



Sign language lexical recognition with Propositional Dynamic Logic

Arturo Tracael Curiel Diaz, Christophe Collet

► To cite this version:

Arturo Tracael Curiel Diaz, Christophe Collet. Sign language lexical recognition with Propositional Dynamic Logic. 51st Annual Meeting of the Association for Computational Linguistics (ACL), Aug 2013, Sofia, Bulgaria. 2, pp. 328-333, 2013. <hal-01193094>

HAL Id: hal-01193094

<https://hal.archives-ouvertes.fr/hal-01193094>

Submitted on 4 Sep 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Open Archive TOULOUSE Archive Ouverte (OATAO)

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

This is an author-deposited version published in : <http://oatao.univ-toulouse.fr/>
Eprints ID : 13051

The contribution was presented at ACL 2013 :
<http://acl2013.org/site/>

To cite this version : Curiel Diaz, Arturo Tlacaelel and Collet, Christophe
Sign language lexical recognition with Propositional Dynamic Logic. (2013)
In: 51st Annual Meeting of the Association for Computational Linguistics
(ACL), 4 August 2013 - 9 August 2013 (Sofia, Bulgaria).

Any correspondance concerning this service should be sent to the repository
administrator: staff-oatao@listes-diff.inp-toulouse.fr

Sign Language Lexical Recognition With Propositional Dynamic Logic

Arturo Curiel

Université Paul Sabatier
118 route de Narbonne, IRIT,
31062, Toulouse, France
curiel@irit.fr

Christophe Collet

Université Paul Sabatier
118 route de Narbonne, IRIT,
31062, Toulouse, France
collet@irit.fr

Abstract

This paper explores the use of Propositional Dynamic Logic (PDL) as a suitable formal framework for describing Sign Language (SL), the language of deaf people, in the context of natural language processing. SLs are visual, complete, standalone languages which are just as expressive as oral languages. Signs in SL usually correspond to sequences of highly specific body postures interleaved with movements, which make reference to real world objects, characters or situations. Here we propose a formal representation of SL signs, that will help us with the analysis of automatically-collected hand tracking data from French Sign Language (FSL) video corpora. We further show how such a representation could help us with the design of computer aided SL verification tools, which in turn would bring us closer to the development of an automatic recognition system for these languages.

1 Introduction

Sign languages (SL), the vernaculars of deaf people, are complete, rich, standalone communication systems which have evolved in parallel with oral languages (Valli and Lucas, 2000). However, in contrast to the last ones, research in automatic SL processing has not yet managed to build a complete, formal definition oriented to their automatic recognition (Cuxac and Dalle, 2007). In SL, both hands and non-manual features (NMF), *e.g.* facial muscles, can convey information with their placements, configurations and movements. These particular conditions can difficult the construction of

a formal description with common natural language processing (NLP) methods, since the existing modeling techniques are mostly designed to work with one-channel sound productions inherent to oral languages, rather than with the multi-channel partially-synchronized information induced by SLs.

Our research strives to address the formalization problem by introducing a logical language that lets us represent SL from the lowest level, so as to render the recognition task more approachable. For this, we use an instance of a formal logic, specifically Propositional Dynamic Logic (PDL), as a possible description language for SL signs.

For the rest of this section, we will present a brief introduction to current research efforts in the area. Section 2 presents a general description of our formalism, while section 3 shows how our work can be used when confronted with real world data. Finally, section 4 present our final observations and future work.

Images for the examples where taken from (DictaSign, 2012) corpus.

1.1 Current Sign Language Research

Extensive efforts have been made to achieve efficient automatic capture and representation of the subtle nuances commonly present in sign language discourse (Ong and Ranganath, 2005). Research ranges from the development of hand and body trackers (Dreuw et al., 2009; Gianni and Dalle, 2009), to the design of high level SL representation models (Lejeune, 2004; Lenseigne and Dalle, 2006). Linguistic research in the area has focused on the characterization of corporal expressions into meaningful transcriptions (Dreuw et al., 2010; Stokoe, 2005) or common patterns across SL (Aronoff et al., 2005; Meir et al., 2006; Wittmann, 1991), so as to gain understanding of the un-

derlying mechanisms of SL communication.

Works like (Losson and Vannobel, 1998) deal with the creation of a lexical description oriented to computer-based sign animation. Report (Filhol, 2009) describes a lexical specification to address the same problem. Both propose a thoroughly geometrical parametric encoding of signs, thus leaving behind meaningful information necessary for recognition and introducing data beyond the scope of recognition. This complicates the reutilization of their formal descriptions. Besides, they don't take in account the presence of partial information. Treating partiality is important for us, since it is often the case with automatic tools that incomplete or unrecognizable information arises. Finally, little to no work has been directed towards the unification of raw collected data from SL corpora with higher level descriptions (Dalle, 2006).

2 Propositional Dynamic Logic for SL

Propositional Dynamic Logic (PDL) is a multi-modal logic, first defined by (Fischer and Ladner, 1979). It provides a language for describing programs, their correctness and termination, by allowing them to be modal operators. We work with our own variant of this logic, the *Propositional Dynamic Logic for Sign Language* (PDL_{SL}), which is just an instantiation of PDL where we take signers' movements as programs.

Our sign formalization is based on the approach of (Liddell and Johnson, 1989) and (Filhol, 2008). They describe signs as sequences of immutable *key postures* and movement *transitions*.

In general, each key posture will be characterized by the concurrent parametric state of each *body articulator* over a time-interval. For us, a body articulator is any relevant body part involved in signing. The parameters taken in account can vary from articulator to articulator, but most of the time they comprise their configurations, orientations and their placement within one or more *places of articulation*. Transitions will correspond to the movements executed between fixed postures.

2.1 Syntax

We need to define some primitive sets that will limit the domain of our logical language.

Definition 2.1 (Sign Language primitives). *Let $\mathcal{B}_{SL} = \{\mathbb{D}, \mathbb{W}, \mathbb{R}, \mathbb{L}\}$ be the set of relevant body articulators for SL, where \mathbb{D} , \mathbb{W} , \mathbb{R} and \mathbb{L} represent the dominant, weak, right and left hands, respectively. Both \mathbb{D} and \mathbb{W} can be aliases for the right or left hands, but they change depending on whether the signer is right-handed or left-handed, or even depending on the context.*

Let Ψ be the two-dimensional projection of a human body skeleton, seen by the front. We define the set of places of articulation for SL as $\Lambda_{SL} = \{\text{HEAD}, \text{CHEST}, \text{NEUTRAL}, \dots\}$, such that for each $\lambda \in \Lambda_{SL}$, λ is a sub-plane of Ψ , as shown graphically in figure 1.

Let \mathbb{C}_{SL} be the set of possible morphological configurations for a hand.

Let $\Delta = \{\uparrow, \nearrow, \rightarrow, \searrow, \downarrow, \swarrow, \leftarrow, \nwarrow\}$ be the set of relative directions from the signer's point of view, where each arrow represents one of eight possible two-dimensional direction vectors that share the same origin. For vector $\delta \in \Delta$, we define vector $\overleftarrow{\delta}$ as the same as δ but with the inverted abscissa axis, such that $\overleftarrow{\delta} \in \Delta$. Let vector $\widehat{\delta}$ indicate movement with respect to the dominant or weak hand in the following manner:

$$\widehat{\delta} = \begin{cases} \delta & \text{if } \mathbb{D} \equiv \mathbb{R} \text{ or } \mathbb{W} \equiv \mathbb{L} \\ \overleftarrow{\delta} & \text{if } \mathbb{D} \equiv \mathbb{L} \text{ or } \mathbb{W} \equiv \mathbb{R} \end{cases}$$

Finally, let \vec{v}_1 and \vec{v}_2 be any two vectors with the same origin. We denote the rotation angle between the two as $\theta(\vec{v}_1, \vec{v}_2)$.

Now we define the set of atomic propositions that we will use to characterize fixed states, and a set of atomic actions to describe movements.

Definition 2.2 (Atomic Propositions for SL Body Articulators Φ_{SL}). *The set of atomic propositions for SL articulators (Φ_{SL}) is defined as:*

$$\Phi_{SL} = \{\beta_{1\beta_2}^\delta, \Xi_\lambda^{\beta_1}, \mathcal{T}_{\beta_2}^{\beta_1}, \mathcal{F}_c^{\beta_1}, \angle_{\beta_1}^\delta\}$$

where $\beta_1, \beta_2 \in \mathcal{B}_{SL}$, $\delta \in \Delta$, $\lambda \in \Lambda_{SL}$ and $c \in \mathbb{C}_{SL}$.

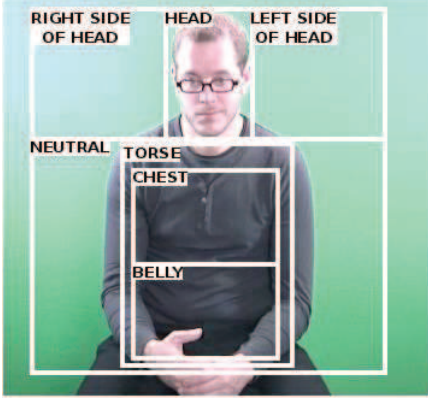


Figure 1: Possible places of articulation in \mathcal{B}_{SL} .

Intuitively, $\beta_1^{\delta}_{\beta_2}$ indicates that articulator β_1 is placed in relative direction δ with respect to articulator β_2 . Let the current place of articulation of β_2 be the origin point of β_2 's Cartesian system (\mathcal{C}_{β_2}). Let vector $\vec{\beta}_1$ describe the current place of articulation of β_1 in \mathcal{C}_{β_2} . Proposition $\beta_1^{\delta}_{\beta_2}$ holds when $\forall \vec{v} \in \Delta$, $\theta(\vec{\beta}_1, \delta) \leq \theta(\vec{\beta}_1, \vec{v})$.

$\Xi_{\lambda}^{\beta_1}$ asserts that articulator β_1 is located in λ .

$\mathcal{T}_{\beta_2}^{\beta_1}$ is active whenever articulator β_1 physically touches articulator β_2 .

$\mathcal{F}_c^{\beta_1}$ indicates that c is the morphological configuration of articulator β_1 .

Finally, $\angle_{\beta_1}^{\delta}$ means that an articulator β_1 is oriented towards direction $\delta \in \Delta$. For hands, $\angle_{\beta_1}^{\delta}$ will hold whenever the vector perpendicular to the plane of the palm has the smallest rotation angle with respect to δ .

Definition 2.3 (Atomic Actions for SL Body Articulators Π_{SL}). The atomic actions for SL articulators (Π_{SL}) are given by the following set:

$$\Pi_{\text{SL}} = \{\delta_{\beta_1}, \rightsquigarrow_{\beta_1}\}$$

where $\delta \in \Delta$ and $\beta_1 \in \mathcal{B}_{\text{SL}}$.

Let β_1 's position before movement be the origin of β_1 's Cartesian system (\mathcal{C}_{β_1}) and $\vec{\beta}_1$ be the position vector of β_1 in \mathcal{C}_{β_1} after moving. Action δ_{β_1} indicates that β_1 moves in relative direction δ in \mathcal{C}_{β_1} if $\forall \vec{v} \in \Delta$, $\theta(\vec{\beta}_1, \delta) \leq \theta(\vec{\beta}_1, \vec{v})$.

Action $\rightsquigarrow_{\beta_1}$ occurs when articulator β_1 moves rapidly and continuously (thrills) with-

out changing its current place of articulation.

Definition 2.4 (Action Language for SL Body Articulators \mathcal{A}_{SL}). The action language for body articulators (\mathcal{A}_{SL}) is given by the following rule:

$$\alpha ::= \pi \mid \alpha \cap \alpha \mid \alpha \cup \alpha \mid \alpha; \alpha \mid \alpha^*$$

where $\pi \in \Pi_{\text{SL}}$.

Intuitively, $\alpha \cap \alpha$ indicates the concurrent execution of two actions, while $\alpha \cup \alpha$ means that at least one of two actions will be non-deterministically executed. Action $\alpha; \alpha$ describes the sequential execution of two actions. Finally, action α^* indicates the reflexive transitive closure of α .

Definition 2.5 (Language PDL_{SL}). The formulae φ of PDL_{SL} are given by the following rule:

$$\varphi ::= \top \mid p \mid \neg \varphi \mid \varphi \wedge \varphi \mid [\alpha] \varphi$$

where $p \in \Phi_{\text{SL}}$, $\alpha \in \mathcal{A}_{\text{SL}}$.

2.2 Semantics

PDL_{SL} formulas are interpreted over labeled transition systems (LTS), in the spirit of the possible worlds model introduced by (Hintikka, 1962). Models correspond to connected graphs representing key postures and transitions: states are determined by the values of their propositions, while edges represent sets of executed movements. Here we present only a small extract of the logic semantics.

Definition 2.6 (Sign Language Utterance Model \mathcal{U}_{SL}). A sign language utterance model (\mathcal{U}_{SL}), is a tuple $\mathcal{U}_{\text{SL}} = (S, R, \llbracket \cdot \rrbracket_{\Pi_{\text{SL}}}, \llbracket \cdot \rrbracket_{\Phi_{\text{SL}}})$ where:

- S is a non-empty set of states
- R is a transition relation $R \subseteq S \times S$ where, $\forall s \in S, \exists s' \in S$ such that $(s, s') \in R$.
- $\llbracket \cdot \rrbracket_{\Pi_{\text{SL}}} : \Pi_{\text{SL}} \rightarrow R$, denotes the function mapping actions to the set of binary relations.
- $\llbracket \cdot \rrbracket_{\Phi_{\text{SL}}} : S \rightarrow 2^{\Phi_{\text{SL}}}$, maps each state to a set of atomic propositions.

We also need to define a structure over sequences of states to model internal dependencies between them, nevertheless we decided to omit the rest of our semantics, alongside satisfaction conditions, for the sake of readability.

3 Use Case: Semi-Automatic Sign Recognition

We now present an example of how we can use our formalism in a semi-automatic sign recognition system. Figure 2 shows a simple module diagram exemplifying information flow in the system’s architecture. We proceed to briefly describe each of our modules and how they work together.

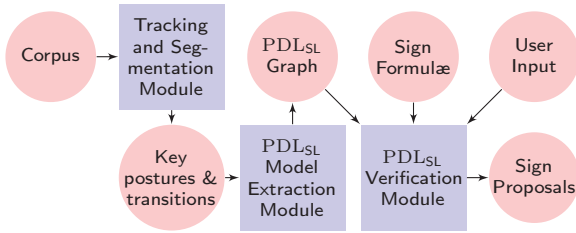


Figure 2: Information flow in a semi-automatic SL lexical recognition system.

3.1 Tracking and Segmentation Module

The process starts by capturing relevant information from video corpora. We use an existing head and hand tracker expressly developed for SL research (Gonzalez and Collet, 2011). This tool analyses individual video instances, and returns the frame-by-frame positions of the tracked articulators. By using this information, the module can immediately calculate speeds and directions on the fly for each hand.

The module further employs the method proposed by the authors in (Gonzalez and Collet, 2012) to achieve sub-lexical segmentation from the previously calculated data. Like them, we use the relative velocity between hands to identify when hands either move at the same time, independently or don’t move at all. With these, we can produce a set of possible key postures and transitions that will serve as input to the modeling module.

3.2 Model Extraction Module

This module calculates a propositional state for each static posture, where atomic PDL_{SL}

formulas codify the information tracked in the previous part. Detected movements are interpreted as PDL_{SL} actions between states.

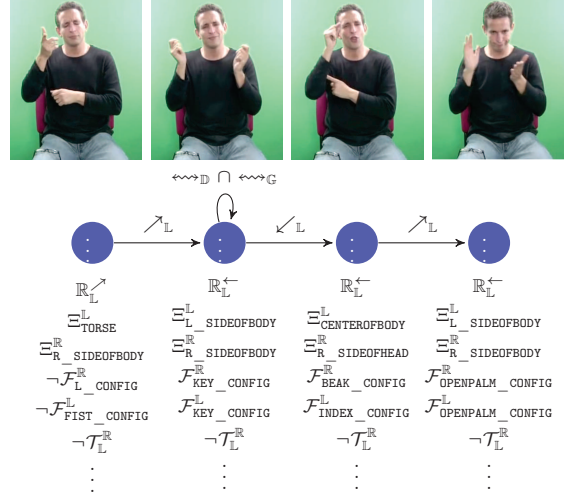


Figure 3: Example of modeling over four automatically identified frames as possible key postures.

Figure 3 shows an example of the process. Here, each key posture is codified into propositions acknowledging the hand positions with respect to each other ($\mathbb{R}_L^{\leftarrow}$), their place of articulation (e.g. “left hand floats over the torse” with $\Xi_{L_TORSE}^{\nearrow}$), their configuration (e.g. “right hand is open” with $\mathcal{F}_{OPENPALM_CONFIG}^{\leftarrow}$) and their movements (e.g. “left hand moves to the up-left direction” with \nearrow_L).

This module also checks that the generated graph is correct: it will discard simple tracking errors to ensure that the resulting LTS will remain consistent.

3.3 Verification Module

First of all, the verification module has to be loaded with a database of sign descriptions encoded as PDL_{SL} formulas. These will characterize the specific sequence of key postures that morphologically describe a sign. For example, let’s take the case for sign “route” in FSL, shown in figure 4, with the following PDL_{SL} formulation,

Example 3.1 (ROUTE_{FSL} formula).

$$(\Xi_{FACE}^{\mathbb{R}} \wedge \Xi_{FACE}^{\mathbb{L}} \wedge \mathbb{L}_R^{\rightarrow} \wedge \mathcal{F}_{CLAMP}^{\mathbb{R}} \wedge \mathcal{F}_{CLAMP}^{\mathbb{L}} \wedge \mathcal{T}_L^{\mathbb{R}}) \rightarrow [\leftarrow_{\mathbb{R}} \cap \rightarrow_{\mathbb{L}}](\mathbb{L}_R^{\rightarrow} \wedge \mathcal{F}_{CLAMP}^{\mathbb{R}} \wedge \mathcal{F}_{CLAMP}^{\mathbb{L}} \wedge \neg \mathcal{T}_L^{\mathbb{R}}) \quad (1)$$

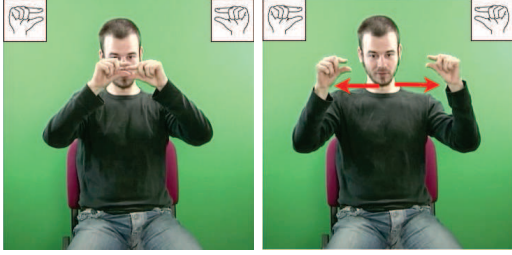


Figure 4: $\text{ROUTE}_{\text{FSL}}$ production.

Formula (1) describes $\text{ROUTE}_{\text{FSL}}$ as a sign with two key postures, connected by a two-hand simultaneous movement (represented with operator \cap). It also indicates the position of each hand, their orientation, whether they touch and their respective configurations (in this example, both hold the same CLAMP configuration).

The module can then verify whether a sign formula in the lexical database holds in any sub-sequence of states of the graph generated in the previous step. Algorithm 1 sums up the process.

Algorithm 1 PDL_{SL} Verification Algorithm

Require: SL model \mathcal{M}_{SL}

Require: connected graph \mathcal{G}_{SL}

Require: lexical database \mathcal{DB}_{SL}

- 1: $\text{Proposals_For}[\text{state_qty}]$
 - 2: **for** state $s \in \mathcal{G}_{\text{SL}}$ **do**
 - 3: **for** sign $\varphi \in \mathcal{DB}_{\text{SL}}$ where $s \in \varphi$ **do**
 - 4: **if** $\mathcal{M}_{\text{SL}}, s \models \varphi$ **then**
 - 5: $\text{Proposals_For}[s].\text{append}(\varphi)$
 - 6: **end if**
 - 7: **end for**
 - 8: **end for**
 - 9: **return** Proposals_For
-

For each state, the algorithm returns a set of possible signs. Expert users (or higher level algorithms) can further refine the process by introducing additional information previously missed by the tracker.

4 Conclusions and Future Work

We have shown how a logical language can be used to model SL signs for semi-automatic recognition, albeit with some restrictions. The traits we have chosen to represent were imposed by the limits of the tracking tools we had to our disposition, most notably working

with 2D coordinates. With these in mind, we tried to design something flexible that could be easily adapted by computer scientists and linguists alike. Our primitive sets, were intentionally defined in a very general fashion due to the same reason: all of the perceived directions, articulators and places of articulation can easily change their domains, depending on the SL we are modeling or the technological constraints we have to deal with. Propositions can also be changed, or even induced, by existing written sign representation languages such as Zebedee (Filhol, 2008) or HamNoSys (Hanke, 2004), mainly for the sake of extendability.

From the application side, we still need to create an extensive sign database codified in PDL_{SL} and try recognition on other corpora, with different tracking information. For verification and model extraction, further optimizations are expected, including the handling of data inconsistencies and repairing broken queries when verifying the graph.

Regarding our theoretical issues, future work will be centered in improving our language to better comply with SL research. This includes adding new features, like incorporating probability representation to improve recognition. We also expect to finish the definition of our formal semantics, as well as proving correction and complexity of our algorithms.

References

- Mark Aronoff, Irit Meir, and Wendy Sandler. 2005. The paradox of sign language morphology. *Language*, 81(2):301.
- Christian Cuxac and Patrice Dalle. 2007. *Problématique des chercheurs en traitement automatique des langues des signes*, volume 48 of *Traitement Automatique des Langues*. Lavoisier, <http://www.editions-hermes.fr/>, October.
- Patrice Dalle. 2006. High level models for sign language analysis by a vision system. In *Workshop on the Representation and Processing of Sign Language: Lexicographic Matters and Didactic Scenarios (LREC)*, Italy, ELDA, page 17–20.
- DictaSign. 2012. <http://www.dictasign.eu>.
- Philippe Dreuw, Daniel Stein, and Hermann Ney. 2009. Enhancing a sign language translation system with vision-based features. In Miguel Sales Dias, Sylvie Gibet, Marcelo M.

- Wanderley, and Rafael Bastos, editors, *Gesture-Based Human-Computer Interaction and Simulation*, number 5085 in Lecture Notes in Computer Science, pages 108–113. Springer Berlin Heidelberg, January.
- Philippe Dreuw, Hermann Ney, Gregorio Martinez, Onno Crasborn, Justus Piater, Jose Miguel Moya, and Mark Wheatley. 2010. The Sign-Speak project - bridging the gap between signers and speakers. In Nicoletta Calzolari (Conference Chair), Khalid Choukri, and *et. al.*, editors, *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, Valletta, Malta, May. European Language Resources Association (ELRA).
- Michael Filhol. 2008. *Modèle descriptif des signes pour un traitement automatique des langues des signes*. Ph.D. thesis, Université Paris-sud (Paris 11).
- Michael Filhol. 2009. Zebedee: a lexical description model for sign language synthesis. Internal, LIMSI.
- Michael J. Fischer and Richard E. Ladner. 1979. Propositional dynamic logic of regular programs. *Journal of Computer and System Sciences*, 18(2):194–211, April.
- Frédéric Gianni and Patrice Dalle. 2009. Robust tracking for processing of videos of communication's gestures. *Gesture-Based Human-Computer Interaction and Simulation*, page 93–101.
- Matilde Gonzalez and Christophe Collet. 2011. Robust body parts tracking using particle filter and dynamic template. In *2011 18th IEEE International Conference on Image Processing (ICIP)*, pages 529–532, September.
- Matilde Gonzalez and Christophe Collet. 2012. Sign segmentation using dynamics and hand configuration for semi-automatic annotation of sign language corpora. In Eleni Efthimiou, Georgios Kouroupetroglou, and Stavroula-Evita Fotinea, editors, *Gesture and Sign Language in Human-Computer Interaction and Embodied Communication*, number 7206 in Lecture Notes in Computer Science, pages 204–215. Springer Berlin Heidelberg, January.
- Thomas Hanke. 2004. HamNoSys—Representing sign language data in language resources and language processing contexts. In *Proceedings of the Workshop on the Representation and Processing of Sign Languages “From Sign Writing to Image Processing. Information*, Lisbon, Portugal, 30 May.
- Jaakko Hintikka. 1962. *Knowledge and Belief*. Ithaca, N.Y., Cornell University Press.
- Fanch Lejeune. 2004. *Analyse sémantico-cognitive d'énoncés en Langue des Signes Française pour une génération automatique de séquences gestuelles*. Ph.D. thesis, PhD thesis, Orsay University, France.
- Boris Lenseigne and Patrice Dalle. 2006. Using signing space as a representation for sign language processing. In Sylvie Gibet, Nicolas Courty, and Jean-François Kamp, editors, *Gesture in Human-Computer Interaction and Simulation*, number 3881 in Lecture Notes in Computer Science, pages 25–36. Springer Berlin Heidelberg, January.
- S. K. Liddell and R. E. Johnson. 1989. *American sign language: The phonological base*. Gallaudet University Press, Washington. DC.
- Olivier Losson and Jean-Marc Vannobel. 1998. Sign language formal description and synthesis. *INT.JOURNAL OF VIRTUAL REALITY*, 3:27–34.
- Irit Meir, Carol Padden, Mark Aronoff, and Wendy Sandler. 2006. Re-thinking sign language verb classes: the body as subject. In *Sign Languages: Spinning and Unraveling the Past, Present and Future. 9th Theoretical Issues in Sign Language Research Conference, Florianopolis, Brazil*, volume 382.
- Sylvie C. W. Ong and Surendra Ranganath. 2005. Automatic sign language analysis: a survey and the future beyond lexical meaning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(6):873–891, June.
- William C. Stokoe. 2005. Sign language structure: An outline of the visual communication systems of the american deaf. *Journal of Deaf Studies and Deaf Education*, 10(1):3–37, January.
- Clayton Valli and Ceil Lucas. 2000. *Linguistics of American Sign Language Text, 3rd Edition: An Introduction*. Gallaudet University Press.
- Henri Wittmann. 1991. Classification linguistique des langues signées non vocalement. *Revue québécoise de linguistique théorique et appliquée*, 10(1):88.