

# DEGELS1: A comparable corpus of French Sign Language and co-speech gestures

Annelies Braffort<sup>1</sup>, Leïla Boutora<sup>2</sup>

<sup>1</sup> LIMSI/CNRS

<sup>1</sup> Campus d'Orsay bat 508, BP 133, F-91403 Orsay cx, France

<sup>2</sup> LPL & Aix-Marseille University

<sup>2</sup> 5 avenue Pasteur, F-13100 Aix-en-Provence, France

annelies.braffort@limsi.fr, leila.boutora@lpl-aix.fr

## Abstract

In this paper, we describe DEGELS1, a comparable corpus of French Sign Language and co-speech gestures that has been created to serve as a testbed corpus for the DEGELS workshops. These workshop series were initiated in France for researchers studying French Sign Language and co-speech gestures in French, with the aim of comparing methodologies for corpus annotation. An extract was used for the first event DEGELS2011 dedicated to the annotation of pointing, and the same extract will be used for DEGELS2012, dedicated to segmentation.

**Keywords:** Corpus annotation, sign language, gestures, multimodality

## 1. Introduction

Few studies have related the gestural organisation of signed languages (SL) with that of spoken languages (Zwitserlood, Özyürek & Perniss, 2003). These areas are less-resourced, lacking in standard for methodologies and tools. The co-speech gesture and sign language communities share a number of research issues related to corpora annotation, and should benefit to share their resources and expertise. Sharing corpora and annotations requires good knowledge of methods and a clear description of the annotation schemes and the criteria used (segmentation, selection of categories, etc.), using guidelines, such as that developed for sign languages (Nonhebel, Crasborn, & Van der Kooij, 2004; Johnston, 2011) and for multimodality (Kita, Van Gijn & Van der Hulst, 1997).

International conferences such as LREC and associated workshops related to multimodality or sign languages offer a place to exchange and share on this topic. We want to offer such an opportunity in France for researchers studying co-speech gestures in French and French Sign Language (LSF). With this aim, we have initiated a series of workshops, organised the first event in 2011, and created a specific corpus to be used for these events called DEGELS1.

In this paper, we describe DEGELS1, a comparable corpus of French Sign Language and co-speech gestures that has been created to serve as a testbed corpus for the DEGELS workshops.

The DEGELS workshop series is presented in section 2. Section 3 described the DEGELS1 corpus created for these workshops, and section 4 provides a report on the first edition of the workshop, DEGELS2011.

## 2. DEGELS workshop series

The objective of the DEGELS (DEfi Geste Langue des Signes) workshop series is to bring together the scientific communities studying co-speech gestures and sign languages in France, around common issues concerning corpus annotation.

This workshop takes the form of a challenge of annotation: we provide video data extracted from a comparable corpus of Sign Language and spoken French including voice and gestures to the participants. The participants must annotate them in a limited time; provide their annotation to the organisers; finally write a paper describing their methodology (theoretical choices, annotation scheme, selection criteria ...). The organisers compare annotations with each other, on a different topic each year, and draft a synthesis of the different approaches. This synthesis and the different annotations are presented at the workshop day with an assessment of the strengths of each approach. This day is an opportunity to discuss the methods and annotation schemes.

For this first edition, DEGELS2011, we have proposed to annotate pointing (Boutora, Braffort & Bertrand, 2011).. Pointing is present in both gestures and sign languages. This phenomenon is described in the literature and is subject to relatively large annotations and analysis compared to other themes. Thus, this was an ideal candidate for this first edition.

## 3. DEGELS1 corpus

DEGELS1 is a video corpus of dialogues in two languages: spoken French and French Sign Language. The theme of the dialog is the same for the two languages: a proposal for places to see in Marseille (the second town of France, in the south east).

We filmed six sessions: three dyads of hearing people and three dyads of deaf signing people. Each dyad consisted of a moderator and a speaker not involved in the project. The moderator was the same for each dyad of a given language: a hearing one for French and a deaf one for LSF. The two moderators had prepared the elicitation protocol to ensure two comparable corpora in French and LSF. The participants were seated opposite to each other.

The recording was performed in the LPL's anechoic room, with two external microphones (for French speech) connected to two cameras providing a front view of the speakers/signers, and a third camera for a global view. For each session, we built a composited view, as shown in Figure 1. All the files are synchronised on the basis of a clap at the start of the recording. Each session in French contains four video files and two speech files. Each session in LSF contains four video files and one text file that provide a translation in French of the LSF production.



Figure 1: Excerpt from DEGELS1 corpus for French (top) and LSF (down).

DEGELS1 corpus has been registered in the Speech & Language Data Repository (SLDR) managed by the LPL, which provides services for sharing linguistic data (primary data, annotations, tools, etc.) and archiving it. The corpus is referenced in international repositories such

as OLAC and VLO with the identifier [oai:sldr.org/sldr000767](http://oai.sldr.org/sldr000767). It is in free access for researchers, under a non-commercial license.

We provided participants to DEGELS2011 workshop with extracts of DEGELS1 of one minute for each language. These extracts were selected in order to include a large number of pointing, while remaining short in duration, given the limited time available to the participants to annotate the corpus. For both languages, the extracts relate to the same portion of route: the road that follows the sea from the Vieux-Port to Marseille's beaches by the coast.

#### 4. Comparison of the DEGELS1 annotations in DEGELS2011

Seven teams have participated to the DEGELS2011 workshop dedicated to pointing on the proposed DEGELS1 extract: three have worked on the French extract, three on the LSF one, and one team has worked on the two languages. That represents 21 participants, mainly linguistic researchers.

We have collected all the annotations and tried to find a way to compare them. Most of the time, the annotations were different on various aspects:

- annotated modalities and articulators: only hand gesture, whole body, etc.; only gestures vs. gestures and speech;
- annotated speakers: only the speaker, or also the moderator;
- annotation software: Anvil, Elan, and other software;
- use of automatic processing or not;
- time spent: from 4 to 90 hours;
- number of annotators: from 1 to 4;
- nature of the categories: descriptive (form) or analytic (linguistic function).

When possible, alignments of the annotations were performed and integrated in annotation software in order the participants to discuss together of the differences during the workshop. This alignment was performed using the export facilities offered by the various annotation software and Perl scripts for format conversion.

For example, Figure 2 shows an extract of an annotation of LSF by two groups, shown in the ANVIL annotation software. The annotation of team 7 was performed using the Elan software. We used the Elan export function in CVS format and a Perl script able to build an ANVIL annotation file with these data. Then we fused the team 3 and team 7 ANVIL annotation files into one ANVIL annotation file

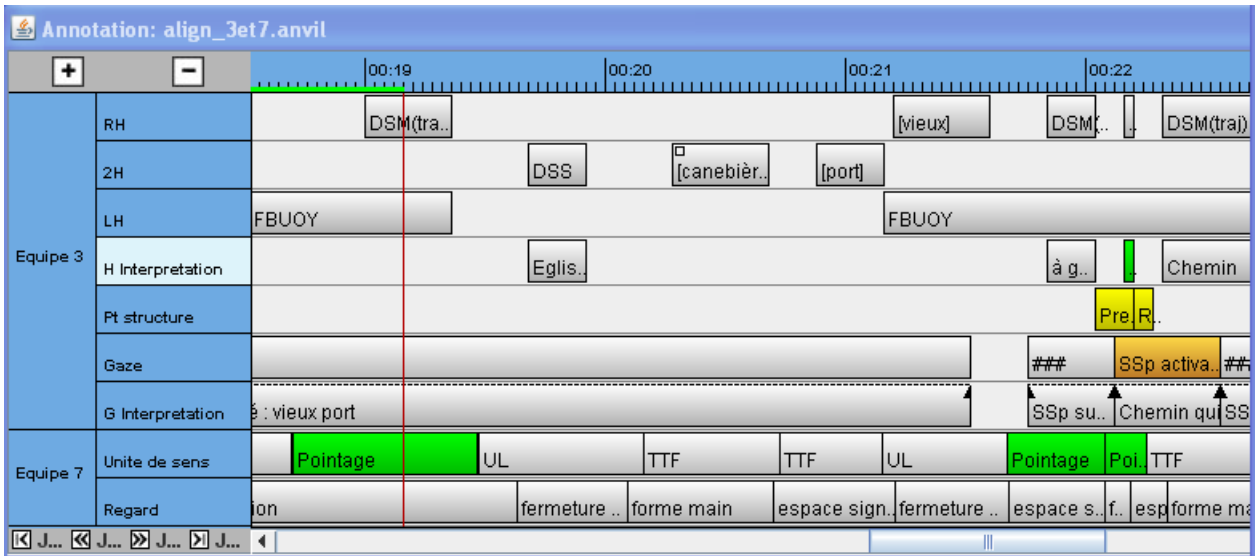


Figure 2: Extract from DEGELS1 annotations for LSF, aligned for two groups.

This annotation extract shows two groups of track: “Equipe 3” and “Equipe 7”.

The group “Equipe 3” is composed of seven tracks:

- “RH”, “2H”, “LH” are used to annotate the gestural units, using glosses for lexicalised signs or linguistic categories for the other gestural units, specifying if the gestural unit has been performed by the right hand, the two hands or the left hand ;
- “H interpretation” gives an interpretation of the meaning of non lexicalised signs ;
- “Pt structure” provides a fine segmentation of pointing temporal phases ;
- “Gaze” is used to annotate the gaze direction and “G interpretation” provides an interpretation of the place looked at.

The group “Equipe 7” is composed of two tracks:

- “Unité de sens” is used to annotate the gestural units with linguistic categories, without providing the glosses or the meaning and without differentiate the hands used ;
- “Regard” provides information on the gaze direction without information on the place looked at.

In addition to the differences observed on the annotation schemes and the level of details between the two annotations, we can observe that:

- the segmentation into gestural units doesn’t include transition for the group 3 (allowing gaps in the tracks) and includes the transitions for the group 7 (no gaps). That reflects different practices in the way to segment the gestural units;
- some gestural units have been categorised as depicting signs for the group 3 and as pointing for the group 7, as for example the gestural unit performed by the signer on the right in the video on figure 2, which frame corresponds to the position of the cursor on the annotation window (Figure 3).

After having asked the informant about his intent in this sequence, it was established that the gesture doesn’t seem to own a pointing function. That reflects the fact that a precise description of the objective criteria used to select categories can help to choose the good one.



Figure 3: Frame extracted from DEGELS1 corpus, showing a depicting structure in LSF, which represents a description of the way a car moves in a sloping road.

More broadly, this first comparison has highlighted practices in each community. Researchers working on gestures apprehend the temporal structure more accurately. Sign Languages (SL) researchers have a more advanced expertise in the study of spatial phenomena and the arrangement of the hands together (bimanual vs. mono, symmetry pattern, etc.) or with another body part. This is not completely a surprise: The gestural studies are used to

explore the relations between gestures and speech, in a temporal reference framework. Conversely, SL researchers, by the very nature of SL that intensively use localisation in space or relatively to the body in their grammar, need to handle three-dimensional spatial representations to accurately model the language. The aim of the next workshops will be to explain in detail and share the best practices of the two communities, and go toward common guidelines when possible.

DEGELS2011 also showed the impact of methodologies and tools on the analysis itself. This would be beneficial to become aware of that aspect, eventually promote practices that do not allow themselves to be compelled by more or less suitable existing tools, and exhibit the needs of researchers in order to move towards the development of tools that really support linguistic exploration.

## 5. Conclusion

In this paper, we have presented the DEGELS1, a comparable corpus of French Sign Language and co-speech gestures that has been created to serve as a testbed corpus for the DEGELS workshops. These workshop series aim to compare annotation and analysis methods for corpora of co-speech gestures in French and French Sign Language. We have also reported on the first edition of the workshop, DEGELS2011 which focused on pointing, a research topic shared by the two communities. The comparison between annotations has led us to propose the topic of segmentation for the DEGELS2012 event, with the aim to precise the objective criteria used by the participants for the different approaches.

We hope that this corpus and these events will help sharing the best methods and practices of the two communities. In a long term, we hope helping to address fundamental questions underlies modern linguistics, and that the inclusion of SL sheds a new light on, such as: Does the modality alone allows us to define what is verbal or nonverbal? What is the coverbal in SL and how can we define it?

## 6. References

- Boutora L., Braffort A. & Bertrand R. (2011). Présentation et premiers résultats du défi d'annotation DEGELS 2011 sur un corpus bilingue de français oral et de langue des signes. DEfi Geste Langue des Signes (DEGELS) 2011, atelier de Traitement Automatique des Langues Naturelles (TALN) 2011. <http://degels.limsi.fr/>
- Johnston T. (2011). Auslan Corpus Annotation Guidelines. <http://www.auslan.org.au/video/upload/attachments/AuslanCorpusAnnotationGuidelines30November2011.pdf>
- Kita, S., Van Gijn, I. & Van der Hulst, H. (1997). Movement Phase in Signs and Co-Speech Gestures, and Their Transcriptions by Human Coders. *Gesture and Sign Language in Human-Computer Interaction*, LNCS/LNAI 1371, pp. 23-35.
- Nonhebel, A., Crasborn, O. & Van der Kooij, E. (2004). Sign language transcription conventions for the ECHO project. Version 9, 20 January 2004. Ms., Radboud University Nijmegen. [http://www.let.kun.nl/sign-lang/echo/docs/transcr\\_conv.pdf](http://www.let.kun.nl/sign-lang/echo/docs/transcr_conv.pdf)
- Zwitserslood I., Özyürek A. & Perniss P. (2003). Annotation of Sign and Gesture Cross-linguistically, 3rd Workshop on the Representation and Processing of Sign Languages, 6th language resources and evaluation conference (LREC) 2008.