An Annotation Formalism for a French–LSF Bilingual Corpus Supporting Sign Language Generation

Sylvie Gibet, Clément Reverdy, Pierre-François Marteau
IRISA, Université Bretagne Sud
Vannes, France
{sylvie.gibet,clement.reverdy,pierre-francois.marteau}@irisa.fr

Abstract

This paper introduces an annotation formalism for bilingual corpora of written French and French Sign Language (LSF), based on a manually-produced, expert transcription of LSF video data. The formalism captures the grammatical specificities of LSF, including spatial and iconic mechanisms, while explicitly encoding features that support motor programs for animated signing avatars. We propose a parameterized gloss-based approach, called *PGloss-LSF*, which integrates syntactic and semantic structures alongside motion features critical for accurate sign synthesis. We illustrate the framework with examples drawn from our bilingual corpus. The annotation process is incremental, ensuring internal consistency and computational tractability through a two-step evaluation: a qualitative assessment aligning generated signs with the annotation language, and a quantitative evaluation via automatic translation using large language models. By bridging the linguistic specificities of sign language with the computational requirements of sign synthesis, this work advances the integration of sign language corpora into multilingual resources and contributes to the standardization of sign language technologies.

Keywords: sign language, annotation, formalism, bilingual corpora, generation

1. Introduction

Sign languages are fully-fledged visual-gestural languages with their own grammar and complex linguistic organization. Unlike spoken languages, they rely on spatial structuring, iconic mechanisms, and the simultaneity of manual and non-manual articulators. Each country-specific sign language constitutes an autonomous linguistic system.

In the case of French Sign Language (LSF), most educational, cultural, and administrative information remains primarily available in written or spoken French, limiting accessibility for many deaf signers. Developing computational tools for translation between written French and LSF therefore represents both a societal need and a scientific challenge. A central difficulty lies in representing LSF utterances in a way that preserves their grammatical structure while remaining operational for annotation and computational processing.

This paper focuses on the translation of written French into LSF and introduces a parameterized gloss-based annotation language, hereafter referred to as *PGloss-LSF*, designed as an intermediate representation between French sentences and LSF utterances. Standard gloss sequences are insufficient to capture key properties of LSF, such as simultaneity, spatial anchoring, and the interaction between manual and non-manual components. We therefore extend gloss notation with explicit parameters encoding spatial organization, co-occurrence of both hands, and asynchronous non-manual features (e.g., facial expressions and torso movements). The proposed formalism is grounded in linguistic descriptions of LSF grammar (Cuxac, 2000; Millet, 2019) and is designed to support annotation practices similar to those used in large sign language corpora. Furthermore, while several non-linear or highly expressive formalisms have been proposed, they often remain difficult to apply efficiently in large-scale manual annotation.

Our objective is to propose a structured intermediate representation that facilitates efficient corpus annotation suitable for LLM-based processing, while providing a sufficiently direct and controllable interface for motor-level specification, as required for avatar animation. By balancing linguistic precision and computational tractability, the proposed specification aims to bridge the gap between linear gloss representations and highly detailed scripting systems oriented toward motion generation. The remainder of this paper is organized as follows. We first review related work, then present the annotation formalism illustrated with examples drawn from our corpus, and finally report our initial evaluation results.

2. Related Work

The representation of sign languages has been extensively studied from a linguistic perspective. Stokoe's foundational work (Stokoe et al., 1976) introduced a phonological decomposition of signs into minimal parameters, later extended by systems such as *HamNoSys* (Prillwitz, 1989) and SignWriting. Multichannel models, including the

Proceedings of the 12th Workshop on the Representation and Processing of Sign Languages, pages 184–192

PDTS framework (Johnson and Liddell, 2011) and tier-based annotation schemes (Johnston and de Beuzeville, 2009), emphasized the coordination of manual and non-manual articulators. In corpus practice, ELAN has become the standard environment for multi-tier annotation, enabling the alignment of glosses, facial expressions, gaze, and body movements. However, these approaches primarily provide descriptive transcription frameworks rather than structured intermediary representations designed for computational reuse.

From a computational standpoint, scripting and animation languages such as *QualGest* (Lebourque et al., 2001), Partition/Constitution (P/C) (Huenerfauth, 2006), *SiGML* (Elliott et al., 2008), and *AZee* (Filhol et al., 2017) provide highly expressive multichannel specifications of sign articulation. While these systems enable precise modeling of signs and utterances, they require detailed manual encoding and are not well suited to large-scale corpus annotation.

Major corpus efforts, particularly for German Sign Language (DGS), have introduced more efficient annotation schemes (Hanke et al., 2020, 2023). The DGS Corpus annotation framework is primarily designed to support the development of a corpus-based lexical database. It focuses on lexical signs and integrates segmentation cues, lemmatisation rules, and specific annotation features such as double glossing and double-token tags. It also accounts for mouthing, including its coarticulation with spoken-language elements, and captures multi-channel constructions enabling the representation of context-sensitive meanings. Recent work such as Easier Notation (Hanke et al., 2023) further extends gloss-based representations within a linear, human-readable paradigm and provides an editing interface.

Other semi-automatic annotation approaches combine computer vision with linguistic constraints (Momeni et al., 2022; Neidle et al., 2018; Metaxas et al., 2018). More recently, neural end-to-end approaches to sign language translation and production (Kahlon and Singh, 2021; Farooq et al., 2021; Núñez Marcos et al., 2023) have enabled automatic gloss prediction and avatar pose generation (Shalev-Arkushin et al., 2023; Yu et al., 2024; Baltatzis et al., 2024). However, these approaches typically rely on linear gloss sequences or holistic pose representations, which under-specify the structured simultaneity and grammatical parameterization central to sign languages.

In this context, our work proposes a parameterized gloss-based annotation language designed as an operational intermediary layer: it preserves linguistic structure, supports annotation practices, and facilitates both corpus enrichment and downstream generation tasks. While our approach shares similar goals with DGS-based notation approaches, the two can be seen as complementary: the latter focus on scripting and editing based on lexical data, whereas our approach emphasizes syntactic and semantic modeling, as well as scalability for LLM-driven text-to-sign translation. In addition, utterances glossed with *PGloss-LSF* are systematically aligned with produced sign language (video or avatar), and the parameterization explicitly incorporates gestural targets directly usable for synthesis.

Our long-term objective is to develop an end-to-end text-to-sign translation system that maps written input to motor-level representations for sign language generation, while ensuring intelligibility through consistent and linguistically grounded annotation.

3. Annotation Formalism

Motivation and Key Features

The main objectives of our annotation formalism are: (i) to enable in-depth analysis of linguistic processes in French and LSF; (ii) to provide an intermediate language oriented toward motor program specification, thereby facilitating the control of signing avatars; (iii) to enable the text-to-sign translation using large language models by following the approach described in (Reverdy et al., 2025). Our formalism serves as an interface between written French and LSF through a parameterized gloss-based annotation language, called *PGloss-LSF*, which aligns French sentences with actual LSF utterances (video or 3D avatar). It preserves key LSF-specific mechanisms, in particular spatialization and iconicity, by supporting the co-occurrence of manual and non-manual components.

In contrast to highly expressive non-linear representation systems, our approach prioritizes scalability and usability for corpus annotation, while preserving key linguistic structures through explicit parameterization. This makes it possible to construct large annotated datasets while maintaining a representation that remains directly exploitable for motor-level specification and avatar control.

Our annotation approach follows an incremental and controlled design. We begin with simple glosses without arguments and progressively introduce additional parameters (arguments with associated values) to capture specific grammatical targets and increasing sentence complexity, while ensuring that all information required for synthesis is specified. In an initial phase, we annotated utterances containing signs in their citation form, followed by structures for which a syntactic formalization could be explicitly defined, such as directional verbs. Since then, we have extended the annotation framework to encompass more complex syntactic and semantic phenomena, including pointing gestures, different forms of classifiers (called *proforms*), co-occurring signs enabling the simultaneous use of the left and right hands, non-manual features (e.g., facial expressions), and role shift expressed through torso movements in dialogic contexts.

At each stage of corpus construction and annotation, a two-step validation protocol has been systematically implemented. First, annotations are reviewed by Deaf linguists or experts with advanced proficiency in LSF to ensure linguistic adequacy and theoretical consistency. Second, a computational validation phase is conducted using LLM-based corpus enrichment methods (Reverdy et al., 2025), enabling cross-verification, detection of inconsistencies, and iterative refinement of the formal representation.

French Word / LSF Gloss Dictionary

The dictionary lists all lexical items in the corpus in the format `<French-word, PGloss-LSF>`. Each gloss is defined by its name and a set of argument–value pairs, represented as:

$$\texttt{[GLOSS-NAME\_arg}_1\texttt{:VAL}_1\texttt{:.\_arg}_n\texttt{:VAL}_n\texttt{]}$$

Sentence Structure

A sentence is represented as a sequence of glosses, each accompanied by arguments specifying a lexical, syntactic or semantic function: for example the source, destination, and object for an indicating or trajectory sign, the spatial location, the used hand, the facial expression, the torso orientation for role shifts, the size and shape specifier, the classifiers, etc. Each argument is encoded by an identifiers such as `_src`, `_tgt`, `_mpr`, `_nsn`, `_id`, `_loc`, `_hd`, `_trs`, `_pnt`, `_sass`, `_exp`. These (argument, value) pairs enable discourse to be articulated around spatial references (*loci*), geometric processes (trajectory verbs, specific handshapes), and processes that enable simultaneity between several signs (using two hands, facial expression) or narrative processes (torso orientation).

The canonical syntactic structure in LSF typically follows a default-order sequence, encompassing time, location, object, agent, and action. Locative information is usually introduced with a general spatial description and gradually narrowed to a specific locus in signing space.

> **French:** *Hier, il neigeait dans la forêt.*
> **English:** *Yesterday, it was snowing in the forest.*
> **LSF:** `[HIER] [FORÊT] [NEIGER]`
> **SL:** `[YESTERDAY] [FOREST] [SNOW]`

In the previous example, the sequence consists of signs in their citation form, each associated with elementary motor programs that are sequentially activated. In the following, examples are presented in English, with corresponding LSF glosses translated into English for clarity.

A key feature of LSF syntax is the semantic distinction between animate (people, animals) and inanimate (objects, places) entities. Gloss ordering reflects this distinction and determines the verb's position relative to its arguments: for inanimate objects, `[OBJECT] [AGENT] [VERB]`; for animate beneficiaries or patients, `[AGENT] [VERB] [BENEFICIARY/PATIENT]`.

> *The woman brings food for the picnic.*
> `[PICNIC] [FOOD] [WOMAN] [BRING]`

> *The teacher asks the student.*
> `[TEACHER] [ASK] [STUDENT]`

Negation

For certain verbs, negation is integrated directly into the sign. Examples of these verbs along with their corresponding negative forms are presented below: `[NEED]` / `[NOT-NEED]`, `[WANT]` / `[NOT-WANT]`, `[CAN]` / `[CANNOT]`, `[BELIEVE]` / `[NOT-BELIEVE]`, `[KNOW]` / `[NOT-KNOW]`, `[HERE]` / `[NOT-HERE]`, `[LOVE]` / `[NOT-LOVE]`

> *You do not like strawberries*
> `[STRAWBERRY] [PRO2] [NOT-LOVE]`

> *He does not want to go to school by bike*
> `[SCHOOL] [BIKE] [GO] [PRO3] [NOT-WANT]`

For other verbs, negation is expressed by an additional sign, such as `[NOT]`, `[NEVER]`, `[NOTHING]`, `[NO-MORE]`, or `[THERE-IS-NOT]`, placed after the relevant group of glosses. Note that these signs are typically accompanied by head movements or negative facial expressions.

> *I do not understand the lesson*
> `[LESSON] [PRO1] [UNDERSTAND] [NOT]`

> *They never eat meat*
> `[MEAT] [PRO6] [EAT] [NEVER]`

Punctuation

Punctuation, annotated as `[/]`, is incorporated in *PGloss-LSF* sequence to segment portions of an utterance. These markers correspond to pauses in the flow of signs, which are necessary to preserve the sentence's semantic structure.

Adjectives

Adjectives follow the words to which they refer.

> *This morning, he admires the tired rose. She prepares to open her delicate petals.*
> `[MORNING] [ROSE] [TIRED] [PRO3] [ADMIRE] [/] [PETALS] [DELICATE] [PRO3] [OPEN] [PREPARE]`

Discretization of the signing space: _loc

The signing space is divided into discrete zones, following the principles defined in (Lebourque et al., 2001) for motor program specification and in (Millet, 2019) for the pre-semantic space. We discretize this space relative to the signer's body using three biomechanical planes –sagittal, frontal, and transverse, corresponding to forward/backward, left/right, and up/down dimensions. Two additional intermediate planes, oriented at 45° to the right and left, further subdivide the space. Within each plane, directions are coded as follows: planes orthogonal to the transverse plane use left L, neutral N (default), and R; depth (amplitude) is coded as on the body B, medium M (default), or extended E; vertical directions are coded as 0 (horizontal), 1 ($\approx 30°$ from transverse, default), and 2 ($\approx 60°$ upward). These three dimensions uniquely define a location in the signing space, with default values used when information is missing.

Locations are defined relative to an absolute reference linked to the signer's center of gravity. If the torso rotates from the frontal position (e.g., during role shift), the signing space rotates accordingly, and previously defined locations retain their coding. The argument _loc specifies the position of an entity within these predefined zones.

> [HOUSE_loc:R1M]
> means the sign HOUSE at the location right, medium height, medium amplitude.

Identity: _id, _nsn

Individuals are generally identified by their French name, which can be spelled out, and by their signed name _nsn. Because the signed name may vary within a narrative while still referring to the same person (e.g., [CURLED] or [LONG] [CURLED]), an _id argument can be used to maintain consistent reference.

> John, signed name CURLED, is represented as [nsn:CURLED_id:JOHN]

Pre-semantic space, pronouns

We define a pre-semantic space, inspired from (Millet, 2019), as illustrated in Figure 1. This space provides a structured set of predetermined areas that enable shared and consistent referencing within a narrative. These areas are represented by conventionalized spatial loci (rather than discourse-established referential loci), and used as pronominal referents or as starting points or end points of executed trajectories, particularly for directional or trajectory verbs.

In LSF, pronouns exploit the pre-semantic space as deixis elements, whose meaning depends on the discourse context. They are realized through pointing or designation gestures toward specific loci referenced in the discourse. Singular pronouns use a pointing gesture with the index finger extended to a pre-semantic locus. In our annotation, these are represented as [PROx], with x = 1, 2, 3.

Specifically, singular pronouns [PRO1] (I/Me), [PRO2] (You), and [PRO3] (he/she/him/her) are realized by pointing to their reserved loci, while plural pronouns use small gestures within the pre-semantic space: a circular gesture for [PRO4] (we) and PRO6 (they/them), or a sweeping gesture for [PRO5] (you).

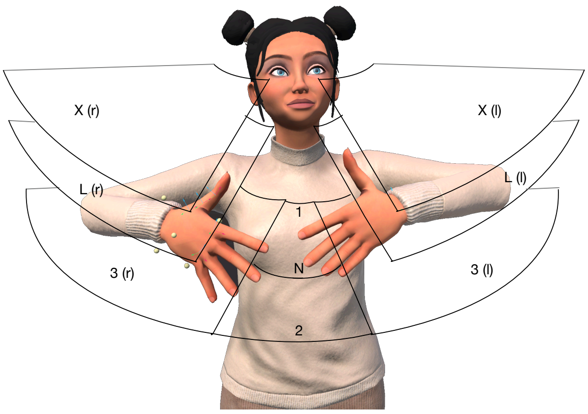

Figure 1: The pre-semantic space, inspired from (Millet, 2019).

Other third-person pronouns vary with animacy and are defined in the pre-semantic space with the prefix PSS:

PSS3: animate (he/she), in space 3, to the left or right of the signer, in the horizontal plane;

PSSX: indeterminate (*someone*), to the left or right of the signer, in the horizontal plane at temple level;

PSSL: inanimate (*it*), to the left or right of the signer, referring to a location, between the two previous planes.

Similarly, additional pre-semantic zones are defined as:

PSS1: near the signer;

PSSN: neutral zone, in front of the signer;

PSSO: distant zone, corresponding to the objective.

> *Information is sent to me.*
> [INFORMATION] [SEND_src:PSSX_tgt:PSS1]

> *I send information to someone.*
> [INFORMATION] [SEND_src:PSS1_tgt:PSSX]

> *He is going to visit Paris.*
> [PARIS] [PRO3] [VISIT] [GO_tgt:PSSL]

Possessive pronouns can be represented as:

TO-PRO1 : *mine*

TO-PRO2 : *yours*

TO-PRO3 : *his/her*
TO-PRO4 : *ours*
TO-PRO5 : *yours (plural/formal)*
TO-PRO6 : *theirs*

Pointing: _pnt

Pointing gestures are a core mechanism for spatial reference in LSF. We focus on simple deictic forms with the index finger or hand, leaving other strategies (e.g., gaze or body orientation) aside. The referential target, not the gesture itself, is the key parameter.

Spatial reference is realized via loci in the discretized signing space, corresponding either to entities previously introduced and anchored in the discourse (person, object, place, etc.) or to a predefined zone of the pre-semantic space. In our annotation language, each locus is encoded as a pointing motor program, computed via an inverse kinematics process, with its spatial parameter given by _pnt followed by a value POINTING-TARGET defined in the spatial coding system.

The house that is there
```
[HOUSE] [THERE_pnt:R1E]
```

The neutral demonstrative pronoun may also be realized by a pointing gesture, with a palm-up handshape, referring to a previously introduced conceptual entity, as illustrated below.

It is something too often forgotten (Le Petit Prince, Saint Exupéry)
```
[THAT_pnt:THING-POINTED] [THING] [FORGET] [OFTEN]
```

Other deictic forms associated with demonstrative pronouns (e.g., THIS-ONE, THAT-ONE) and demonstrative determiners (e.g., THIS, THAT, THESE) are not addressed here.

Directional or trajectory verbs in LSF: _src, _tgt, _mpr

In LSF, there exists a set of verbs, called *directional (or indicating) verbs*, which are syntactically structured as follows:

```
[VERB_src:SOURCE_tgt:TARGET_mpr:OBJECT]
```

with

- VERB : the action performed
- SOURCE : the agent of the action
- TARGET : the beneficiary of the action
- OBJECT : the entity on which the action is performed (optional)

The arguments _src and _tgt can take predefined location in the pre-semantic space, with values such as _src:PSS1, _tgt:PSS3, or the identifier of an already referenced object, e.g., _id:APPLE1. The third argument is addressed in the paragraph on proforms.

I tell you a story.
```
[STORY] [TELL_src:PSS1_tgt:PSS2]
```

He lends me his computer.
```
[COMPUTER] [TO-PRO3] [LEND_src:PSS3_tgt:PSS1]
```

The arguments _src and _tgt can also be used for verbs with trajectories.

Paris, many tourists go there.
```
[PARIS] [THERE_pnt:PSSL] [TOURIST] [MANY]
[GO_src:N_tgt:PSSL]
```

Note that verbs other than directional or trajectory verbs do not incorporate the agents and beneficiaries of the action in their realization.

In Brussels, we eat fries.
```
[BRUXEL1S] [FRIES] [PRO4] [EAT]
```

Manual Proforms _mpr

PProforms are linguistic mechanisms grounded in the iconic dynamics of sign languages. Manual proforms act as pronouns, maintaining lexical and referential continuity and supporting syntactic structure. Bodily proforms involve the signer's embodied engagement, whereby the signer adopts the role of a character in the narrative.

From a syntactic perspective, specific hand configurations reproduce or maintain lexical properties within a sign or across discourse. These forms correspond to manual proforms (also called classifiers [1] in English). In our annotation language, manual proforms are specified using the argument _mpr, followed by the name of the proform prNAME (distinct from the citation form of the sign), and optionally by arguments specifying its location and identifier.

We distinguish several types of manual proforms:

1. **Verb-incorporated proforms**, particularly in directional or trajectory verbs involving an object argument; they may also occur in dynamic pointing constructions.

2. **Proforms representing an entity**, annotated as independent glosses. These proformed signs are primarily used to position entities relative to one another and may refer to either animate or inanimate referents.

3. **Quasi-lexicalized dynamic proforms**, including person proforms (with variants such as standing, seated, lying, etc.), and vehicle proforms.

[1] Classifiers include both proforms and size-and-shape specifiers.

Examples of the first category are shown below:

```
I give him an apple.
[APPLE] [GIVE_src:PSS1_tgt:PSS3_mpr:prAPPLE]
```

```
He gives you the book (Renard's book).
[BOOK]                  [mpr:prBOOK_id:RENART-BOOK]
[GIVE_src:PSS3_tgt:PSS2_mpr:RENART-BOOK]
```

In the first example, the proform prAPPLE is integrated into the verb [GIVE], producing a handshape that encodes the object along the verb's trajectory. In the second example, the handshape represents the book identified by RENART-BOOK. The proform may optionally precede the verb to emphasize the object. It should be noted that a list of pre-defined proforms is used in our corpus, each associated with specific hand configurations.

Proforms of the second category frequently involve co-occurring signs, i.e., signs performed simultaneously with both hands. Examples are presented in the following section, which introduces the concept of manual simultaneity. Proforms of the third category are not addressed here due to their high variability and the need for separate analysis.

Simultaneity of left/right hand: _hd

This mechanism involves producing a sign with one hand while the other maintains an utterance-level element, typically via a proform –for example, executing a sign with the right hand while the left hand holds a proformed sign. The argument _hd specifies the hand (L or R). Once assigned, the hand retains its properties (location, handshape, orientation, etc.) until released or reassigned by a subsequent gloss.

A hand can be released in two ways: (i) implicitly, when a subsequent gloss is produced without the _hd parameter; (ii) explicitly, using the tag [/hd:L] or [/hd:R].

```
I am going to the house.
[HOUSE_loc:L1] [mpr:prHOUSE_loc:L1_id:HOUSE1_hd:L]
[GO_src:PSS1_tgt:HOUSE1_hd:R]
```

Through spatial anchoring, the sign [HOUSE] is first produced at location L1. The proform prHOUSE is then maintained by the left hand at the same location, while the verb GO is executed by the right hand along a trajectory from PSS1 toward the spatial locus associated with HOUSE1.

```
I forgot the keys at the house.
[HOUSE]    [FORGET]    [mpr:prHOUSE_id:HOUSE1_hd:L]
[KEY_hd:R] [OPEN-KEY_loc:HOUSE1_hd:R]
```

In this last example, the left hand maintains the proform prHOUSE (identified as HOUSE1), while the right hand sequentially produces the signs [KEY] and [OPEN-KEY]. The opening action is directed toward the spatial locus associated with HOUSE1, thereby encoding the locative relation through simultaneity.

The spatial relationship of one object relative to another can be encoded using two proforms, each performed by one hand. A non-exhaustive set of relative position vocabulary includes: ABOVE, BELOW, OVER, UNDER, LEFT, RIGHT, NEXT-TO, IN-FRONT-OF, BEHIND, TOP, ON, INSIDE, etc.

```
The bowl (proform) is on the table (proform)
[BOWL]     [TABLE]     [prm:prTABLE_id:TAB1_hd:L]
[prm:prBOWL_id:BOL1_hd:R_loc:TAB1*ABOVE]
```

In this example, we first produce the sign [BOWL], followed by [TABLE], and then position the proform prTABLE slightly before prBOWL, applying the modifier ABOVE to indicate the spatial relation.

```
There is a tree; at the top of the tree, there is a nest in which there is a bird.
[TREE_loc:L]       [mpr:prTREE_id:ARB1_loc:L_hd:L]
[THERE_pnt:ARB1*TOP_hd:R_id:HA]      [NEST_loc:HA]
[mpr:prNEST_id:NEST1_hd:L_loc:HA]
[BIRD_loc:HA*INSIDE_hd:R]
```

In this example, the sign [TREE] is first produced by the two hands at a left locus (spatialization) and then maintained as a proform with the left hand. The right hand points to the top of the tree, establishing the location labeled HA. Next, the sign [NEST] is produced by the two hands at the left locus identified by HA and maintained as a proform with the left hand. Finally, the sign [BIRD] is executed by the right hand inside the nest.

This utterance can alternatively be produced without pointing, by anchoring the relevant elements with the hands in the signing space:

```
[TREE]                    [mpr:prTREE_id:ARB1_hd:R]
[mpr:prNEST_id:NEST1_loc:ARB1*TOP_hd:L]
[BIRD_hd:R_loc:NEST1*INSIDE]
```

Role shift (torso orientation): _trs

When a narrative involves multiple characters, the narrator may alternately assume the role of each character, especially in dialogue situations. In such cases, torso rotation from left to right signals the change of character.

Each role shift is annotated with an opening tag indicating torso orientation [trs:VALUE], with VALUE = L, R, or N, and a corresponding closing tag [/trs]. The opening tag can optionally include an _id argument to identify the signing interlocutor.

Because this argument typically applies to a sequence of glosses, [/trs] marks the end of the sequence. An optional _id parameter may also be included to disambiguate the interlocutor when identities change within the narrative. Examples of dialogue with role shifts (from the Tales of the Fox) are provided below. Examples:

```
Merchant 1: Look! A fox! I think it is dead!
[trs:R] [LOOK] [FOX] [DEAD] [PRO1] [THINK] [/trs]
```

– Let's stop and take it!
`[trs:L] [PRO4] [STOP] [TAKE_src:R_tgt:L_mpr:prTHING] [/trs]`

– Look at the fur! It is beautiful, red with a white neck!
`[trs:R] [FUR] [LOOK] [BEAUTIFUL] [RED] [WITH] [NECK] [WHITE] [/trs]`

– Merchant 2, angry: Damned fox! We hope the eels will suffocate you!
`[trs:L] [DAMN_exp:MEAN] [FOX] [/] [EELS] [PRO2] [SUFFOCATE] [PRO4] [HOPE] [/trs]`

Facial expression: `_exp`

In LSF, facial expressions are syntactically relevant and fulfill three main functions: (i) marking sentence modality (e.g., interrogative, exclamative, imperative); (ii) expressing adverbial or adjectival meaning; and (iii) conveying the signer's attitude toward the utterance content (e.g., emotion, boredom, concern, contempt).

Facial expressions may co-occur with a single gloss or extend over multiple glosses simultaneously with manual signs. We therefore annotate them using the argument `_exp` in two ways: When the facial expression applies to a single gloss, it is included directly within that gloss, e.g., `[GLOSS_exp:INTENSE]`; When it spans multiple glosses or an entire clause, we use opening and closing tags.

The examples below illustrate these three situation:

`[exp:QUESTION] [GLOS1] [GLOS2] [GLOS3] [/exp:QUESTION]`

This cat is very cute.
`[CAT_id:CAT1] [THIS_pnt:CAT1] [CUTE_exp:VERY]`

He eats messily.
`[PRO3] [EAT_exp:DISGUST]`

You don't eat soup?
`[exp:DOUBT] [SOUP] [PRO2] [EAT] [NOT] [/exp:DOUBT]`

4. Resultats and Evaluation

4.1. Bilingual Corpus

We rely on a bilingual corpus (French sentences paired with LSF parameterized glosses) which has been created specifically for our *SignToKids* project. It comprises materials produced by Deaf educators from French institutions[2], as well as transcribed children's tales and literary works.

This corpus include several sub-corpora:

- Sentences illustrating grammatical structures such as negation, interrogation, and indicating verbs.
- LSF narratives of the Tales of Renart.
- Short LSF stories demonstrating spatialization mechanisms and proforms (e.g., persons, vehicles).
- Several excerpts from *Le Petit Prince* by Antoine de Saint-Exupéry.

All corpora were annotated and signed by native Deaf signers and recorded using motion capture or video. In total, the dataset comprises approximately two hours of signed LSF utterances.

4.2. Evaluation

The evaluation of the proposed formalism follows a twofold approach. First, a qualitative linguistic assessment is conducted in which a Deaf LSF teacher generates signed utterances based on the written transcription. The alignment between these generated productions and the corresponding LSF video recordings is subsequently verified manually, with corrections implemented whenever discrepancies are identified.

Second, we conducted a quantitative evaluation using an automatic text-to-sign translation system. Unseen French sentences were translated into LSF glosses and analyzed for structural consistency with the specification.

The initial evaluation used 165 sentence pairs and was later expanded to 1,650 pairs (Reverdy et al., 2025), covering three-argument glosses with varied spatial loci in the pre-semantic space. Gloss prediction (the semantic label) reached an 18% error rate (Levenshtein distance), while parameters (argument and value) achieved F1-scores of 97% and 98%, respectively.

The annotation scheme was then extended to include additional grammatical mechanisms (pointing, manual proforms for inanimate objects, left/right hand co-occurrence, facial expressions, and role shift), resulting in a dataset of approximately 3,000 sentence pairs. Evaluation of this expanded set is ongoing.

5. Conclusion and perspectives

This work has led to the development of a formalism for the annotation of LSF, called *PGloss-LSF*, systematically aligned with LSF video data. The formalism is based on glosses enriched with explicit linguistic and articulatory parameters. Unlike linear gloss sequences, which fail to capture the structural specificities of sign languages, it integrates syntactic and semantic grammatical descriptions

[2]Institut National des Jeunes Sourds and Centre Gabriel Deshayes

that can be dynamically adapted to the discourse context. Its key contribution lies in its ability to explicitly model motor control parameters, thereby bridging linguistic representation and animated sign synthesis.

At this stage, the majority of the formalized grammatical rules of LSF are incorporated into the framework. Future work will extend the annotation scheme to cope with additional sign language mechanisms that exhibit greater variability, such as dynamic proforms and highly iconic constructions. Ultimately, the proposed formalism is designed to be adaptable to other sign languages.

Beyond linguistic documentation, the framework supports computational applications including automatic sign language translation, avatar-based sign generation, and assistive communication technologies. By addressing the challenges of bilingual alignment and semantic interoperability, our *PGloss-LSF* formalism provides a principled and extensible foundation for annotating, integrating, and sharing data across heterogeneous linguistic resources.

6. Acknowledgements

The *SignToKids* project is funded by the French National Research Agency (ANR). Several research teams are involved in this project: the IRISA–CNRS laboratory, the MoDyCo–CNRS laboratory, and the National Institute for Deaf Youth (INJS). We would also like to thank the Deaf and signing participants who contributed to the project: Nathalie Irdel, from the Gabriel Deshayes Center, Kristof Gendreau Touchais from the University of Poitiers, and the Deaf/French teaching staff at INJS. They all contributed to the design, recording, and annotation of the corpus.

7. Bibliographical References

Vasileios Baltatzis, Rolandos Alexandros Potamias, Evangelos Ververas, Guanxiong Sun, Jiankang Deng, and Stefanos Zafeiriou. 2024. Neural sign actors: A diffusion model for 3d sign language production from text. *arXiv 2312.02702*.

Christian Cuxac. 2000. *La langue des signes française (LSF) : les voies de l'iconocité*. Faits de langues. Ophrys.

Ralph Elliott, John RW Glauert, JR Kennaway, Ian Marshall, and Eva Safar. 2008. Linguistic modelling and language-processing technologies for avatar-based sign language presentation. *Universal Access in the Information Society*, 6(4):375–391.

U. Farooq, M. Rahim, N. Sabir, A. Hussain, and A. Abid. 2021. Advances in machine translation for sign language: approaches, limitations, and challenges. *Neural Computing and Applications*, 33:14357–14399.

Michael Filhol, John McDonald, and Rosalee Wolfe. 2017. Synthesizing sign language by connecting linguistically structured descriptions to a multi-track animation system. In *International Conference on Universal Access in Human-Computer Interaction*, pages 27–40. Springer.

Thomas Hanke, Lutz König, Reiner Konrad, Maria Kopf, Marc Schulder, and Rosalee Wolfe. 2023. EASIER notation – a proposal for a gloss-based scripting language for sign language generation based on lexical data. In *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, ICASSP 2023 - Workshops, Rhodes Island, Greece, June 4-10, 2023*, pages 1–5, Rhodes, Greece.

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

Matt Huenerfauth. 2006. *Generating American Sign Language classifier predicates for English-to-ASL machine translation*. Ph.D. thesis, University of Pennsylvania.

Robert E Johnson and Scott K Liddell. 2011. A segmental framework for representing signs phonetically. *Sign Language Studies*, 11(3):408–463.

Trevor Johnston and Louise de Beuzeville. 2009. Researching the linguistic use of space in Auslan: Guidelines for annotators using the Auslan corpus. Technical report, Department of Linguistics, Macquarie University, Sydney.

Navroz Kahlon and Williamjeet Singh. 2021. Machine translation from text to sign language: a systematic review. *Universal Access in the Information Society*.

Thierry Lebourque, Sylvie Gibet, and Pierre-François Marteau. 2001. High level specification and animation of communicative gestures. *Journal of Visual Languages and Computing*, 12:657–687.

Dimitris Metaxas, Mark Dilsizian, and Carol Neidle. 2018. Scalable ASL sign recognition using model-based machine learning and linguistically annotated corpora. In *Workshop on the Representation and Processing of Sign Languages (LREC 2018)*, pages 127–132, Miyazaki, Japan.

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

Liliane Momeni, Hannah Bull, K. R. Prajwal, Samuel Albanie, Gül Varol, and Andrew Zisserman. 2022. Automatic dense annotation of large-vocabulary sign language videos. In *European Conference on Computer Vision – ECCV 2022*, pages 671–690. Springer.

Carol Neidle, Augustine Opoku, Gregory Dimitriadis, and Dimitris Metaxas. 2018. New shared & interconnected asl resources: Signstream® 3 software; dai 2 for web access to linguistically annotated video corpora; and a sign bank. In *Workshop on the Representation and Processing of Sign Languages (LREC 2018)*, Miyazaki, Japan.

Adrián Núñez Marcos, Olatz Perez-de Viñaspre, and Gorka Labaka. 2023. A survey on sign language machine translation. *Expert Syst. Appl.*, 213.

Siegmund Prillwitz. 1989. *HamNoSys: Version 2.0; Hamburg Notation System for Sign Languages; An Introductory Guide*. Signum-Verlag.

Clément Reverdy, Sylvie Gibet, and Pierre-François Marteau. 2025. Translating sentences from written french into LSF annotation. In *Adjunct Proceedings of the 25th ACM International Conference on Intelligent Virtual Agents, IVA Adjunct 2025, Berlin, Germany, September 16-19, 2025*, pages 13:1–13:6.

Rotem Shalev-Arkushin, Amit Moryossef, and Ohad Fried. 2023. Ham2pose: Animating sign language notation into pose sequences. In *IEEE/CVF Conf. on Computer Vision and Pattern Recognition, CVPR 2023, Vancouver, BC, Canada, June 17-24, 2023*, pages 21046–21056.

William C Stokoe, Dorothy C Casterline, and Carl G Croneberg. 1976. *A dictionary of American Sign Language on linguistic principles*. Linstok Press.

Zhengdi Yu, Shaoli Huang, Yongkang Cheng, and Tolga Birdal. 2024. SignAvatars: A Large-scale 3D Sign Language Holistic Motion Dataset and Benchmark. *arXiv 2310.20436*.