Formalising Sign Language Depiction, Characterising Categories and Measuring Iconicity with AZee

Michael Filhol, Emmanuella Martinod
Université Paris-Saclay, CNRS, Laboratoire Interdisciplinaire des Sciences du Numérique (LISN)
507, rue du Belvédère, 91400, Orsay, France
michael.filhol@cnrs.fr, emmanuella.martinod@gmail.com

Abstract

This paper deals with depiction in (French) Sign Language, the formal account AZee can provide, and how it compares, validates or simplifies the linguistic notions of classifiers and iconic structures. It reports on a partial encoding work on *Mocap1*, a corpus with a high density of depicting structures, following the same method that led to the first AZee reference corpus *40 brèves*. The approach does not postulate classifiers or iconic structures as entities separate from lexical signs, and nonetheless manages to model the corpus data. We discuss the entailed possibility to rediscover some of the useful categories, and if so define them from AZee's premises. We also specify how a formal metric can be specified to measure iconicity in signed data. While this paper is of linguistic interest as it compares to existing theories, it also provides a concrete step to covering depicting discourse with AZee, therefore enable automatic SL animation of depiction.

Keywords: French Sign Language, Formal Representation, SL Synthesis, AZee, Iconicity, Depicting Signs

1. Introduction

Sign Languages (SLs) depend on visual-gestural modality. This modality allows iconicity in the language, where elements resemble what they refer to, even outside of conventionalised signs. In the history of SL study, this non-lexical iconicity has often been overlooked, sometimes even rejected or considered anecdotal (Frishberg, 1975). Although with varying degrees of interest, it is clearly more considered today. This led to the development of a large range of concepts to describe it.

These descriptions are however not formal to the extent that they could be used for automatic generation with avatars for example. In the domain of SL generation, iconicity is indeed still largely set aside because of its complexity. To this end, we have been developing the AZee framework, which provides a formal linguistic description, readable by a computer to fully specify avatar animation. This paper contributes to developing AZee further by confronting a corpus containing heavily iconic signing, a genre still only partially covered.

First, we present the main concepts used in linguistics to describe iconic structures in SLs. After introducing AZee as a theoretical framework and a language, we report on the corpus we encoded and how these structures were handled with AZee. We then propose a comparison to existing linguistic frameworks for their description, attempting to rediscover categories such as "classifiers" which have been essential in linguistic theories.

2. State of the Art

2.1. Related Work in Linguistics

Iconic signed parts have traditionally been analysed distinguishing low-level combining features, with special interest in handshapes, from the higher level structures in which these are used. We follow the same distinction in the two following sections. Then, we focus on the frequently mentioned links between iconic structures and the lexicon.

2.1.1. Iconic Handshapes

The most widely used concept is that of *classifier*. This terminology comes from an analogy with the description of some spoken languages like Navajo (Supalla, 1978). In those languages, *classifier morphemes* are used to refer to a class of entities. Applied to SL description, *classifier* stands for a meaningful handshape that expresses information about an entity: its size, shape, or location (Emmorey, 2003; Perniss et al., 2007).

Through time, many different types of classifiers have been defined. Today, a salient distinction seems to have emerged between *entity* (or *object) classifiers* and *handling classifiers*: the first concerns the representation of a shape, whereas the second concerns the representation of the handling of a shape. Several phonological differences have been observed between the two groups (Coppola and Brentari, 2014; Brentari and Goldin-Meadow, 2017).

Another term sometimes referring to such iconic representation of a shape in SL is *proform*. Several authors have used this term (Sutton-Spence

Proceedings of the 12th Workshop on the Representation and Processing of Sign Languages, pages 164–173
16 May 2026. ©ELRA Language Resources Association (ELRA), 2026

and Woll, 1999; Johnston and Schembri, 1999), but strongly promoting it is the "Semiological approach", which gives iconicity a crucial role in shaping SL discourse and grammar (Cuxac, 2000; Sallandre, 2014). It proposes a systematic role of eye gaze in the structures the proforms are used in, which we further present in the following section. In general though, we note little difference between the concepts of proform and classifier.

In any case, an iconic handshape is never produced alone: it is necessarily part of a structure that can be analysed as such (Valli and Lucas, 2000). Let us see how these structures are categorised in the literature.

2.1.2. Iconic Structures

From the 1990s, iconic structures have been described in SLs analysed separately, but a range of terminology has evolved to refer to those structures everywhere, with various definitions (Emmorey, 2003): classifier signs, classifier verbs, verbs of motion and location, classifier predicates, spatial-locative predicates, polymorphemic predicates, polysynthetic signs, productive signs, polymorphemic signs, etc.

Given the diversity of this terminology, a standardization was attempted by Johnston (2024), in part inspired by the work of Liddell (2003) in a cognitive-functionalist framework, i.e. emphasizing the *function* of linguistic structures. This document, last updated in 2024, serves as a reference for a large part of recent work in linguistics, especially when it comes to annotation of SL corpora. It distinguishes between depicting signs and non-conventional signs.

Depicting Signs

Alternatively named *classifier signs*, these structures include movement depictions ("DSM") and location depictions ("DSL"), in which the entity that moves or is located in the signing space is identified "from the context and/or by the handshape". They further include size and shape depictions ("DSS"), which "describe an object by showing a salient physical feature of it and often occur immediately before or after the object is named with a sign". A last possible type is ground depiction ("DSG"), representing the physical background, or a reference point in a situation.

Non-conventional Signs

These structures include "enactments", where the signer mimics an action using the whole upper body, without using lexical (i.e. conventional) signs.[1] Enactments can involve handling classifiers thus constitute "handling depiction". The other type is "gestures", which have a role of modulating and expressing the mood or attitude of the signer, or regulating the discourse and interaction.

[1]Enactment is also frequently called *Constructed Action* (CA) in the literature, see Cormier et al. (2015).

Johnston (2024)	Sem. approach
DSS	SST
Combinations of DSG, and/or DSM, and/or DSL	ST
Non-conventional Sign: Enactment and/or Gesture	PT + sub-types

Table 1: Descriptive concepts for iconic structures

In parallel, the semiological approach uses another terminology in which the criterion of non-conventionality of some structures is not at the core of the description. They are called *transfers* because of the cognitive process of transferring human experience into visual linguistic structures. The three main types of transfer are: size and shape transfer (SST), to show the size and/or shape of an entity, situational transfer (ST), which shows the movement of an actant in relation to a ground, and personal transfer (PT), where the signer embodies an individual or an entity. This last category has at least five sub-types which we do not develop here. Those three possible transfers require very precise use of manual and non-manual parametric components. In particular, a specific eye gaze pattern is observed: the signer's gaze seems indeed directed towards the hands *versus* towards the addressee for SST and ST, and the signer's gaze is the one of the embodied entity for PT.

We tried to establish correspondences of these main various descriptive concepts in table 1. Yet, the range of descriptive concepts often overlap without matching exactly, and frequently include combinations of sub-types, so the correspondence remains somewhat approximate. Despite standardization attempts, there is currently no official consensus to describe these elements which are nevertheless massively present in SLs.

2.1.3. Iconic vs. Lexical Continuum

The complexity of the linguistic literature on iconic handshapes and structures and their difficult categorisation also comes from the possibility to combine those iconic elements with conventional lexical signs. For example one can make use of a conventional sign while in the middle of an enactment. This would correspond to a non-conventional sign for Johnston (2024), and to a sub-type of PT ("semi-PT") in the Semiological approach (Sallandre, 2014). In the classification described by

Cormier et al. (2015), this case would correspond to a *Reduced CA*.

This kind of combination is part of SL. It is also linked with its diachronic evolution. Indeed, classifiers would be central in the emergence of lexical signs (Brennan, 1990; Fusellier-Souza, 2004; Johnston and Schembri, 1999; Schembri, 2003; Zwitserlood, 2012). This process sometimes makes it difficult to determine with certainty whether one is dealing with a lexicalized sign or not yet. Here too, categorisations diverge depending on the theoretical framework of authors and their respective focus, but they globally tend to overlap. The overall similarity is that even if researchers seem aware of the fact that the frontier between iconic and lexical is not sealed, their description overall include elements of two different natures. In any case, those categorisations are still missing the formal representation that is necessary for the purpose of avatar animation.

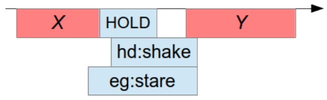
2.2. AZee Framework and Corpus Creation

AZee is a formal framework to represent Sign Language utterances and discourse, designed to make both following statements true (Filhol and Hadjadj, 2016):

- compositionality in meaning is accounted for, i.e. if an interpretable piece of signing S is built from a smaller interpretable piece with form s', then the formal representation of s' is visible in that of S;
- all body articulations are determined by their representation, in other words automatic avatar animation of any represented utterance is possible, which makes the approach testable.

This is done by encoding linguistic rules, each capturing an association between an identified meaning and form articulations in the target language. Such a rule is labelled according to the attached meaning, and the form specified in the native AZee language, which is able to express body posture constraints, geometric objects in the signing space, and synchronise actions on a timeline. The form–meaning association can be of any language level (including non-manual aspects), carry any semantic granularity, and produce a form constraining any body parts. It can also be parametrised, provided the form can be expressed as a function of the parameters in a stable way. For example:

- a rule “`red`” is specified for the trivial, invariable association between the meaning “red” and the conventional form one finds in a dictionary;



“hd”: head; “eg”: eye gaze

Figure 1: Form timeline for `not-but`(X, Y)

- the more complex, parametrised rule “`not-but`(X, Y)” associates the meaning “not X but Y” and the timeline shown in fig. 1 (arguments X and Y are open to nested signed content);
- or concerning the matter in this paper, rules making iconic use of space like `place-object`, which captures the common meaning factor between placing a plate on a table, a frame against a wall, etc. and the common form factor: a quick eye gaze to the object location loc, immediately followed by a straight movement of body parts jointly configured (denoting a class of object), settling at loc.

These rules then serve as the basis for “discourse expressions”, i.e. combinations representing signed utterances of arbitrary length. For example, rules `red`() and `with-surprise`(sig) can be combined to mean “really red?!”: `with-surprise`$(sig =$ `red`$())$. The examples above would also be useful to compose the meaning: “not [here / at point A], but surprisingly [there / at point B] instead”. Both arithmetically resolve to a timeline containing all of the forms contributed by the respective rules used in the expression.

The approach has been tested both theoretically, through linguistic coverage of the French Sign Language (LSF) videos of the *40 brèves* corpus (LIMSI and LISN, 2022), and technically, through actual avatar synthesis of AZee-encoded expressions (Filhol and McDonald, 2018). The corpus coverage test was published in (Challant and Filhol, 2024), and consists in the AZee encodings of 120 videos of about 30 s each, totalling 1 hour of LSF. It evaluated a coverage of 98%, through timing of the missing pieces, encoded as “ellipses”[2]. It is today a reference corpus for AZee.

2.3. What is Missing and New Goal

The genre of the *40 brèves* material has clear benefits to serve as a reference. It comprises

[2]An ellipsis is an expression made of a single dummy rule application, covering a non-encoded part of the signing and placed to keep the overall expression valid and compilable. Its label is not one known and taken from the production set, rather it is chosen ad-hoc to describe the meaning of the omitted part, and enclosed in brackets to make it recognisable. For example:
`:[cars everywhere stuck in traffic]`

Figure 2: Example of elicitation material used in the descriptive task of *Mocap1*

Figure 3: Example of `class-flat-round-large`

monologues facing the camera of error-free SL, not produced by simultaneous/on-the-fly interpretation, and provides a non-restricted semantic field. Besides, since spatial, iconic constructions are ubiquitous in SL, instances can be found in it too. But those are known to reach discourse proportions that are a lot higher in SL, especially when depicting scenes and relative positions of objects or places, which are not strongly exhibited by the *40 brèves*. A goal then was to extend the test above and apply it to a genre of signing containing a high density of these iconic depictions, eventually to evaluate the AZee capability on such data.

3. Experiment: *Mocap1* and First Pass

The corpus chosen to represent iconic depiction is *Mocap1*, created through controlled elicitation of image descriptions (LIMSI and CIAMS, 2020). It consists in 187 LSF videos, of 8 signers each freely describing 25 different images such as that in fig. 2, with an average duration of 25 s per video.

We applied the AZee methodology to write the discourse expressions for the contained videos, exactly as was done to reach the *40 brèves* milestone. That is, given the current state of the productions set, either:

- find the combination of rules both accounting for the meaning interpreted of the video and justifying the observable forms;
- or if this is impossible due to missing rules, mine for new rules in SL data to extend or refactor the set, following the published methodology (Hadjadj et al., 2018);
- lastly, if a discourse segment remains difficult to account for with a stable combination of known, new or refactored linguistic rules, capture the whole segment in an ellipsis.

At this point we have reached 60 encoded videos of the corpus, which corresponds to 30 minutes, and one third of the full *Mocap1* material. In total, 22,307 operations and values compose the 60 corresponding expressions, of which 5,471 are rules of the production set, and 35 are ellipses. The remaining 16,801 are native AZee operations (e.g. `translate`) and values (e.g. numericals). No funding is currently supporting further progress, but these numbers to us are enough to create interest in the data. We therefore chose to report on this first pass and look into the way iconic structures are handled.

4. Discussion and Comparison to Literature

Section 2.1 presented classifiers as manual forms with attached semantics (usually a category of object, according to its shape), recruitable in structures that fall out of the conventional lexicon, and yet are non-ambiguous between signers. This section studies how these forms and the structures they appear in are accounted for in the AZee material resulting from the encoding of Mocap1.

4.1. Iconic Handshapes

In AZee, such a reusable, stable link between a form and an attached semantics precisely makes the case for a linguistic rule in the production set. An example in LSF, and a widely shared one across SLs, is the form–meaning binding between the manual configuration of fig. 3 and a large flat round object.

The form, as any other postural configuration in AZee, is specified as a set of articulatory constraints applying simultaneously. In the case of fig. 3, the form:

- constrains the appropriate finger joints to make the thumb and index bend and the other fingers close;
- orients the index proximal joints along opposite normal vectors to make the hands enclose a disc, in a plane whose orientation might also be relevant (e.g. plate on table vs. on wall);
- places each index base point at a distance from the centre of the disc, that distance giving an idea of its size, fingers in opposite directions.

The location point, the distance from it to each hand, and the disc plane orientation are three variables affecting the form and the meaning jointly. This is captured by arguments for the rule, each named after its interpreted meaning and typed accordingly: the location loc (of AZee native type `POINT`), the orientation of the plane in which it lies, represented by a normal vector $nrmsurface$ (type `VECT`), and a $size$ argument (of type `NUM`, compatible with lengths).

The linguistic rule `class-flat-round-large` is defined as a function of those parameters, yielding the body configuration (native type `POSTURE`) directly oriented and positioned in space according to the given argument values. Other rules accounting for traditional classifiers were of course found useful to encode the full corpus, e.g. `class-flat-surface-narrow` (configures the index and middle fingers straight in the plane of the surface, other fingers closed), or `class-parallel-lines` (four fingers extended and spread, thumb folded in).

We see how no supplementary concept like "classifier" is required to account for such entities. They just fall in the general formal approach to linguistic rule specification in AZee. But after describing *Mocap1* where many are present, an interesting question arises about AZee's account of such documented feature, and whether this category can be rediscovered or characterised in formal terms within the framework, which does not postulate their existence as a distinct case in the first place.

The first thing to notice is that they are modelled as functions yielding instant body postures, i.e. constrain parts of the body into timeless configurations. This is unlike the other linguistic rules defined so far, which yield timelines, i.e. intervals with durations allowing to control posture changes over time.

We also notice that while their arguments can vary, all of them have a location argument of geometric type `POINT`. They often have supplementary spatial orientation or size features, but none that can be made generic in the sense of being applicable to all of them, or interpretable with the same meaning. For example `class-flat-round-large` can be oriented in a plane (e.g. horizontal table on which a plate sits), and `class-person-standing` can be oriented to face a certain direction (index finger pad is understood to be the face of the person), but these are rule-specific interpretations. Besides, `class-flat-round-large` allows a size argument, whereas `class-person-standing` does not. Although a person could semantically be given a size, the form of the latter cannot linguistically depend on such variable, and is only meaningfully variable in location and index facing. The location argument though, is always present. It is in fact a requirement for use in certain structures as we are about to discover. Hence we can tentatively characterise a classifier (or a proform), in formal terms based on AZee specifications, as an AZee function with a signature compatible with `(loc:POINT)` $\rightarrow$ `POSTURE`.

Interestingly, the literature overwhelmingly speaks of *handshapes* to define classifiers, whereas in this framework any set of articulators can be constrained to yield the `POSTURE`. This can involve both hands (e.g. `class-flat-round-large`), more than the hand (e.g. handshape + forearm and its relevant orientation for a tree), or indeed anything else (e.g. the head or an arm alone). It is now worth looking at where these rules are used in the corpus, for insight on the structures that contain them.

4.2. Iconic Structures

Of the iconic structures listed in 2.1, we only found 5 role shifts in the corpus. This has to do with its descriptive and non-narrative genre, and leads us to concentrate on the other structures: those categorised by Johnston as "depicting signs".

Let us present examples of AZee code representing such structures. Expression 1 below corresponds to a standing person facing backwards (towards the signer), placed at a point named P defined earlier in the code. Following Johnston's guidelines, such an instance would be annotated "DSL" on the strong hand. Expression 2 shows a similar action on the strong hand, this time performed while the weak hand is holding a fixed anchor, namely a bridge instantiated as a narrow strip (straight index and middle fingers) oriented flat along a left–right axis. The strong hand placed a statue (instantiated as a standing person) facing backwards from behind the bridge. This is another type of classifier predicate, one that would be annotated with a "DSG" on the weak hand, and a "DSL" on the strong. Expression 3 encodes a production where the signer depicts the arm of a crane with a forward and upward movement of the strong hand in a cylindrical configuration ("C"). For Johnston, this would be an instance of a "DSS".

```
(1) :place-object
      'class
      .prf-person-standing
        'facing
        bwd
      'loc
      ^P

(2) :landmark-in-place
      'lm
      :class-flat-surface-narrow
        'loc
        [bridge anchor point]
        'hand
        w
```

```
        'nrmsurface
        dwn  % horizontal surface
        'dirsurface
        rgt  % strip direction
      'sig
      :place-object
        'loc
        translate| ^Mssp| bwd| tiny
        'class
        .class-standing-person
          'facing
          bwd
(3) :deploy-shape
      'path
      [fwd & upward from near chest]
      'class
      .class-cylindrical-medium
        'hand
        s
        'nrmsurface
        [fwd & upward]
```

In the encoded data overall, all iconic structures have been covered by combinations of the following rules[3]:

- `place-object`($class$, loc, $?nrmsurface$), mentioned in §2.2;
- `move-object`($class$, $path$), which depicts a movement of an instance of $class$ by making it follow the given $path$;
- `deploy-shape`($class$, $path$), which depicts a fixed line or surface in space, and produces a non-accelerated movement along the given path with a $class$ argument denoting the kind of shape or texture depicted;
- `mult-in-a-row`($item$, $path$), which depicts a set of countable objects of $class$, distributed along the given $path$;
- `mult-around`($item$, loc, $?nrmsurface$, $?radius$), which depicts a set of countable objects of $class$, distributed on a surface around loc;
- `landmark-in-place`(lm, sig), which means sig in a context where a landmark lm is active/present, for example an object in space;
- classifiers as per defined above, i.e. yielding postures from (at least) a location point like `class-person-standing` or `class-flat-round-large`;
- already known rules combining timelines, namely `each-of`/`all-of` (creating sets from parts), `simultaneous` (meaning two pieces $sig1$ and $sig2$ simultaneously true), `side-info` (adding non-focused/incidental information $info$ about a $focus$) and `instance-of` (specifying a $type$ for a focused element elt).

We first notice that except for several discovered of the classifier kind, these rules were already known or tentatively proposed in prior work, either to complete the *40 brèves* corpus or to address avatar animation with the Paula system (Filhol and McDonald, 2018). The result is that all of the spatial iconicity covered in the corpus is accounted for through combinations of only very few rules, forming a highly productive system.

Besides, all of the depicting structures encoded involved at least one of the five first items in the list above, which are rules with mandatory arguments of a geometric type (POINT, VECT or PATH). The classifiers, also frequent, verify this condition too given that they have at least a point argument loc. The others however, although seen in many iconic structures, are also largely used in non-iconic, non-spatialising contexts. In light of this, we propose the following criterion to qualify a rule as "iconic": having at least one mandatory geometric argument. This outlines a subset of "iconic" rules in the production set. But a noteworthy fact is that structures making use of such an iconic rule almost always contain other, non-geometric ones like `all-of` or `simultaneous`. The next section deals with this typical kind of mix.

4.3. Iconic vs. Lexical Continuum

At the Structures Level

The fact that no a-priori categorisation is assumed between rules makes the definition of lexical vs. iconic units almost redundant on the elementary level in AZee. But this again also consequentially allows to look for a formal definition of the categories accepted in the literature, and perhaps discover an explanation for the borderline cases, such as semi-PT, or reveal a continuum and a way to measure positions thereon.

In the literature, mixes of lexical elements and iconic structures must be analysed as a language phenomenon recruiting elements from different natures to constitute a new kind of structure. In AZee, no such separation exists in the first place, rather rules are applied as arguments for others, regardless of their iconic weight. Those of the so-called lexicon (no mandatory arguments) and those labelled iconic above (with mandatory geometric arguments) may just take part in building discourse expressions. For example, expression 2 in §4.2 contains

[3] Rule arguments prefixed with '?' are optional arguments, i.e. those for which a default value is specified, and which can therefore be left out when the rule is applied in an expression. Non-prefixed arguments are mandatory, i.e. must be given a value when the rule is applied.

4 rules, namely `landmark-in-place`, `class-flat-surface-narrow`, `place-object` and `class-standing-person`. The last three of those are iconic as per our criterion above (they have a mandatory geometric argument in their signature: all of them at least a point named loc), whereas `landmark-in-place` is not (its arguments are a timeline and a posture, neither geometric objects).

Now the utterance could typically be signed with the conventional sign for "statue" preceding the strong hand placement (standing person), without breaking the weak hand landmark. This would be encoded by expanding the sig part signed while the lm landmark is held, turning it into another typical pattern in our data:

```
:landmark-in-place
  'lm
  [class-flat-surface-narrow; no change]
  'sig
  :instance-of
    'type
    :statue
    'elt
    [place-object; same as prev. "sig"]
```

This inserts two rules in the expression, one lexical (`statue`, no arguments) and one not (`instance-of`, two timeline arguments). Neither verifies the "iconic" criterion proposed above, so the number of iconic rules is unchanged, while the total rule count went from 4 to 6. The ratio of these two figures N_{iconic}/N_{total} therefore went down from .75 to .5, reflecting a heavier use of non-iconic rules in the mix. We propose that this ratio conveniently measures the iconicity of a signed segment on a $[0..1]$ continuum, through counting rules used in its AZee-encoded expression, instead of defining a-priori labels and creating new categories as new mixes are discovered.

At the Discourse and Corpus Level

Intuitively, one would expect to observe very different AZee rules in the encoding of *Mocap1* corpora, compared to that of *40 brèves*. We did the comparison, which showed that among the ten most commonly used rules, three are identical. Those are `info-about` (focused information about a preceding topic), `side-info` (additional information given about a preceding focus), and `in-context` (setting a context for a following process). These three rules therefore appear most useful both in informative and descriptive contexts. For instance, `info-about` is used in *40 brèves* to specify that a man is French, and in *Mocap1* to specify that the lower part of a wardrobe has a sliding door.

However, some of the top 10 rules in *40 brèves* are not among the *Mocap1* top 10. Vice versa, *Mocap1* exhibits many `deploy-shape`, `prf-flat-surface`, `simultaneous`, `place-object`, unlike *40 brèves*. These rules are all present in both, but their ranking is drastically different.

We also observe differences in proportions of lexical signs between the two corpora. Considering an AZee rule "lexical" if it is specified without any mandatory argument, *40 brèves* contains 43% of them, whereas *Mocap1* only holds 10%. This allows to confirm, through AZee representations, that lexical signs are therefore less prevalent in *Mocap1*, and propose a formal metric for iconicity of a corpus.

In summary, the same production rules are used to represent utterances from the two distinct corpora, although these rules are applied in different proportions. First, this shows that discourse is almost never *only* about lexical elements. Second, that iconic structures can be formalised with the same structuring principles than the fully conventional ones. AZee encodings allows us to observe and quantify the proportion of these features.

4.4. Limitations and Propositions

A few elements still need improvement or refinement.

Firstly, we have encountered cases where the known production set was insufficient to fully account for an utterance, in other words where more rule discovery is needed. One is the desired semantic difference between actual and potential movement or action. For example, sliding doors of a cupboard were often signed with simultaneous flat surfaces moving, but the AZee expression structure `simultaneous(move-object(...), move-object(...))` reads as doors actually moving whereas the form seems not to allow this interpretation, and rather to impose that the doors *can* move. We hypothesise that a rule is to be discovered, that distinguishes these interpretations. The dynamics of the body and its detachment vs. engagement in the articulation should further be investigated, and may generalise to a "potentiality" rule usable everywhere in LSF, not just for such object movement.

Another case of missing rule was to justify movements repeated at fixed locations, when they are interpreted as continuous processes. While only four cases were encountered in the corpus, to do with flowing or running water, they are enough to reveal that we could not correctly encode such production with the known rule set. This actually takes us back to an unexplored case in our prior study on repeated movements. Two rules had emerged then (`mult-in-a-row` and `mult-around`), in which the countable repetition in the form mapped to a countable set of objects distributed in space. The missing case here means something different because there is no countable object set. Exploring

Figure 4: Example of an iconic structure formalized with a complex `all-of`

this path will probably not only provide a new rule capturing the missing meaning, but also as usual lead to the discovery of some subtle dynamic form difference, useful to refine our movement specifications and produce better animation quality.

Secondly, we observed recurring patterns of descriptions, and one particularly frequent: an `all-of` root applied to a list in which:

- every item applies the `simultaneous` rule, with both of its arguments $sig1$ and $sig2$ applying `deploy-shape`;
- the same rule is used for every $class$ argument contained;
- every $sig1/sig2$ pair recruits opposite hands for their deployment;
- every $path$ argument in a $sig1$ is symmetric to that in its sibling $sig2$.

This complex expression pattern is used 167 times to depict things like the multiple sides of a window or a carpet (e.g. fig. 4), the general case being a symmetric shape broken down into edges forming its outline. This is entirely compatible with the semantics of the encoded expression, as `all-of` creates and focuses a set from its parts, namely in such instances an oblong shape from three simultaneous edge deployments. It also produces the correct form, as was demonstrated with the "bedroom walls" example by Filhol and McDonald (2020).

The problem we report here is therefore not a theoretical one, rather to the contrary given the productivity of the system. But such a single piece of signing easily takes up 100 lines of code, with a lot of tedious repetition for the specification of the geometric parts. This to us reveals something impractical if not equipped with at least some form of shorthand. A new function can be introduced for this case, say "`deploy-edges-symmetric-shape`", factoring the few arguments needed to produce the whole structure. This is in a way a new form–meaning binding, but unlike those added to the production set so far, it only wraps an internal complex structure; it is not necessary to encode the utterances it covers. To be faithful to the formal AZee proposition unbroken so far, we therefore propose that such function be introduced to drastically collapse the code and make it more human-readable, but not considered as a linguistic rule of the production set. We will call such functions "wrappers", and encode them separately from the production set, but allow them anywhere in discourse expressions.

Lastly, the case of moving objects has in many places been insufficiently addressed by the only candidate rule at this point: `move-object`. It was introduced with Filhol and Mcdonald (2018) for animating simple object displacements ("DSM"), but it relies on a `PATH` argument, along which a point is dynamically moved for a `POSTURE` to depend on. The problem is that this does not allow things like rotations, e.g. a standing tree falling flat, or even complex compositions of movements, e.g. a car taking a turn while skidding. We understand here that any geometric change in position or rotation can in fact be signed to mean itself, not to mention that any synchronisation between changes can too, if there are several. No linguistic rule or set of rules can therefore cover all cases of object movement. To allow full coverage, some of the geometric specification of what moves and when will ultimately have to be provided in the encoded expression. This means allowing for segments where native AZee objects (points, vectors...) and operators (translations, rotations...) are freely written instead of constraining the expression to rules of the production set. In such segments, the encoding of the dynamic points and vectors transforming over time would be delegated to the author, not the result of linguistic rules.

However, we hypothesise that a generic form–meaning is still to be defined, devoid of geometric handling (factored out), but allowing precisely to interpret that such a segment is being performed, where free-hand geometric changes are happening. Again some subtle body dynamics marking the switch to this depicting mode is probably to be discovered, along with the expected and documented eye gaze detaching from the addressee and looking at the signed scene. We also propose that this switch might be very similar to that happening at the start of role shifts, for which a similar system is needed. Indeed the content of an iconically reported action cannot be the result of a defined linguistic rule; it must be provided by the author as it is the semantic contents itself. But the very fact that we can segment role shifts to interpret them as such may come from some dynamics. This would have to come from a rule enclosing the free-hand content, which itself would have to be provided with native specifications. In our view, this would formally make role shifts and DSMs similar to a wide extent, only one would be enclosing geometric changes and the other posture changes, perhaps even wrapped under a same rule.

5. Conclusion and Future Work

AZee had proved good coverage of news-type discourse, which only contained a few instances of iconic structures. This work has dealt with heavily iconic discourse, to encompass more of what SL produces when not restricted to a genre or use case. While this is often considered as a leap into a new kind of discourse, with new constructs and requiring new descriptive theories, the presented work showed that AZee on the contrary allows to do so without new concepts, and hardly even needs more rules than those already known prior to this work. While this is in itself of theoretical value, we also highlight the concrete benefit of entirely specifying output forms, which enables avatar animation of every encoded expression, including the fine dynamics and crucial eye gaze control where necessary.

In the future, we plan to expand the description in AZee of *Mocap1*. The goal would be to provide an AZee reference corpus for its depicting genre, similarly to what *40 brèves* brought for all-purpose, informative discourse. We should also address yet other discursive genres, in particular narrative discourse containing more role shifts, which were mostly lacking in *Mocap1*.

6. Bibliographical References

Mary Brennan. 1990. *Word formation in BSL*. Doctoral dissertation, department of linguistics, Stockholm University.

Diane Brentari and Susan Goldin-Meadow. 2017. Language emergence. *Annu. Rev. Linguist.*, 3:363–388.

Camille Challant and Michael Filhol. 2024. Extending AZee with Non-manual Gesture Rules for French Sign Language. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation*, (LREC-COLING 2024), pages 7007–7016, Turin, Italy. European Language Resources Association.

Marie Coppola and Diane Brentari. 2014. From iconic handshapes to grammatical contrasts: longitudinal evidence from a child homesigner. *Frontiers in psychology*, 5.

Kearsy Cormier, Sandra Smith, and Zed Sevcikova Sehyr. 2015. Rethinking constructed action. *Sign Language & Linguistics*, 18(2):167–204.

Christian Cuxac. 2000. *La langue des signes française (LSF): les voies de l'iconicité*. Faits de Langue. Ophrys, Paris.

Karen Emmorey. 2003. *Perspectives on classifier constructions in sign languages*, lawrence erlbaum associates, publishers edition. Psychology Press, Mahwah, New Jersey, London.

Michael Filhol and Mohamed Nassime Hadjadj. 2016. In *Proceedings of the workshop on Representation and Processing of Sign Languages*.

Michael Filhol and John McDonald. 2018. Extending the azee-paula shortcuts to enable natural proform synthesis. In *Proceedings of the workshop on Representation and Processing of Sign Languages*, pages 45–52, Miyazaki, Japan.

Michael Filhol and John Mcdonald. 2018. Extending the azee-paula shortcuts to enable natural proform synthesis. In *Workshop on the Representation and Processing of Sign Languages*.

Michael Filhol and John McDonald. 2020. The synthesis of complex shape deployments in sign language. In *Proceedings of the workshop on Representation and Processing of Sign Languages*, Marseille, France.

Nancy Frishberg. 1975. Arbitrariness and iconicity: historical change in american sign language. *Language*, page 696–719.

Ivani Fusellier-Souza. 2004. *Sémiogenèse des langues des signes: étude de langues des signes primaires (LSP) pratiquées par des sourds brésiliens*. Ph.D. thesis, Université Paris 8.

Mohamed Hadjadj, Michael Filhol, and Annelies Braffort. 2018. Modeling french sign language: a proposal for a semantically compositional system. In *International Conference on Language Resources and Evaluation*, Miyazaki, Japan. ELRA.

Trevor Johnston. 2024. Auslan corpus annotation guidelines. *Auslan Corpus*.

Trevor Johnston and Adam Schembri. 1999. On defining lexeme in a signed language. *Sign language & linguistics*, 2(2):115–185.

Scott K. Liddell. 2003. *Grammar, gesture, and meaning in American Sign Language*. Cambridge University Press.

Pamela Perniss, Roland Pfau, and Markus Steinbach. 2007. Can't you see the difference? sources of variation in sign language structure. *Visible Variation: Comparative Studies on Sign Language Structure*, 188:1.

Marie-Anne Sallandre. 2014. *Compositionnalité des unités sémantiques en langues des signes. Perspective typologique et développementale*. Thèse hdr, Université Paris 8.

Adam Schembri. 2003. Rethinking 'classifiers' in signed languages. In *Perspectives on classifier constructions in sign languages*, page 3–34, Mahwah, NJ: Lawrence Erlbaum Associates. Emmorey, K. D.

Ted Supalla. 1978. Morphology of verbs of motion and location in american sign language. *Proceedings of the Second National Symposium of Sign Language Research and Teaching*, page 27–45.

Rachel Sutton-Spence and Bencie Woll. 1999. *The linguistics of British Sign Language: an introduction*. Cambridge University Press.

Clayton Valli and Ceil Lucas. 2000. *Linguistics of American sign language: An introduction*. Gallaudet University Press.

Inge Zwitserlood. 2012. Classifiers. In *Sign language: An international handbook*, page 158–186. Mouton de Gruyter.

7. Language Resource References

LIMSI and CIAMS. 2020. *MOCAP1*. ISLRN 502-958-837-267-9. ORTOLANG (Open Resources and TOols for LANGuage) –www.ortolang.fr.

LIMSI and LISN. 2022. *40 brèves*. v2, ISLRN 988-557-796-786-3. ORTOLANG (Open Resources and TOols for LANGuage) –www.ortolang.fr.