A Model for Sign Language Grammar

Boris Lenseigne*,Patrice Dalle*

*IRIT-TCI, Université Paul Sabatier
Toulouse
France

Abstract

Sign language processing is often focused on individual sign processing which means that one needs an exhaustive description of those gestures. In this paper we propose a computational model of sign language utterances based on the construction of the signing space and a grammatical description: the iconicity theory. This model leads to an entity-and-relationship description of the meaning of the discourse which can be attached to a qualitative description of the gestures that where used to produce that utterance.

1. Introduction

Sign languages (SL), such as the French sign language (FSL), use gestures instead of sounds to convey a meaning, they are deaf peoples natural language. Unlike oral languages, SL are characterised by a great multilinearity due to the fact that the signer can simultaneously use several body parts to communicate: hands configuration, localisation and motion, facial expression, body motion,. . .

Most of the time one considers two levels of language: standard utterances that only use standard signs, the ones that can be found in dictionaries and iconic utterances where most of the meaning relies on iconic structures. Iconic structures are widely used in spontaneous SL so they need to be taken in account in automatic SL processing systems.

Due to the fact that iconic sentences are often considered as being a pantomime, many linguists deny them the language status so most of existing works focus on standard signs description (Stokoe et al., 1965)(Prillwitz and al., 1989). But C. Cuxac (Cuxac, 1999) has shown that iconicity is used in standard sentences as well so that it has been possible to propose a global description of the FSL grammar. This grammar is based on the construction of the signing space (the space surrounding the signer where signs are produced).

Such a dichotomy can be found in automatic SL processing systems too. Most of the time, they focus on standard sentences without taking in account the spatial structure of the sentence, as (Liang and Ouhyoung, 1998)(Vogler and Metaxas, 2003) in the field of automatic translation. Only a few ones pay attention to that aspect of SL: (Braffort, 1996)(Sagawa et al., 1997) for translation or (Lebourque and S., 1999)(Huenerfauth, 2004) for automatic SL generation.

2. Goal of the paper

In the case of the analysis of a SL image sequences, significant informations on sign parameters are difficult to identify without using a specific acquisition system (several cameras, high resolution cameras,. . .). Considering this, a good way to help to reconstruct the meaning of the discourse is to have a preliminary idea of the entities that where evoked in the sentence and of their semantic relationships. Moreover a grammatical model of the structure of the SL sentence allows us to use a prediction/verification approach where low level image processing is only used to verify whether the data corroborates the hypothesis concerning the meaning of the sentence.

The goal of the paper is to present a model of the structure of FSL that enables us to compute components (hands, face, body) descriptions, from hypothesis on what is going to be said in the sentence. The use of such a model should lead to descriptions suitable for image processing possibilities.

2.1. Our approach

On the grammatical level, the meaning of the FSL sentence can be understood considering the entities and their semantic relationships. A cognitive approach of language production (Thom, 1972-1977)(Langacker, 1987) has shown that there were only a few possible relationships that could be evoked in a sentence:

- *temporal relationships* that can be either absolute or relative to the current time of the production.
- *Spatial locations* between two entities.
- *Actions* that can link several entities.

Moreover, the iconicity theory by C. Cuxac makes it possible to identify those entities and relationships from the construction of the signing space and describes the way the signer uses gestures to build that space.

Our model integrates the knowledge of FSL grammar as it is described in this theory. To avoid the risk of an abusive simplification implied by an incomplete description, it focuses on representing a subset of that grammar. This subset concerns sentences produced in the context of a timetable, that means sentences that brings on play persons, places, dates and actions.

2.2. General structure of the model

Our FSL grammatical model comprises two parts. The first part consists in a *symbolic representation of the signing space* that contains the spatial organisation of the entities and their relationships according how they were evoked in the sentence, the second one in a *behaviour model* that uses the FSLgrammar description to define the way components of the signer's body have to behave in order to modify the signing space in a particular way.

Entity	Potential relationships			
	Relative temporal location	Absolute temporal location	Spatial location	Action
Date	×	×		
Place			×	
Animate			×	×
Person			×	×
Action		×		×
Object			×	×
Implicit			×	×

Table 1: Different kinds of entities that may be evoked in a signed sentence and relationships that can exist between them.

3. Modelling the signing space

3.1. Entities and relationships

In the FSL, entities are evoked through signs and located in the signing space so that their relative position will correspond with the spatial relationships between those entities in the real world. Temporal relationships are evoked through entities that are located on "time lines". Binary actions are evoked through directional verbs and more complex ones by grammatical structures called "transfers" (Cuxac, 1999). The different kinds of entities depend on the kind of the relationships in which each entity may be involved: *dates* can be involved in temporal relationships, *places* in spatial relationships; *animates* can perform an action or be located relatively to another entity, *actions* can be referenced as a moment in time or as one of the protagonists of an action. The specificities of the FSL grammar require to consider some additional kind of entities: one needs to make a distinction between entities that whenever involved in a complex action are evoked by the signer taking their role(*persons*[1]) and the entities that cannot be evoked that way (*objects*). Finally, due to the temporal ordering of the signs, one needs to take into account the case of actions that are evoked before one of their protagonists, the type of this entity is *implicit*.

Table 1 gives an overview of the different kinds of entities that can be evoked depending on the relationships that may link them together.

3.2. Signing space representation

The symbolic representation of the signing space consists of a cube surrounding the signer, regularly divided into *Site*(s)[2]. Each location may contain a single *Entity*, each *Entity* having a *Referent*. A *Referent* is a semantic notion that can be found in the discourse. Once it has been placed in the signing space, it becomes an *Entity* and has a role in the sentence. So that, building a representation of a sign language sentence consists in creating a set of *Entities* in the *SigningSpace*. The meaning contained in this signing space construction is represented in terms of *Entities*(s) whose *Referent*(s) can have successively different *function*(s) during the construction of the sentence (*locative, agent,*...). A set of rules maintains the consistency of the representation by verifying that enough and coherent information has been provided when one needs to create a new entity in the signing space. The figure (fig. 1) describes the global architecture of the model in UML notation standard.

[1]In FSL, persons are not necessarily humans, they can be assimilated to animals or even objects of the real world in humorous stories for example.

[2]Terms written using a *slanted* font are elements of the model.

4. Behaviour model

4.1. A short survey of FLS Grammar

The rules of FSL grammar we consider intend to describe each possible modification of the signing space. As modifying the signing space only consists in creating new entities, our model focuses on the gestures that are used to create those entities. Without taking in account lexical knowledge, it is not possible yet to make a distinction between entities that are neither dates nor actions. So that creating such an entity relies on a generic mechanism. Creating an entity of a given type relies on the following mechanisms:

- Creating a generic entity: generally speaking, entities are created and localised in the signing space by signs that can be performed either directly in the desired location or localised on the signer's body for lexical reasons. In the second case, the production of the sign is followed by an explicit designation of the desired location.

- Creating a date: in our reduced context, dates are explicitly evoked by standard signs, performed in a neutral location (if front of signer's chest) and located simultaneously on one of the time lines.

- Creating an action: binary actions are evoked through directional verbs, which imply gestures that explicitly connect two locations containing entities in the signing space. For complex actions, "great iconicity" structures such as those where the signer plays the role of one of the action's protagonist have to be used. Such complex actions do not appear in the context of our application.

The formalisation of that grammar relies on the fact that each of those mechanisms can be described by a gesture sequence.

4.2. Describing the construction of the signing space

A modification in the signing space is defined by the kind of the entity that is created and its localisation. The behaviour model attaches to each kind of entity a gesture sequence that describes the state of the components involved and the way they are synchronised.

The computational representation of that grammar relies on a description logic formalism and uses the CLASSIC knowledge representation system (Brachman and al., 1991). This system allows the representation of FSL grammar as a set of hierarchically organized concepts.

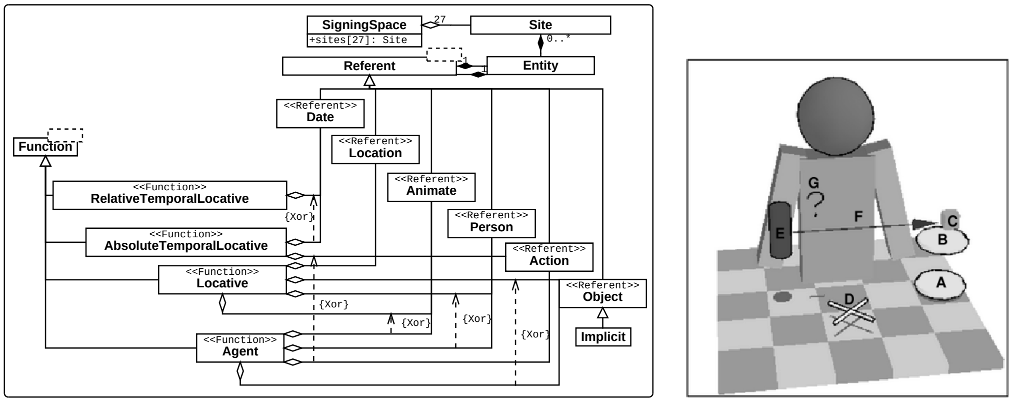

Figure 1: **On the left:** UML class diagram from the symbolic representation of *SigningSpace*. *SigningSpace* is regularly divided into *Sites*. Each *Site* may contain a single *Entity* whose *Referent* will have several *Function*(s) during the production of the sentence.**On the right:** An example of the construction of the signing space that corresponds to the FSL question (the order of the signs has been respected): *"In the city of Toulouse (A), in the movie theatre called Utopia (B), the movie that plays (C), on Thursday February 26th at 9.30 pm (D), the one (E) who made it (F), who is it (G) ?"*. In this figure, one can see that the sentence is realized by putting the different entities in place in the space surrounding the signer and that their respective place is relied to the semantic relationships among these items.

Concept are structured objects constituted with roles (concepts of a given type) and associated with automatic inference mechanisms and user-defined propagation rules.

4.2.1. Formalisation of an entity creation

With the description logic formalism, describing the creation of an entity consists in defining a set of concepts with specific constraints on some of their roles:

1. The concept representing the creation of an entity is called *ACTS* (ACtion Transforming Signing space). It is described by a location, a temporal interval and a gesture sequence.

2. Gesture sequences consist in a list of component descriptions associated with constraints on the values of the component roles.

3. Additional knowledge propagation rules concern vertical information propagation from an *ACTS* description to gestures defined in the corresponding sequence (e.g. the localisation of the hand must be the same as the one of the entity). Horizontal information propagation mechanisms are used between different gesture descriptions in the same sequence (e.g. both hands must have the same location). Finally gestures synchronisation rules are based on Allen's algebra operators (Allen, 1990).

This formalisation leads to a global representation of the FSL grammar as a concept hierarchy associated with additional propagation rules sets.

4.2.2. Global structure of the FSL language model

The concept hierarchy that describes the FSL grammar model is given in figure 2: for each kind of entity, there is a specialisation of the *ACT* concept with a specific *GestureSequence*. This sequence can be derived depending on the different ways to create an entity of that type. Gestures that can be found in *GestureSequence*s are specialisation of generic *Component* descriptions that include additional constraints on their roles.

5. Using the model

5.1. Inference during a FSL sentence analysis

At runtime, individuals, concept instances, are created. The information is propagated through those individuals according to the concepts descriptions and the rules that constitute the behaviour model.

Preliminary evaluations have shown that our model enables the inference of component descriptions from knowledge on the creation of a new entity in the signing space. During a prediction/verification analysis cycle (Cf. 2.), those descriptions are used to determine which properties have to be verified on the data. If the data provide additional informations, CLASSIC provides an additional classification mechanism that automatically changes the type of the individual when its properties correspond to the ones of a different, but more specialised, concept.

5.2. Expressiveness of the model

Our FSL grammar model was originally intended to represent a subset of that grammar, so that it is useful to point out which kind of sentences can be described with that model and which cannot.

The symbolic representation of the signing space does only represent the spatio-temporal organisation of the entities that where evoked so that the only limitation of that part of the model is related to the signing space discretisation. However, the construction of the signing space cannot be considered as an exact transcription of the meaning of the sentence: as the model does not take into account the lexicon, it is not possible to determine exactly the type of an entity so that it has to be infered from the successive functions of that entity during the production of the sentence (table 1). Such a mechanism can lead to the misunderstanding of the sentence as some entities will be assigned a wrong type due to the fact that the signed sentence only evokes the significant set of entities and relationships according to the intention of the signer.

In our evaluation, the behaviour model only represents

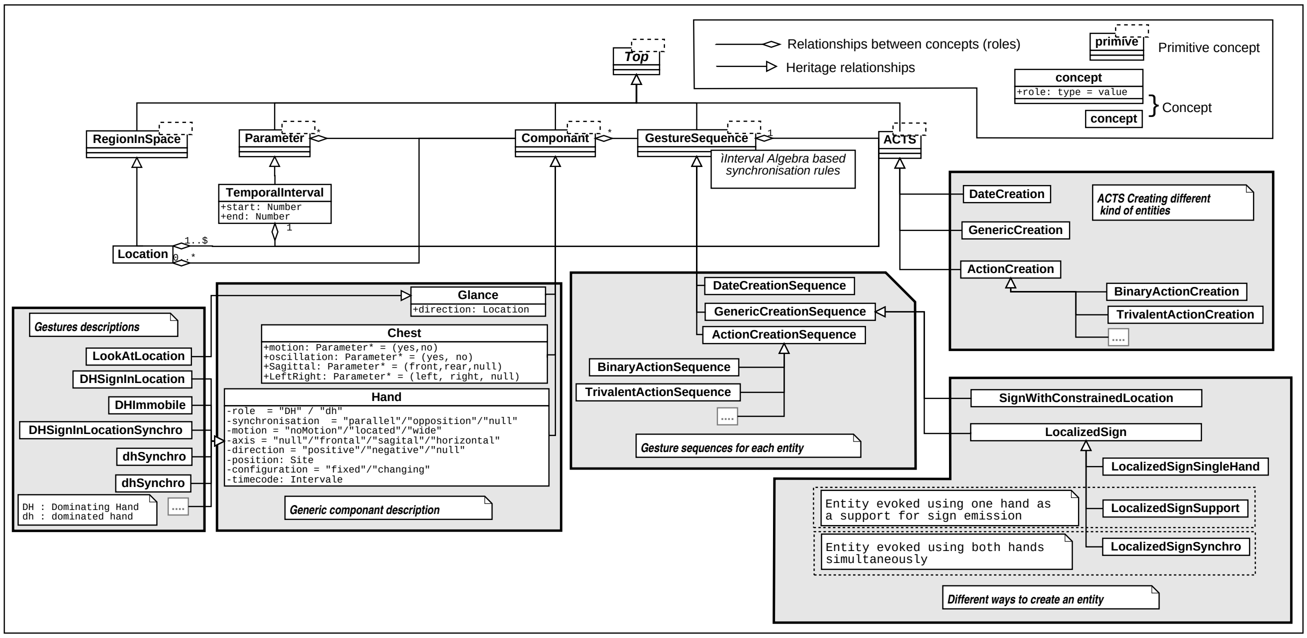

Figure 2: UML class diagram describing the concept hierarchy associated with FSL grammar model in description logic formalism.

knowledge concerning the grammatical structures that can be found in sentences in the specific context of a timetable so that it does not take into account structures such as those produced in complex iconic sentences. However, the model is based on an iconic approach of SL so that the grammatical mechanism it describes can be found in standard and iconic sentences as well.

6. Conclusion

We presented in this paper a preliminary work concerning the formalisation of a global model of the structure of FSL utterances that makes it possible to connect a representation of the meaning of the sentence through the construction of the signing space and the gestures that were used to produce this meaning through a behaviour model of the components of the signer's body. Our goal was to analyse a sentence without taking in account the lexicon so that we could avoid using exhaustive and expensive gesture identification methods. Such a representation can be sufficient in a limited context where the combinatory of possible sentences is reduced. If not, our model can be used to identify the general structure of the sentence and thus help sign recognition.

The model we propose does not take into account the usual dichotomy between standard and iconic sentences so that it is suitable for spontaneous FSL sentence interpretation and can be extended to take in account complex sentences such as those including transfers.

Moreover, even if that model was originally designed for FSL analysis, it appeared that it could be used for FSL sentence generation as it provides a way to describe the spatial structure of the sentence and to connect it to the gestures that will be produced to build that structure. Finally, the model encounters great interest from linguist working on FSL, as that formalisation is a way to evaluate the linguistic model and may be a help for further research on SL grammar.

7. References

Allen, J.F., 1990. Towards a general theory of action and time. In J. Allen, J. Hendler, and A. Tate (eds.), *Readings in Planning*. San Mateo, CA: Kaufmann, pages 464–479.

Brachman, R.J. and al., 1991. Living with classic: When and how to use a kl-one-like language. In J. Sowa (ed.), *Principles of Semantic Networks: Explorations in the representation of knowledge*. San Mateo, California: Morgan-Kaufmann, pages 401–456.

Braffort, A., 1996. Argo : An architecture for sign language recognition and interpretation. In P. Harling and al. (eds.), *Progress in Gestual Interaction*. Springer.

Cuxac, C., 1999. French sign language: proposition of a structural explanation by iconicity. In Springer: Berlin (ed.), *Lecture Notes in Artificial Intelligence : Procs 3rd Gesture Workshop'99 on Gesture and Sign-Language in Human-Computer Interaction*. Gif-sur-Yvette, France: A. Braffort, R. Gherbi , S. Gibet , J. Richardson, D. Teil.

Huenerfauth, M., 2004. Spatial representation of classifier predicates for machine translation into american sign language. In *Workshop on Representation and Processing of Sign Language, 4th Internationnal Conference on Language Ressources and Evaluation (LREC 2004)*. Lisbon Portugal.

Langacker, R., 1987. *Foundations of cognitive grammar: Theoretical Prerequisites*. Stanford, CA: Stanford University Press. Vol 1, 1987 (Hardcover), 1999 (Paperback).

Lebourque, T. and Gibet S., 1999. High level specification and control of communication gestures : the gessyca system. In *Computer Animation'99*. Geneva: Computer Graphics Society (CGS) and the IEEE Computer Society.

Liang, R.H. and M. Ouhyoung, 1998. A real-time continuous gesture recognition system for sign language. In 3^{rd} *International conference on automatic face and gesture recognition*. Nara, Japan.

Prillwitz, S. and al., 1989. *HamNoSys. Version 2.0; Hamburg Notation System for Sign Languages. An introductory guide*. Hamburg : Signum.

Sagawa, H., M. Takeuchi, and M. Ohki, 1997. Description and recognition methods for sign language based on gesture components. In *in Proceedings of IUI*. orlando, Florida.

Stokoe, D., D. Casterline, and C. Croneberg, 1965. *A dictionnary of american sign langage*. Gallaudet College Press.

Thom, R., 1972-1977. *Stabilité structurelle et morphogénèse*. Paris: Édisciences, 2nd edition.

Vogler, C. and D. Metaxas, 2003. Handshapes and movements: Multiple-channel asl recognition. In G Volpe A. Camurri (ed.), *Proceedings of the Gesture Workshop'03, Lecture Notes in Artificial Intelligence*, volume 2915 of *Lecture Notes in Computer Science*. Genova, Italy: Springer.