# Representation and Synthesis of Geometric Relocations

**Michael Filhol[1], John McDonald[1,2]**
[1] Laboratoire Interdisciplinaire des Sciences du Numérique (LISN), CNRS, Université Paris–Saclay, Orsay, France
[2] School of Computing, DePaul University, Chicago, IL, USA
michael.filhol@cnrs.fr, jmcdonald@cdm.depaul.edu

**Abstract**

One of the key features of signed discourse is the geometric placements of gestural units in signing space. Signers use the geometry of signing space to describe the placements and forms of objects and also use it to contrast participants or locales in a story. Depending on the specific functions of the placement in the discourse, features such as geometric precision, gaze redirection and timing will all differ. A signing avatar must capture these differences to sign such discourse naturally. This paper builds on prior work that animated geometric depictions to enable a signing avatar to more naturally use signing space for opposing participants and concepts in discourse. Building from a structured linguistic description of a signed newscast, they system automatically synthesizes animation that correctly utilizes signing space to lay out the opposing locales in the report. The efficacy of the approach is demonstrated through comparisons of the avatar's motion with the source signing.

**Keywords:** AZee, Sign Language, Avatar, Geometric Constructions, Relocation

## 1. Introduction

In the history of Sign Language study, there has been a long tradition of distinguishing the units composing signed sequences between conventional "lexical" units and the productive "depicting" ones (Cuxac, 2000; Johnston, 2010). The former are usually found in picture dictionaries, each given a fixed/canonical form (citation form), and labelled with a written language word equivalent (gloss) used in video stream annotations. These sometimes allow for variations such as hand location or movement direction change, which theories tend to analyse as spatial agreements ("relocated signs", "directional verbs") (Lillo-Martin and Meier, 2011; Quer, 2011; Wilbur, 2013), but much of the associated form remains invariant.

The latter usually escape the traditional grammatical qualifiers like verb or behaviour like agreement, and are described less systematically than, say, manual parameter values for signs. They make a productive use of space, and their iconicity is usually accepted, if not put forward, as their primary feature (Cogill-Koez, 2000; Liddell, 2003a; Liddell, 2003b). In contrast, the fact that the dictionary-type entries are often iconic too is mostly incidental to their common theoretical descriptions (Baus et al., 2013; Padden et al., 2013). Interestingly, we note that in closed vocabulary annotations, labels used for depicting units usually describe a wide generic category, e.g. "classifier placement" or "size and shape specification", covering highly variable forms and meaning. In contrast, glosses have a specific meaning, generally conveyed by the written word, and are coupled with forms that are largely invariant.

These categories can inform the way an avatar will need to move to correctly communicate the intended meaning. As signing is studied, however, the division between these two categories can become difficult to identify in practice due to the manner in which signers structure discourse. Further, computer animation sometimes has its own questions, techniques and considerations that make it difficult to align synthesis with traditional linguistic divisions.

This paper presents a linguistic framework and corresponding animation techniques that support the various relocating structures described above, and demonstrates their expressive efficacy by applying them to real-life discourse. The approach is validated by mapping the components of these descriptions to animated forms that are combined on the avatar to reproduce such discourse.

## 2. Related Prior Work

Over the recent past, we have been working to bridge the Sign description model AZee with the Paula Sign synthesis platform (Filhol et al., 2017). Both systems share two fundamental organising principles: multi-linear scheduling of motion processes as opposed to synchronous sequences, and a holistic view of the body as opposed to a-priori partitioning of the body articulators. The bridge leverages the principle of "the coarser the better", by which larger animation blocks yield more natural-looking animations, while also recognising the infinite variability of some SL productions, caused for example by the use of a continuous space in geometric constructions. Based on a top-down search for parameterised "shortcuts", it trades off composition of minimal units with reuse of coarser ones when available.

The stones were largely unturned in the area of depicting, i.e. non-glossable, productions, which led us to explore phenomena such as placements of objects (Huenerfauth, 2006; López-Colino and Colás, 2011; Filhol and McDonald, 2018) and deployments of shapes (Filhol and McDonald, 2020). The latest test has managed the synthesis of a description comprised mostly of depicting units, rendering an utterance of 20+ seconds (McDonald and Filhol, 2021). With this work, we could demonstrate that AZee provides efficient abstraction of articulated forms into semantic combinations, covering much of the language, even dealing with features left uncovered by other approaches. Plus, we have shown that this coverage was achieved with a limited number of combining production rules, making the case that AZee is a productive system. This point was recently generalised to the 120 videos of the *40 brèves* corpus[1],

[1]Parallel corpus of news items in French and French Sign Lan-

totalling 1 hour of LSF, and manually encoded in AZee (Challant and Filhol, 2022). The resulting AZee expressions were found to cover 94% of the signing time, with just 30+ combining production rules.
We also demonstrated that Paula was able to render natural and meaningful movement from the compiled AZee discourse expressions directly. This is thanks to AZee's specification not only of the forms to articulate but also the timing information to synchronise them, a feature that is essentially absent from gloss strings.
Finally, we observed that production rules that were initially found to order discourse constituents, were actually useful on all levels and in all types of productions, even the purely geometric ones. For example, rules `all-of` and `each-of` structured the table scene description on various levels of nested lists. For example, the utterance contained an `each-of` list of items on a table, which itself contained lists of plates, chairs, glasses, etc. Interestingly, `all-of` was also used to create a "knife & fork" compound to form a sign for "cutlery", which is close to the level usually considered lexical. We even found this rule inside units typically annotated as non-lexical and productive, as we studied shape deployments[2]. So we ended up with rules that are widely applicable regardless of divisions between lexical and non-lexical, between glossable and non-glossable, or between depicting and non-depicting.
This paper pushes our research on the use of space and geometric constructions further, this time taking advantage of the recently published corpus of 120 AZee expressions mentioned above.

## 3. Relocations in AZee

Looking at the AZee expressions recently added to the *40 brèves* repository[3], we can extract two types of AZee patterns that result in geometric relocations in the produced forms.
One is the use of the rule `about-point`, whose arguments are a locus point *pt* of the signing space and a score (signed utterance) *locsig*. Its form description involves a torso rotation and a brief glance towards the locus at the very beginning of the rotation, and *locsig* being signed normally, although seemingly displaced towards *pt*. The gestural units perceived as "at the locus" are performed with the torso rotation sustained until their completion. Its semantics is that *locsig* is information about what is anchored at *pt*. It is very similar to `info-about` (score arguments *topic* and *info*), except that the topic about which the information is given lies in the reference of the anchor, not in a signed argument.
A second AZee pattern found to relocate items is for a rule to use a geometric argument directly, e.g. a point of the signing space, on which its produced form depends. This generally affects the placement or movement of the more distal effectors controlled by the rule, like finger tips or hands, with less torso involvement and no gaze redirection. This pattern will typically be used for instances of units traditionally called "relocated", "directional" or "pointing" signs. Note that such arguments in AZee can be defined as optional (and given a default value, i.e. one to use if none is given when the rule is applied), or mandatory.
Let us take the "1R-JP" entry of the *40 brèves* corpus to exemplify these cases. The news report in question describes a French citizen who was held hostage in Iraq for 35 days, was released, and is about to return to France.

> *Le Français Bernard Planche, 52 ans, retenu en otage en Irak pendant 35 jours, a retrouvé samedi la liberté près de Bagdad, et devrait très prochainement regagner la France.*

The signed translation has the following structure (see fig. 1 for snapshots):

1. Iraq established on the right of the signing space (from the beginning to 2 s);
2. Bernard Planche, including nationality and age, established on the left (2–7.5 s);
3. reporting that he was held hostage 35 days, still signed on the left-hand side except for the last sign meaning "hostage", which ends on the right (7.5–11 s);
4. pausing with hands retracted together (11–11.3 s);
5. establishing a place near Baghdad that he is freed from, back on the right-hand side of the signing space (11.3–16 s);
6. reporting that he will be returning to France soon with a sign for "return" performed from right to left (from 16 s to the end).

The overall expression, given in file "1R-JP.az" of the corpus repository, connects the six segments mostly by means of the `context` rule, which we have covered in prior publications.
The first pattern with `about-point` is found in three instances, namely over segment 1 (locus on the right), over the pair of segments 2 and 3 (on the left), and over segment 5 (on the right again). The AZee expression for segment 1 for example is the following, where `category` with arguments *cat* and *elt* means "*elt* as an instance of *cat*", `pays` means "country" and `Irak` "Iraq". The `Rssp` reference stands for a point on the right of the signing space.

```
:about-point
  'pt
  ^Rssp
  'locsig
  :category    (*)
    'cat
    :pays
    'elt
    :Irak
```

guage, available at https://www.ortolang.fr/market/corpora/40-breves.

[2] See the description of the wall sections in (Filhol and McDonald, 2020)—direct link to video: https://zenodo.org/record/3904430/files/bedroom-walls.mp4?download=1.

[3] Permanent link to the corpus version used for this work: https://hdl.handle.net/11403/40-breves/v2.

1. “country” 2. “person” 3. “hostage” 4. (pause) 5. “nearby” 6. “return”

Figure 1: Video snapshots of 1R-JP, one per segment described in section 3.

The second pattern is seen in various places. Below are three excerpts from the whole expression.

```
:otage     :retour     :pointage index
  'loc       'source     'target
  ^Rssp      ^Rssp       ^Lssp
             'dest
             ^Lssp
```

The first expression uses rule `otage`, which means “hostage” and produces the appropriate sign, allowing for a location point argument *loc* used to control the final position of the hands. The corresponding instance in the video is at 8 s (fig. 1.3). The `Rssp` argument point (signer’s right-hand side of signing space in AZee) accounts for the fact that the sign ends on the right-hand side, which is interpreted as the location of the associated event. Note that neither the spine rotation nor the leading glance of the previous pattern are involved.

Similarly, the second expression uses rule `retour` for “return”, which allows two point arguments *source* and *dest*, affecting the start/end points of the finger movement. The corresponding instance is visible at 17 s in the video (fig. 1.6).

Finally, the third expression captures the index pointing sign with rule `pointage index` and `Lssp` as its obligatory argument *target*, modifying the direction of the index. The corresponding instance is at 6 s in the video, and a similar one, pointing to the right (`Rssp`), can be found at 14 s.

Let us now look more closely at the AZee expressions shown above, and in particular how the part to which the geometric change is applied is inserted in the expression. The `about-point` expression, generating pattern 1, locates Iraq around `Rssp`. That geometric change is applied to what would otherwise be encoded as the sub-expression marked `(*)`. This sub-expression is inserted in the `about-point` operation with no change to it whatsoever. What the AZee formalism is telling us here is that this is an external change; nothing was changed inside of the relocated discourse itself. In contrast, the other three expressions relevant to the second form pattern, exhibit changes to, or the addition of, internal parameters of the original expressions. With no geometric change, `:otage` and `:retour` would be valid expressions, signed in the middle of the signing space (often referred to as neutral). In our current definition of `pointage index`, parameters could not be left out (one does not point at no target using that form), but any geometric change to it would still be encoded as a change to its *target* argument. This contrast in the formal representations indicates that geometric changes can be categorised in two groups: the external changes, i.e. operations applied to expressions as a whole, and internal ones, i.e. changes to the expression itself.

## 4. Animating Geometric Changes

In order to animate the full 1R-JP expression, the avatar must handle all of the contained geometric changes. We will see that the two types of modifications, *internal* and *external* defined above, need to be treated in distinct ways due to differences in how the human body moves to make these changes happen.

### 4.1. Animating External Changes

The `about-point` construction occurs several times in the discourse as elements relating to Iraq are placed to the right in segments 1, 3 and 5, and elements relating to France are placed to the left in segment 2. Analyzing the motion of the signer’s body during these segments reveals the following common characteristics:

- the whole production of the sign is rotated to the relevant side of signing space with no other modification to the sign movement, supported by a shallow turn of the torso;
- the hands are not re-oriented to remain facing forward or to the side relative to the viewer, as can be clearly seen in the production of `une personne` in fig. 1.2 where the handshape is now facing to the right instead of forward as it would be in an unaltered production of `une personne`;
- there is an initial glance of short duration ($< .5$ s) at the beginning of the segment towards the target area.

To effectuate this change, all the avatar has to do is to rotate the torso and the shoulder joints without changing the rest of the motion in the production. Such a change is called a forward kinematic (FK) movement since the angles are being directly affected (Parent, 2012).

In many ways, this technique is similar to the forward kinematic rotations that were used in (Wolfe et al., 2012) to animate turn taking in reported dialog, except that the rotation of the torso is shallower because the signer is not actually shifting their torso to assume the position of a participant, nor is the signer’s gaze locked on the target. In that prior

(a) Citation form (b) Relocated form

Figure 2: Relocation of `une personne` to the left.

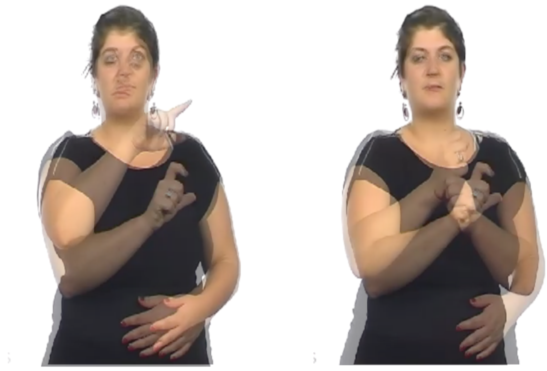

Figure 3: Variance in placement of signs.

work, the torso rotation carried all of the motion, whereas here, more of the rotation happens at the shoulder joints. In summary, the Paula avatar system can respond to AZee's `about-point` rule applied to a point $p$ by taking the following actions:

1. compute the overall angle $\theta$ that needs to happen to shift the center to align with $p$;
2. apply $.45\theta$ to the spine's twist angle;
3. apply the remainder of $\theta$ to rotate the shoulder joints, and thus the arms, laterally;
4. raise or lower the elbow using the shoulder's swivel angle (Parent, 2012) by $.25\theta$ so that as the arm moves across the body the upper arm will not collide with the torso;
5. a small rotation of $.15\theta$ in the wrist is introduced to counter the arm's swivel rotation without bending the wrist too much;
6. apply a gaze of short duration with the rotation largely centered on the eyes rather than on the avatar's neck.

The results of these actions can be seen in fig. 2 wherein `une personne` is rotated to left signing space.

This technique stands in sharp contrast to the inverse kinematics (IK) computations used in prior work on geometric placements in both depictions (Filhol and McDonald, 2020) and agreement structures (Wolfe et al., 2012). The goal of IK is to compute the armature's angles based on the desired hand placement and orientation of the hand or finger.

There is another observation that supports the choice of FK here versus the IK techniques chosen for depictions. The placements of the hands show significant variance in position as the signer is describing Iraq and France on either side. Figure 3 compares the placement of `une personne` to `pointage index` in the left image, and to the form meaning "years old" generated by the rule `cpt-années` in the right image. These placements all arise in discourse segment 2. The overlays show significant differences in both vertical and lateral positions of the hands. This variance is generally higher than would be found in depictions. In addition, the gaze is far less engaged than the sustained eye contact and neck rotations seen in that prior work.

### 4.2. Animating Internal Changes

The other three structures that cause geometric changes to signing in this discourse are `retour`, `otage` and `pointage index`. However, instead of the geometric information for the changes being provided at a higher level in the AZee expression (i.e. a containing structure like `about-point`), the geometric locations come internally in these expressions rather than externally. The only differences between them are that `retour` takes both a `source` and a `target` point, whereas the other two only take one[4]. To see how these changes occur on the body and may be implemented in the avatar, let us take `retour` as an example. Its motion is different from the prior `about-point` relocations since it affects the start and the end of the motion differently. It has very little accompanying torso rotation, and no redirection of the gaze. For these reasons, this motion falls under the traditional agreement pattern and Paula can shortcut to the system (Wolfe et al., 2012), wherein the artist not only animates a citation form of the sign, but provides a generic curve profile for the arms to follow, which is retargeted using IK. The rule `otage` can also be animated in this way.

The pointing signs actually have the same type of motion in the arms and torso as the `about-point` construction, and are even accompanied by the same type of glance. Since the handshape is pointing towards the target rather than being placed at the target, it can be animated with the same FK process by the avatar. The fact that the location is an internal parameter of the rule `pointage index`, instead of being a general process imposed externally, allows us to use a different animation scheme in this instance. It is worthy to note that the resulting avatar motion is different from the IK process that has been used previously for pointing signs in ASL (Wolfe et al., 2012). The pointing motions in

[4]Actually, the sign for `otage` is also contained in a surrounding `about-point` structure and therefore has its starting point altered externally. It also has one other interesting feature: the signer changes the form to mimic the action of the person being taken hostage at gunpoint causing changes to handshape, hand orientation and other parameters. These elements were included in the pre-animated sequence for `otage` for this demonstration.

that work were similar to the modifications to `otage` and `retour` above, and were accompanied by a stronger fixed gaze. Further study into pointing signs will be needed to determine whether the same type of motion and glance are used more widely across sign languages.

## 5. Discussion

The full synthesized discourse applies these processes to seamlessly synthesize the geometric changes discussed above. It can be viewed at `https://zenodo.org/record/6547654` and frames corresponding to the signer's positions in 1 are shown in fig. 4.

Throughout the animation, AZee rules such as `info-about`, `context` and `side-info` not only structure individual phrases but contribute associated nonmanual signals and rhythm that natively provide prosodic structure to the discourse. Also, we can notice an expressive mixing of the various patterns for locating signs or pieces of the discourse:

- segment 1 uses `about-point` with none of its content affected by internal changes;
- in segment 2, rule `pointage index` is applied inside an `about-point`, both with the same point argument `Lssp`;
- in segment 3, rule `otage` is applied to `Rssp` while nested in an `about-point` applied to `Lssp`;
- in segment 6, rule `return` uses two internal point arguments outside of any `about-point` operation;
- the very last unit, meaning "France", is performed with neither pattern applied.

All of these combinations are made possible by the existence of a single generic rule `about-point`, in addition to individual sign arguments.

Here again AZee proves to be a very productive system, i.e. very expressive, describing a wide range of types of communication while consisting of a limited set of combining rules. It is important to note that in the 120 examples contained in the *40 brèves v2* corpus of AZee expressions, about-point ranks third in application count (531 applications in total), after `info-about` and `side-info`. Thus, `about-point` is extensively used in common discourse, and the phenomenon of turning to a point reference after a quick eye glance is not rare. In contrast, the other pattern of direct point arguments, traditionally viewed as "agreement" structures, is surprisingly anecdotal, and do not generalize easily.

## 6. Conclusion and Future Work

The synthesis of the news report "1R-JP" provides a rich environment to test the ability of the AZee–Paula bridge to synthesize highly structured discourse featuring different types of geometric relocations. In particular, we can see that if the avatar is to convincingly reproduce such signing, it much be able to handle both precise IK styles of relocations used in depictions, as in prior efforts, and FK styles as in the `about-point` relocations explored here.

In addition, through this work, it can be seen that an avatar will struggle to animate even discourse such as the "1R-JP" example, that is of the kind that is often considered glossable, if the linguistic description is limited to a stream of glosses, albeit provided with a set of instructions for relocation of each gloss. Some relocations arise from external processes, i.e. larger in scope, and others from internal ones, i.e. only applying to the sign or a part of it. This distinction greatly impacts how the avatar must respond to produce natural movement.

This paper is a continuation of our effort to explore the description and synthesis of sign language, and there are still many unexplored avenues that will be pursued in future work. But, as in prior work in this area, we see that the interplay between linguistics and animation continues to be a rich field of study yielding insights on both sides.

## Acknowledgement

This work is supported in part by the EASIER (Intelligent Automatic Sign Language Translation) Project. EASIER has received funding from the European Union's Horizon 2020 research and innovation programme, grant agreement n° 101016982.

## 7. Bibliographical References

Baus, C., Carreiras, M., and Emmorey, K. (2013). When does iconicity in sign language matter? *Language and cognitive processes*, 28(3):261–271.

Challant, C. and Filhol, M. (2022). A first corpus of azee discourse expressions. In *Language Resources and Evaluation Conference (LREC), Representation and Processing of Sign Languages, Marseilles, France*.

Cogill-Koez, D. (2000). A model of signed language classifier predicate as templated visual representation. *Sign Language & Linguistics*, 3:209–236.

Cuxac, C. (2000). *Langue des signes française, les voies de l'iconicité*, volume 15–16. Ophrys.

Filhol, M. and McDonald, J. (2018). Extending the azee-paula shortcuts to enable natural proform synthesis. In *Workshop on Representation and Processing of Sign Language, International Conference on Language Resources and Evaluation (LREC)*, pages 45–52.

Filhol, M. and McDonald, J. (2020). The synthesis of complex shape deployments in Sign Language. In *Proceedings of the 9th Workshop on the Representation and Processing of Sign Languages*, Marseille, France.

Filhol, M., McDonald, J., and Wolfe, R. (2017). Synthesizing Sign Language by connecting linguistically structured descriptions to a multi-track animation system. In *11th International Conference on Universal Access in Human-Computer Interaction (UAHCI 2017) Held as Part of HCI International 2017*, volume 10278 of *Universal Access in Human–Computer Interaction. Designing Novel Interactions*, Vancouver, Canada, July. Springer.

Huenerfauth, M. (2006). *Generating American Sign Language Classifier Predicates for English-to-ASL Machine Translation*. Ph.D. thesis, Pennsylvania, USA.

(a) "country" (b) "person" (c) "hostage"

(d) "pause" (e) "nearby" (f) "return"

Figure 4: Frames from the Paula Animation of 1R-JP.


Johnston, T. (2010). From archive to corpus: Transcription and annotation in the creation of signed language corpora. *International journal of corpus linguistics*, 15(1):106–131.

Liddell, S. K. (2003a). *Grammar, gesture, and meaning in American Sign Language*. Cambridge University Press.

Liddell, S. K., (2003b). *Perspectives on classifier constructions in sign languages*, chapter Sources of meaning in ASL classifier predicates, pages 199–220. Erlbaum, Mahwah, NJ.

Lillo-Martin, D. and Meier, R. P. (2011). On the linguistic status of 'agreement' in sign languages. *Theoretical Linguistics*, 37(3/4):95–141.

López-Colino, F. and Colás, J. (2011). The synthesis of lse classifiers: From representation to evaluation. *Journal of Universal Computer Science*, 17(3):399–425.

McDonald, J. and Filhol, M. (2021). Natural synthesis of productive forms from structured descriptions of sign language. *Machine Translation*, 35(3):363–386.

Padden, C. A., Meir, I., Hwang, S.-O., Lepic, R., Seegers, S., and Sampson, T. (2013). Patterned iconicity in sign language lexicons. *Gesture*, 13(3):287–308.

Parent, R. (2012). *Computer animation: algorithms and techniques*. Newnes.

Quer, J. (2011). When agreeing to disagree is not enough: further arguments for the linguistic status of sign language agreement. *Theoretical Linguistics*, 37(3/4):189–196.

Wilbur, R. B. (2013). The point of agreement: changing how we think about sign language, gesture, and agreement. *Sign Language & Linguistics*, 16(2):221–258.

Wolfe, R., McDonald, J. C., Toro, J., and Schnepp, J. (2012). A proposal for making corpora more accessible for synthesis: A case study involving pointing and agreement verbs. In *sign-lang@ LREC 2012*, pages 163–166. European Language Resources Association (ELRA).