (don't know), [NE-PAS-AVOIR-BESOIN] (don't need). We have also added the verb [NE-PAS-AIMER] (dislike). Par exemple, dans la phrase "Le garçon n'aime pas facebook" (The boy doesn't like facebook), la négation s'exprime en LSF par [FACEBOOK][DISLIKE], dislike being represented by a sign, negative form of like. For example, in the sentence "Le garçon n'aime pas facebook" (The boy doesn't like facebook), the negation is expressed in LSF as [FACEBOOK][NE-PAS-AIMER], the negation being incorporated into the verb [NE-PAS-AIMER] (dislike), whose trajectory is reversed relatively to the verb [AIMER] (like). In another example, the sentence in French "Il n'a pas besoin qu'on lui dise deux fois" (He doesn't need to be told twice) can be translated in LSF as [REPETER][DEUX][FOIS][IL][NE-PAS-AVOIR-BESOIN], where the negation is directly incorporated into the verb [NE-PAS-AVOIR-BESOIN].

In both negation categories, we have defined 16 positive and 16 negative assertive sentences for each verb, and we have selected 24 interrogative sentences.

For indicating verbs, we selected 128 sentences repeated twice. For example, the French sentence "Je te raconte une histoire" (I am telling you a story), can be syntactically modified by replacing the pronouns "Je" (I) and "te" (you), as in "Elle me raconte une histoire" (She is telling me a story), or "Tu lui racontes une histoire" (You are telling her a story). In LSF, the syntactic structure with pronoun changes results in indicating verbs whose trajectories are modified, with personal pronouns (Je, Elle, Tu) or complement pronouns (te, me, lui) resulting in "pre-semantized" Loci placed at specific places in the signing space. We have built a total of 128 different sentences with indicating verbs, declining the agent/object/beneficiary actants in each type of sentence. In the sub-corpus STK1.1, we have built a total of 220 phrases. All these sentences have been repeated twice in the recording session.

To this corpus we have added a list of words from a lexicon, which are chosen in such a way as to be able to construct new sentences in relation to the initial ones, thanks to the use of our concatenative synthesis system *SignCom* (Gibet et al., 2011; Naert et al., 2021).

Finally, we considered pointing to be fundamental to syntax, in particular to ensure reference. Pointing in LSF can be used in many semantic-syntactic contexts. In addition, there are many different ways of pointing. We have supplemented STK1.1 with a set of simple pointing gestures (index hand configuration) whose loci are randomly distributed in the signing space. About 40 pointing gestures have been executed.

**STK1.2** In order to have utterances more spontaneous (less controlled), with various grammatical structures proper to LSF (for example morpho-syntactic variations, iconic descriptions, static and dynamic classifiers), we produced STK1.2, composed of three tales that are usually studied in Cycle II for hearing children. The first two tales are French tales taken from the "Roman de Renart", a medieval collection of animal stories written by various authors. We have selected the two stories "Renard et la queue du loup" (Renard and the wolf's tail), and "Renard et les marchands" (Renard and the merchants). These tales were designed and adapted in French and in LSF, and validated by a deaf teacher in LSF. Both of them have a duration of about 10 minutes. Elicitation was then achieved by producing small written sequences (about 10 sentences per sequence) that tell the story, associated to gloses and illustrated by images. Sequences of questions are signed at the end of each tale.

The third tale follows another approach as it is directly signed in LSF by the deaf signer and then transcribed into written French. This tale is an adaptation of the tale "Le vilain petit canard" written by Hans Christian Andersen. Its duration is about 10 minutes. The three tales contain a total of 184 utterances.

## 4. Recording and Motion Dataset

Recording our corpus and motion dataset requires careful focus not only to the motion capture processing chain, but also to the development of tools for visualizing and editing motion, including the design of appropriate avatars.

### 4.1. Motion capture recording and processing

**Recording setup.** The *Motion-Up*[3] company was involved in the project to capture the data, model the 3D avatar and do the rigging and skinning. A total of one hour of data has been recorded over two sessions (STK1.1 and STK1.2). The MoCap system used for both recording sessions was an Optitrack 18-cameras "Prime 22" with reflexive markers, recording at 240 Hz.

To capture facial expressions, we explored two solutions: i) a MoCap-based solution, following the setup and approach described in (Reverdy, 2019), and ii) a commercial software (Faceshift (Weise et al., 2011)) that uses an RGBD video as input. Both options were tested during the initial session, and we chose to use the Faceshift solution exclusively for STK1.2, for reasons of simplicity and efficiency. Indeed, Faceshift provides, through a calibration process, a modeling of the human head with 51 morphtargets, and the automatic transformation of video into 51 blendshape curves.

In addition the scene was recorded by two RGB video cameras (60hz) from two different points of view in order to facilitate the annotation task. To make possible the synchronization between all devices, it was asked to the signer to perform at the beginning and the end of each recording sequence a specific mouth and hands movements.

**Markers setup.** 35 markers were placed all over the body, 40 on the signer's face (only during the first session) and 20 on each hand (a total of 75 markers on the body and hands). The number, size and shape of the markers used for each location is a trade-off between ease of tracking and comfort for the signer.

**Elicitation protocol.** The STK1 corpus was signed by a deaf signer who is also a theater actress. She participated in the design of the corpus content, in both LSF and French, and therefore had a good knowledge of it before each of the recording sessions. A slideshow of the sequences of sentences was projected at a distance of around 3 meters in front of the signer. For memorization purposes, she was instructed to sign the sentences presented in both written French and in the form of a sequence of glosses that she had previously transcribed and, where possible, illustrated with images.

The corpus was divided into sections of 1 to 4 slides, depending on the content of the corpus, with an average of 10 sentences per section. Each slide was composed of a set of sentences or isolated signs. Each section was repeated twice and recorded in one MoCap sequence. To achieve a certain quality of data, we made several takes for each section, but retained only one take for post-processing.

**Post-processing.** The data obtained after recording usually consists of a set of unlabeled marker trajectories. It is possible that, during recording, a marker is temporarily lost by the cameras for a short period (occultation), or that markers are exchanged along the trajectories.

Data cleaning involves labeling the markers along the trajectories and reconstructing the missing parts of the trajectories (gap-filling). In order to efficiently post-process the data, MotionUp has developed a software tool that reduces post-processing times by a 4x factor compared with the software *Motive 3.0.2 (Optitrack system)*. In particular, a new translation-and-rotation-invariant algorithm has been integrated into this Motion-Up software to automatically label the markers and reconstruct the trajectories. This very tedious task

[3]https://www.motion-up.com/


Figure 2: An overview of the editing and visualization application.

required 0.6 working days per minute recorded for the body and hands.

Another tedious task was to clean up the facial data, as the mono-view recording device was prone to occlusions. These gaps had to be filled manually to ensure the quality of the animation. This task took an average of 0.72 days per recorded minute.

### 4.2. Avatar design, Editing and Visualization Software

**Avatar Design.** The final objective of the *Sign-ToKids* project is to provide digital tools for learning both LSF and written French. From the point of view of educational exercise construction, we favor the inclusion of 3D virtual character animations as learning material, due to their interactive and playful aspect (possibility to modify the avatar's appearance) and the ease of editing this type of animation (e.g. camera movement, choice of reading speed).

The design of the avatar required careful technical consideration in order to limit the post-processing required for retargeting (adaptation to the signer's morphology) and to avoid animation artifacts that would damage the credibility and intelligibility of the signed gestures.

Morphologically, the joints and bone lengths have been finely designed to match the signer's skeleton. Furthermore, the face has been designed to be comparable to the signer's face, so that in the event of contact between hands and face, there is no gap or interpenetration.

**Editing and visualization software.** In addition to markers auto-labeling and gap-filing, Motion-Up software offers several other functionalities, including avatar visualization, editing and ultimate motion correction. An illustration is shown in Figure 2. This software features its own real-time iterative skeleton solver, so that the resulting animations can be played back, according to the selected SL sequence and the 3D avatar, and viewed in real time. The solver relies on an iterative process aimed at preserving the accuracy of spatial configurations such as contact between fingers, hands or face. Two mutually dependent steps were used, the first one for the palm transformation, the second for the fingers. This software also facilitates the editing of the recorded gesture afterwards, in order to correct certain undesirable behaviors (defaults generated by the MoCap installation, or signing defaults such as erroneous blinking or a systematic tendency to look in one direction).

## 5. Conclusion and Perspectives

This article presents an initial corpus STK1 and its corresponding dataset developed within the *Sign-ToKids* project. This corpus covers various grammatical phenomena essential to the joint learning of LSF and French. It also contains LSF adaptations of tales studied by hearing children aged 9 to 12. The chosen capture method is MoCap, due to the various objectives and constraints inherent in this project.

Approximately one hour of data was recorded and post-processed, using a motion capture system, including body and hand motion, as well as facial expression and gaze direction. The skeleton data was reconstructed, and a 3D avatar was modeled and rigged to this data. An application was then developed to visualize, correct and annotate the various animations of this avatar corresponding to the recorded data. The linguistic annotation of this corpus has yet to be carried out with an annotation scheme inspired by the approach developed for the German Sign Language (DGS) corpus (Hanke et al., 2020).

The second part of this corpus, completing the recordings, is scheduled for the end of 2024. In the near future, we plan to expand this corpus by automatically producing, through generative AI, French and LSF-glossed sentences, and then by automatically translating the glossed sentences into LSF.

## 6. Acknowledgements

The *SignToKids* project is funded by the French National Research Agency (ANR). Several research teams participate in this project: IRISA-CNRS lab., MoDyCo-CNRS lab., and the national institute for young deaf (INJS). We would also like to thank Nathalie Irdel, from the Morbihan Deaf Association, for her help in designing and recording the *STK LSF* dataset.

## 7. Bibliographical References

A. Balvet, C. Courtin, D. Boutet, C. Cuxac, I. Fusellier-Souza, B. Garcia, M.T. L'Huillier, and M.A. Sallandre. 2010. The creagest project: a digitized and annotated corpus for french sign language (lsf) and natural gestural languages.

É. Bertin-Lemée, A. Braffort, C. Challant, C. Danet, B. Dauriac, M. Filhol, E. Martinod, and J. Segouat. 2022. Rosetta-lsf: an aligned corpus of french sign language and french for text-to-sign translation. In *13th Conference on Language Resources and Evaluation (LREC 2022)*.

A. Braffort. 2016. Mocap1: A motion capture corpus of french sign language for interdisciplinary studies. In *Journées d'études DIG: La dynamique Interactionnelle du Geste" Making sense together"*.

H. Brock and K. Nakadai. 2018. Deep jslc: A multimodal corpus collection for data-driven generation of japanese sign language expressions. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*.

H. Bull, A. Braffort, and M. Gouiffès. 2020. Mediapiskel-a 2d-skeleton video database of french sign language with aligned french subtitles. In *Proceedings of the 12th Conference on Language Resources and Evaluation (LREC 2020)*, pages 6063–6068.

Z. Cao, T. Simon, S.E. Wei, and Y. Sheikh. 2017. Realtime multi-person 2d pose estimation using part affinity fields. pages 7291–7299.

C. Cuxac. 2000. *La langue des signes française (LSF) : les voies de l'iconocité (French) [French Sign Language: the iconicity ways]*. Faits de langues. Ophrys.

K. Duarte and S. Gibet. 2010. Heterogeneous data sources for signed language analysis and synthesis: The signcom project. In *7th Int. Conf. on Language Resources and Evaluation (LREC 2010)*, volume 2, pages 1–8. European Language Resources Association.

J. Fink, B. Frénay, L. Meurant, and A. Cleve. 2021. Lsfb-cont and lsfb-isol: Two new datasets for vision-based sign language recognition. In *2021 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE.

S. Gibet, N. Courty, K. Duarte, and T. Le Naour. 2011. The signcom system for data-driven animation of interactive virtual signers : Methodology and evaluation. In *ACM Transactions on Interactive Intelligent Systems*, volume 1.

T. Hanke, M. Schulder, R. Konrad, and E. Jahn. 2020. Extending the public dgs corpus in size and depth. In *sign-lang@ LREC 2020*, pages 75–82. European Language Resources Association (ELRA).

P. Jedlička, Z. Krňoul, J. Kanis, and M. Železnỳ. 2020. Sign language motion capture dataset for data-driven synthesis. In *Workshop on the Representation and Processing of Sign Languages (LREC2020)*, pages 101–106.

Y. Kartynnik, A. Ablavatski, I. Grishchenko, and M. Grundmann. 2019. Real-time facial surface geometry from monocular video on mobile gpus. *arXiv preprint arXiv:1907.06724*.

M. Kopf, M. Schulder, and T. Hanke. 2022. The sign language dataset compendium: Creating an overview of digital linguistic resources. In *Workshop on the Representation and Processing of Sign Languages (LREC2022)*, pages 102–109.

F. Lefebvre-Albaret, S. Gibet, A. Turki, L. Hamon, and R. Brun. 2013. Overview of the sign3d project high-fidelity 3d recording, indexing and editing of french sign language content. In *Third International Symposium on Sign*

*Language Translation and Avatar Technology (SLTAT) 2013*.

P. Lu and M. Huenerfauth. 2010. Collecting a motion-capture corpus of american sign language for data-driven generation research. In *Proceedings of the NAACL HLT 2010 workshop on speech and language processing for assistive technologies*, pages 89–97.

P. Lu and M. Huenerfauth. 2014. Collecting and evaluating the cuny asl corpus for research on american sign language animation. *Computer Speech & Language*, 28(3):812–831.

A. Millet. 2019. *Grammaire descriptive de la langue des signes française: dynamiques iconiques et linguistique générale*. UGA Editions.

L. Naert, C. Larboulette, and S. Gibet. 2020. Lsf-animal: A motion capture corpus in french sign language designed for the animation of signing avatars. In *Proceedings of The 12th Language Resources and Evaluation Conference*, pages 6008–6017.

L. Naert, C. Larboulette, and S. Gibet. 2021. Motion synthesis and editing for the generation of new sign language content. *Machine Translation*, 35(3):405–430.

L. Naert, C. Reverdy, C. Larboulette, and S. Gibet. 2018. Per channel automatic annotation of sign language motion capture data. In *Workshop on the Representation and Processing of Sign Languages: Involving the Language Community (LREC 2018)*, Miyazaki, Japan.

F. Nunnari, M. Filhol, and A. Heloir. 2018. Animating azee descriptions using off-the-shelf ik solvers. In *Workshop on the Representation and Processing of Sign Languages: Involving the Language Community (SignLang 2018), (LREC 2018)*, pages 7–12.

G. Pavlakos, V. Choutas, N. Ghorbani, T. Bolkart, A.A.A. Osman, D. Tzionas, and M.J. Black. 2019. Expressive body capture: 3d hands, face, and body from a single image. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 10975–10985.

C. Reverdy. 2019. *Annotation et synthèse basée données des expressions faciales de la Langue des Signes Française*. Ph.D. thesis, Université de Bretagne Sud.

T. Weise, S. Bouaziz, H. Li, and M. Pauly. 2011. Realtime performance-based facial animation. *ACM transactions on graphics (TOG)*, 30(4):1–10.