



**HAL**  
open science

## **SL teacher : a framework for teaching LSF on the web**

Jean-Paul Sansonnet, Annelies Braffort, Jérémie Segouat, Cyril Verrecchia

### ► **To cite this version:**

Jean-Paul Sansonnet, Annelies Braffort, Jérémie Segouat, Cyril Verrecchia. SL teacher : a framework for teaching LSF on the web. International Conference on ICT and Accessibility, Jan 2009, Hammamet, Tunisia. hal-01634123

**HAL Id: hal-01634123**

**<https://hal.science/hal-01634123>**

Submitted on 20 Dec 2023

**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.

# SL Teacher: a Framework for teaching LSF on the Web

Jean-Paul Sansonnet, Annelies Braffort, Jérémie Segouat, Cyril Verrecchia  
LIMSI-CNRS BP 133 F-91403 Orsay cedex  
*firstname.lastname@limsi.fr*

## ***Abstract***

*In this paper, we present SL Teacher, a framework for teaching French sign language using web 2.0 and virtual signer technologies. We present here the main features of the web-based framework that supports the sign lexicon animation and the dynamic sign language utterances generation. After describing the technical issues of our proposed framework, we discuss the tradeoffs and the ensuing simplifications applied in order to insure a good level of interactivity with the users in the low bandwidth context of web systems. We present the symbolic representation of the sign language utterances and the associated construction procedure. Finally the implementation, accessible on the web, is illustrated here with an example of a simple teaching task that requires all the features of the SL Teacher framework.*

## **1. Introduction**

### **1.1 Objective**

This paper presents the first step of a project dedicated to French Sign Language (LSF) learning supported by means of a Web environment. E-learning tools dedicated to sign language (SL) can be useful both for deaf and hearing people: as a first language to the deaf or as a second language to the hearing. As for any other natural language, learning the lexicon is not sufficient, and exercises, e.g. about utterance organization, are of primary importance. For SL, that entails 1) visual-based applications are required, due to the gestural/visual channel used by these languages, and 2) strong interactive capabilities are also required as in any e-learning tool.

New technologies such as Web 2.0, virtual signer, language processing and automatic generation allow now envisioning such applications. Hence the two main complementary features of the SL Teacher project are: 1) a web-based framework for teaching SL 2) by means of a virtual signer.

In section 1.2 we present several kinds of existing e-learning applications and technologies. Section 2 presents the technical issues of our proposed framework and section 3 describes the methodology developed for the dynamic construction and the animation of SL-utterances. Finally, section 4 gives an example of typical SL exercises that can be developed within this framework.

### **1.2 E-learning applications dedicated to SL**

Today, SL tools have been deployed in two main environments: *Web pages* and *standalone applications*.

Most operational web-based applications dedicated to SL focus on bilingual dictionaries. They are based on the display of *video recordings* of human subjects. Up now, little work has been carried on web-based virtual signers.

*Virtual signer applications* have been deployed in both Web pages and standalone environments. They rely on two main approaches: pre-synthesized animations and generated animations.

- *Pre-synthesized animations* are first synthesized, and then selected according to the context. For example, motion capture can be applied on realistic or non realistic virtual signers. The one in figure 1, is used in Mathsigner, an interactive learning tool to improve the mathematical abilities of deaf children [1]. These animations are not accurate (many ‘cleaning’ post-processes have to be performed on the signal) and do not take into account significant features of SL: e.g. gaze, facial expressions, and torso movements. Moreover, as the animations rely on a numerical signal, it is hard to manipulate them, as opposed to language generated animation methods.

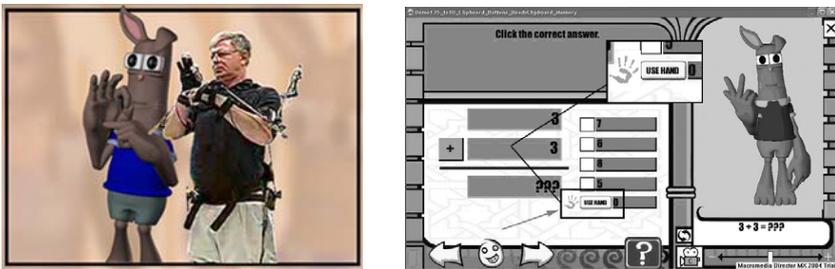


Figure 1: Mathsigner: a) Motion capture phase b) Application interface.

- *Generated animations* are an alternative approach consisting in automatic and real-time generation of animations from *description languages*. For example, the SigML description language [2], based on the HamNoSys phonetic system dedicated to sign languages [3], is used as input in a given signing animation software. It allows the generation of animations that provide flexibility, but at the cost of realism. Java applets or VRML-based applications allow interaction within a web page, such as the Signing Science Dictionary [4] (see figure 2).



Figure 2: The Signing Science Dictionary.

Regardless the chosen approach for animation, the main drawback of all these applications is that the virtual signer is encased into a ‘box’, separated from the rest of the application,

and its interaction with the content is minimalistic. Actually, it is limited to the control of the display of the virtual signer (moving, rotating, scale, speed control etc.). In these applications, the signing agents have no access to the informational content of the web page.

The architecture of any SL Teacher application entails two complementary main issues:

1) How can we propose an architecture that allows realistic virtual signers to be fully and effectively integrated onto web pages? So that people can have an easy internet-based access to SL tools.

2) How can we propose an architecture that allows the dynamic building and the subsequent animation of SL utterances? So that people can practice SL-utterance construction and can also observe (at leisure) their animation in lively exercises.

The two following sections present the way we have dealt with these issues.

## **2. An architecture for fully integration of virtual signer on web pages**

### **2.1 Supporting framework: The DIVA toolkit**

DIVA stands for DOM Integrated Virtual Agents, emphasizing the unique feature of DIVA virtual agents that are completely integrated with the DOM (Document Object Model) tree structure of web pages. It has been designed at LIMSI laboratory [5] with the three main objectives:

- A toolkit dedicated to Assisting Conversational Agents, a subclass of Embodied Conversational Agents (ECAs) [6] whose function is to provide assistance for novice users in front of unfamiliar applications. Note: in this particular context, textual natural language interaction plays a primary role; this issue will not be discussed here even if it is relevant when multimodal deictic assisting agents are concerned;
- An open programming framework, developed for teaching and research purposes, making it easy and quick to develop and deploy new experimental ECAs in web-based applications;
- A completely DOM-integrated software architecture making it easy for the DIVA ECAs to access, both in read/modify modes, to the inner structure of the applications/services supported by the web pages.

The web architecture of DIVA is displayed in figure 3a. It is composed of two main layers:

- A server layer dedicated to data base resources management and symbolic computing;
- A rich-client layer supporting: The specific application/service web page; the animation of the graphic characters; the process of the textual natural language interaction.

The DIVA toolkit is a full Web 2.0 rich client technology based on JavaScript, Java Server Pages and Mathematica from Wolfram Research. It was intended to work with IE 6+ and Mozilla-based navigators (Firefox, Safari and Chrome). DIVA can be deployed in two main modes: 1) the rich-client mode enables web-apps to work locally, i.e. without server capabilities; 2) the symbolic-server mode offers more powerful web architecture but it requires an additional Mathematica engine.

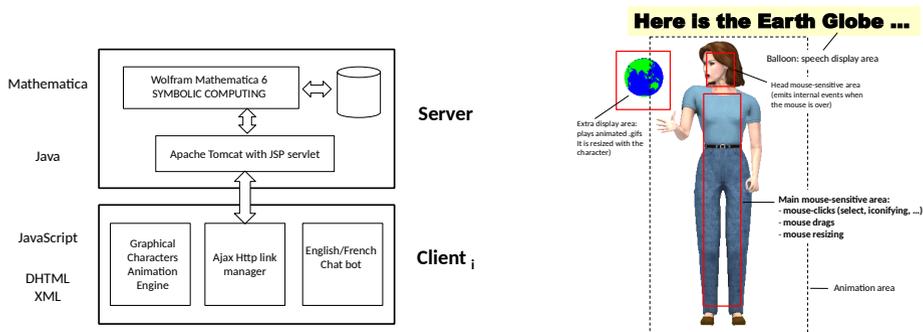


Figure 3: a) Web architecture of DIVA. b) The DOM structure of a signing avatar.

## 2.2 DOM-integrated virtual characters

The first objective of the DIVA toolkit is to develop a fully DOM-integrated environment for animated graphic characters in order to provide a *high degree of interaction* with the DOM-structure and the informational content of the webpage. Therefore, DIVA characters are part of the DOM structure that contains several DOM objects of type `<div>` supporting various capabilities: display, drag&drop, resizing, iconifying, animation, speech display, etc. (see figure 3b.).

Due to the unique feature of DIVA agents where everything is supported by `<div>` objects in the DOM of the client page, users can interact with the agents while using multimodal inputs: by typing utterances in the chat box but also with the mouse acting upon the graphic display in various ways. Another advantage of the full integration of the DIVA agents in the DOM structure is that the rich-client application can have a direct and complete access to the structure and the runtime of the application.

The animation engine also takes advantage of this integration: It is a JavaScript program that directly loads .png files in `<div>` objects and then animates them, at the frame level. This makes it possible to compute animations movies for *realistic characters* and to compose them in JavaScript locally, on the client page. However, good quality realistic characters require fine grained pictures that can slacken the loading phase of the client pages. In a web-based environment we have to deal with such tradeoffs such as the quality of the graphic display against the loading speed. We had to consider both the pixel size of the frames and the number of frames for one animation.

Because we need to animate a virtual signer in a web environment, we have a technological constraint to take into account: Internet bandwidth constraints can make the loading phase of a SL Teacher page quite slow if the realistic avatar picture files are too large. Presently, the animations items are composed of 6, 15 or 30 frames in 500x500 pixels in .png format (.png provides the required transparency for the avatars to move freely over the web page). As for the animation of the characters on screen, although it is supported by a rich-client engine written in JavaScript, the animations are quite fluid, especially with Mozilla based navigators and this is bound to improve with the new generation of navigators integrating JavaScript accelerators (like Google Chrome).

The contributions of this approach are related to the interaction capabilities and the easiness to design new applications. The drawbacks are related to the animation engine. For the time being, only pre-synthesized animations approach can be used. The key issues related to the SL utterance generation and animation, in this context, are discussed in the following section.

### **3. Dynamic building and animation of SL utterances**

#### **3.1 Key issues and simplifications**

For the linguistic side, several complex issues have also to be taken into account when dealing with SL utterances generation. SL is a *visual* and *gestural* language that uses manual and non-manual features. A SL utterance is composed of a sequence of gestural units (signs) involving the upper body, the arms and hands, the face and gaze. All these elements are articulated in the space located in front of the signer, so-called signing space. Moreover, signs have variable realisations, highly depending on the meaning context [7]. Thus, SL utterance synthesis implies to design specific spatio-temporal grammar rules, quite different from the ones in use for vocal language generation [8]. Moreover, the coarticulation process influences both manual and non-manual features when a sign is performed in an utterance. This phenomenon has been studied for few years in SL synthesis research field [9], but there are several studies in SL recognition [10] etc. It is a complex process that has not yet been modelled accurately so far.

For all these reasons, we have decided to manage the generation of the SL-utterances as follow:

- A SL utterance is built by concatenation of signs.
- Each sign is displayed as a predefined animation, built using the rotoscopy technique, allowing a good realistic rendering (see [11] for more details).
- For all the signs the realisation of which varies regarding to the context (e.g. pointing), we predefine several realisations (e.g. six for pointing, three to the left and three to the right); note that their combinatory is limited because we restrict to context-dependent utterances.
- For the coarticulation sign animations, we use one or more intermediary postures (chosen dynamically, according to the context): typically, hands are in front of the chest for signs, and the dominant hand is in front of the shoulder for dactylogy.
- Lastly, in order to build animations as realistic as possible, each utterance contains prologue and epilogue stances, allowing the virtual signer to begin and end the utterance in a rest posture (denoted ‘Stand-still’ where the avatar fidgets just a little to suggest life).

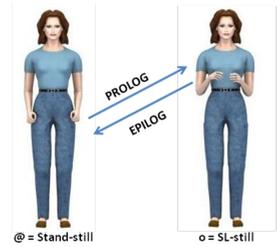
#### **3.2 Constitution of SL-utterances**

SL Teacher utterances are based on four kinds of entities:

1) Two coarticulation postures denoted:

- @ = the Stand-still posture (arms down near the waist);
- o = the SL-still posture (hands in front of the navel).

- 2) Two animations associated with the postures:  
 PROLOG moves the arms from the Stand-still posture @ to the SL-still posture o;  
 EPILOG moves the arms from the SL-still posture o to the Stand-still posture @;



- 3) Constant SL signs (taken from the LSF lexicon), for example:  
 HOME denotes the animated movie associated with “home” in LSF;  
 DOG denotes the animated movie associated with “dog” in LSF;  
 EAT denotes the animated movie associated with “to eat” in LSF;  
 Etc.
- 4) The `_variable_` signs have a *dynamically* computed value which is a constant SL sign.  
 Variable signs are used when their value depends on the context. For example:  
`_POINTING_` is a variable deictic gesture towards a clicked object: its value depends on the relative spatial position of the agent and the object so it has to be chosen between six different directions, entailing six different realisations;  
`_SIGNNAME_` is a variable that contains the “sign name” animation associated with a specific person known in the application;  
`_GOTO_` is a directional verb which beginning and ending spatial location depends on the beginning and ending towns, country etc. world locations in the signing space.

Example:

Suppose we want the avatar to point with its finger on the picture of a specific person displayed on the screen and to give his/her associated sign-name in an utterance like “His sign-name is ‘chirping-bird’ (first author sign-name)”. We dynamically build the following symbolic representation of the SL-utterance S which is composed of six animations containing three static signs and three variable signs:

$$S = [ @ \text{ PROLOG } o \text{ _POINTING_ } o \text{ SIGNNAME } o \text{ _OFHIS_ } o \text{ _SIGNNAME_ } o \text{ EPILOG } @ ]$$

In reaction to some GUI event (like the user clicking on the picture of a person on screen), the variables `_POINTING_`, `_OFHIS_` and `_SIGNNAME_` are *instantiated in context*. Then the animation sequence S is built and played immediately.

#### 4 Example: LSF teaching for hearing beginners

While designing an e-learning application, the main issue relates to pedagogical aspects which are not yet well ascertained. For example, in France, LSF has only been officially recognised in 2005. The pedagogical references are just being elaborated by working groups of the French Ministry of Education and are correlated to the European reference for the learning of language in general. Hence, pedagogical issues are not discussed in this paper but will be subject to further studies.

In order to put our framework to the test, we are developing several demos and experiments that can be accessed at the LIMSI Gestural Agents webpage [12]. In this

section, we will restrict to the description of one typical experiment dedicated to the support of LSF teaching for hearing people which are beginners in LSF. The proposed page could be part of a future web-based course.

In the web page displayed in figure 4, the pedagogical intention is to make the student familiar with the specific ordering of the syntagms in a SL-utterance, namely the so-called “WHEN-WHERE-WHO-WHAT(ACTION)” typical order [13] — without forgetting that LSF being a natural language it cannot be restricted to strict formal rules; because of the naturalness and the iconicity of the language, there are other manners to perform those written sentences.

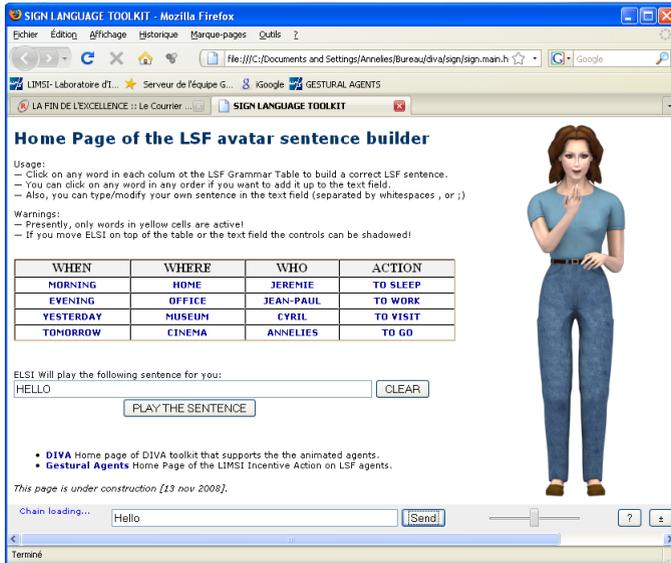


Figure 4: SL Teacher web-page with ELSI virtual signer playing a SL-utterance. Note that here the number of sign-words is limited for the experiment.

For example, the utterance: “*Jeremie work at home this morning*” is (preferably) expressed in LSF as: < MORNING – HOME – JEREMIE – TOWORK >. The table on screen enables the students to compose various SL-utterances by picking sign-words in the table (column after column), thus respecting a typical LSF ordering. Notice that we don't deal with French to LSF translation as we follow the sign order given by the user through the interface.

Then the students can make the character ELSI to play their dynamically composed SL-utterances, *at leisure*, that is they can repeat, slow down, accelerate, etc., the animation.

## 5. Conclusion

This paper presents SLTeacher, a framework for the teaching of the French Sign Language using web 2.0 and virtual signer technologies. We discussed the key issues and the tradeoffs that were related to the constraints in a web environment: the SL utterances are

composed of concatenated pre-defined animations of signs which realisations are chosen regarding the context. Intermediate stances are used as replacement of real coarticulation.

At this moment, the implementation on the DIVA framework of a first SL teaching page with a typical exercise has been done, and our first prototype can be accessed at the DIVA SL Teacher Web page [14]. In a preliminary qualitative evaluation by a deaf person (fourth author), animations are considered fluid, and intermediate postures do not disturb the clear understanding of the SL-utterances. Real evaluation by deaf users is planned during this year. This first actual experiment will allow us to start an evaluation process, in order to validate, adapt or correct some of the choices we made, such as those of the intermediate stances; the first step being to define an adequate evaluation methodology.

In a further step, we will have to extend SLTeacher capabilities, in collaboration with deaf SL teachers, and include SL processing in order to provide for example automatic corrections.

## 6. References

- [1] Adamo-Villani N., B. Benes, M. Brisbin, [A natural interface for sign language mathematics](#), ISVC 2006, Lake Tahoe, 1, 70-79. Berlin, Germany: Springer-Verlag LNCS, 2006
- [2] Elliott, R. and Glauert, J.R.W. and Jennings, V. and Kennaway, J.R., "An Overview of the SiGML Notation and SiGML Signing Software System", *In Fourth International Conference on Language Resources and Evaluation, LREC 2004*, Eds. Streiter, O. and Vettori, C., Lisbon, Portugal, pp. 98-104, 2004
- [3] Prillwitz S, Leven R, Zienert H, Hanke T, Henning J, et al. (1989) Hamburg notation system for sign languages: An introductory guide. International Studies on Sign Language and the Communication of the Deaf, vol 5. Institute of German Sign Language and Communication of the Deaf, University of Hamburg.
- [4] Signing Science URL, <http://signsci.terc.edu/>
- [5] DIVA toolkit homepage, <http://www.limsi.fr/Individu/jps/online/diva/divahome/index.html>
- [6] Cassell, J., Sullivan, J., Prevost, S., and Churchill, E., eds. Embodied Conversational Agents. MIT Press, 2000.
- [7] Cuxac, C. The Expression of Spatial Relations and the Spatialization of Semantic Relations in French Sign Language. Language Diversity and Cognitive Representations. Benjamins : Amsterdam, 123-142, 1999.
- [8] Braffort A., Dalle P., "Sign language applications: preliminary modeling". In: International journal "Universal Access in the Information Society" (UAIS), Special issue 6/4 "Emerging Technologies for Deaf Accessibility in the Information Society", Springer 2007.
- [9] Segouat J., A Study of Sign Language Coarticulation, 10th International ACM SIGACCESS Conference on Computers and Accessibility, ASSETS-2008, Halifax, CA. 2008.
- [10] Vogler, C., Metaxas, D. N. Toward Scalability in ASL Recognition: Breaking Down Signs into Phonemes GW '99: Proceedings of the International Gesture Workshop on Gesture-Based Communication in Human-Computer Interaction, Springer-Verlag, 1999, 211-224
- [11] Sansonnet J.-P., Braffort A., Martin J.-C., Verrecchia C., [Towards interactive web-based signing avatars](#), The 8th International Gesture Workshop, GW 2009, Bielefeld Germany Febr 2009
- [12] Gestural Agents URL, <http://www.limsi.fr/~jps/online/diva/geste/geste.main.html>
- [13] Companys M., LSF mode d'emploi. L'expression par la pensée visuelle, Monica Companys Ed., 2003.
- [14] SLTeacher homepage: <http://www.limsi.fr/~jps/online/diva/sign/sign.main.htm>