

- APPEL À PROPOSITIONS 2000 -  
 « LANGAGE ET COGNITION »

**RAPPORT DE FIN DE RECHERCHE**

**Projet n° : LACO 39**

**Titre du projet :**

Langues des signes : analyseurs privilégiés de la faculté de langage ; apports croisés d'études linguistiques, cognitives et informatiques (traitement et analyse d'image) autour de l'iconicité et de l'utilisation de l'espace.

**Durée du projet :** 24 mois

**Sous-thème(s) dont relève ce projet :**

- Langue des signes
- Sémantique-syntaxe
- Lexique-morphologie
- Traitement de l'Image
- Notation et transcription de corpus

**Responsable scientifique :**

*Nom et prénom* CUXAC Christian  
*Fonction* Professeur d'Université en Sciences du Langage  
*Organisme* Université Paris 8 - UMR 7023  
*Adresse* 2, rue de la Liberté, 93526 - Saint-Denis - Cedex  
*Téléphone* 01 49 40 65 35  
*Télécopie* 01 49 40 67 90  
*Mél* ccuxac@univ-paris8.fr

**Equipes partenaires :**

<i>Laboratoire</i>	<i>Equipe</i>	<i>Coordonnées</i>
CAMS/Paris IV	LaLIC	96 bd Raspail - 75 006 Paris
LIMSI/CNRS	Geste et Image	Bat 508- Campus Universitaire 91 403 Orsay Cedex
IRIT	Traitement et Compréhension d'Image	Université Paul Sabatier 118 route de Narbonne - 31062 Toulouse Cedex

**Organisme gestionnaire de l'opération : Université Paris 8**

## **résumé signalétique**

### **Résumé**

Les analyses linguistiques des LS (langue des signes) sont particulièrement délicates à mener, car la transcription linguistique et l'étiquetage informatique des séquences vidéo se heurtent à des spécificités telles que la quadrimensionnalité du support et la simultanéité de paramètres sémiotiques (signes manuels, regard, mimique faciale, mouvements pertinents du visage et du corps). Par ailleurs, pour les chercheurs spécialisés en traitement d'image, l'analyse de séquences vidéo de signeurs en LS présente l'intérêt particulier de se pencher sur des images véhiculant un sens explicite, et produites en appliquant les règles d'une grammaire.

Notre projet, issu de la rencontre de ces disciplines, s'est élaboré en trois temps :

Recensement des différentes démarches d'analyse d'un corpus de LS, des formalismes existants ainsi que des environnements informatiques les mettant en œuvre.

Réalisation d'un corpus vidéo de locuteurs en LS inédit sur le plan international.

Construction d'un éditeur de partition permettant de visualiser et de manipuler des séquences vidéo de LS et de réaliser des transcriptions de ces vidéos selon des critères utiles à la fois aux linguistes et aux informaticiens.

Notre projet doit déboucher sur une automatisation partielle par traitement d'images.

### **Mots-clés**

LSF (Langue des Signes Française)

Iconicité

Grammaire cognitive

Représentations spatiales

Traitement de l'image

**Nombre total de participants au projet : 18**

Chercheurs :

Sciences du Langage : 1

Informatique : 1

Enseignants-chercheurs et assimilés:

Sciences du Langage : 1

Informatique : 4

Doctorants et autres étudiants :

Sciences du Langage : 6

Informatique : 5

**Nombre total d'hommes-mois : 114**

# SOMMAIRE

<b>LISTE DES PARTICIPANTS</b>	<b>5</b>
<b>1 PRÉSENTATION DE L'ACTION LS-COLIN</b>	<b>6</b>
1.2 PRÉSENTATION DE L'ACTION LS-COLIN	6
1.2.1 HYPOTHÈSES SUR LA LANGUE DES SIGNES	6
1.2.2 OBJECTIFS DE L'ACTION	8
<b>2 ÉTAT DE L'ART ET DÉMARCHE RETENUE</b>	<b>10</b>
2.1 RELEVÉ DES DIFFÉRENTS SYSTÈMES D'ÉCRITURE ET DE TRANSCRIPTION EXISTANTS	10
2.1.1 LES SYSTÈMES DE NOTATION MANUELS	10
2.1.2 LES SYSTÈMES DE NOTATION INFORMATISÉS	11
2.2 LES TRANSCRIPTIONS RÉALISÉES DANS LE CADRE DE LS-COLIN	12
2.3 CONCLUSION	18
2.4 GRILLE D'ANALYSE (VALABLE POUR UN MONOLOGUE)	19
<b>3 CONCEPTION ET ÉLABORATION DU CORPUS</b>	<b>20</b>
3.1 SPÉCIFICATIONS DE RÉALISATION	20
3.1.1 LES LOCUTEURS :	20
3.1.2 LES TECHNIQUES D'ACQUISITION :	20
3.2 RÉALISATION	21
3.2.1 CES JOURNÉES ONT ÉTÉ ORGANISÉES PAR :	21
3.2.2 CONDITIONS TECHNIQUES :	21
3.3 ÉVALUATION	23
<b>4 RÉALISATION DE L'ÉDITEUR DE PARTITION</b>	<b>24</b>
4.1 SPÉCIFICATIONS	24
4.1.1 RÔLE DE L'ÉDITEUR	24
4.1.2 ASPECT GRAPHIQUE DE L'ÉDITEUR	24
4.1.3 DESCRIPTION DÉTAILLÉE DE LA PARTIE TRANSCRIPTION	25
4.1.4 GÉNÉRICITÉ, FLEXIBILITÉ, OUVERTURE	26
4.2 APPORTS DU TRAITEMENT D'IMAGES	26
4.2.1 POSSIBILITÉ D'AUTOMATISATION PARTIELLE PAR TRAITEMENT D'IMAGES	27
4.2.2 LE TRAITEMENT D'IMAGES DANS LE PROJET LS-COLIN	28
4.3 RÉALISATION	31
4.4 ÉVALUATION	31
4.5 AMÉLIORATIONS POSSIBLES	31
<b>5 EXPLOITATION DU CORPUS</b>	<b>33</b>
5.1 PREMIÈRE TRANSCRIPTION MANUELLE DU CORPUS LS-COLIN	33
5.1.1 OBJECTIFS	33
5.1.2 DÉMARCHE	33

<b>6</b>	<b>RÉSULTATS OBTENUS</b>	<b>41</b>
<b>6.1</b>	<b>DES ÉTAPES NÉCESSAIRES POUR ÉTUDIER LES LANGUES DES SIGNES...</b>	<b>41</b>
<b>6.2</b>	<b>... EN TANT QU'ANALYSEURS ÉPISTÉMOLOGIQUES...</b>	<b>43</b>
6.2.1	... DE LA FACULTÉ DE LANGAGE...	43
6.2.2	...ET DE LA COGNITION.	49
<b>7</b>	<b>CONCLUSION ET PERSPECTIVES</b>	<b>50</b>
	<b>LISTE DES PUBLICATIONS DU GROUPE LS-COLIN</b>	<b>51</b>
	<b>BIBLIOGRAPHIE</b>	<b>52</b>
	<b>ANNEXES</b>	<b>53</b>
<b>A.</b>	<b>FICHE D'ENGAGEMENT DU SIGNOREUR</b>	<b>54</b>
<b>B.</b>	<b>TABLEAU DE SYNTHÈSE : LES PRODUCTIONS EN LSF DU CORPUS LS-COLIN</b>	<b>55</b>
<b>C.</b>	<b>TABLEAU DE SYNTHÈSE : LES PRODUCTIONS DU CORPUS LSP ET LIBRAS</b>	<b>56</b>
<b>D.</b>	<b>PUBLICATIONS DU GROUPE LS-COLIN</b>	<b>57</b>

## LISTE DES PARTICIPANTS

NOM et Prénom	Laboratoire	Discipline	Nombre de mois consacrés au projet.
Christian Cuxac	Sciences du Langage, Université Paris 8	linguistique	
Ivani Fusellier		id	
Nathalie Monteillard		id	
Marie-Anne Sallandre		id	
Gwénaëlle Jirou		id	
Annie Risler		id	
Fanch Lejeune	CAMS-LaLIC, université Paris 4 et LIMSI, CNRS, Orsay	informatique	
Annelies Braffort	LIMSI, CNRS, Orsay	informatique	
Annick Choisier		id	
Christophe Collet		id	
Rachid Gherbi		id	
Patrice Dalle	IRIT-TCI, université Toulouse 3	informatique	
Guillemette Jausions		id	
Boris Lenseigne		id	

# 1 Présentation de l'action LS-COLIN

## Introduction

C'est dans le cadre de l'A.C.I. Cognitive 2000 « Langage et Cognition » qu'est né le projet LS-COLIN : « Langues des signes : Analyseurs privilégiés de la faculté de langage; apports croisés d'études linguistiques, cognitives et informatiques (traitement et analyse d'image) autour de l'iconicité et de l'utilisation de l'espace ». Il regroupe des équipes des universités de Paris 8 (Sciences du langage), Paris 4 (CAMS-LaLIC), Toulouse3 (IRIT-TCI) et du CNRS (LIMSI). La responsabilité scientifique est assurée par Christian Cuxac, linguiste, professeur à l'université Paris 8.

## 1.2 Présentation de l'action LS-COLIN

### 1.2.1 Hypothèses sur la Langue des Signes

L'étude des langues des signes permet d'entrevoir au plus près ce qu'est une langue, en remettant en question les frontières habituelles de ce que nous considérons comme appartenant au domaine du "linguistique".

#### 1.2.1.1 *Le cadre théorique*

Nous proposons une analyse linguistique des langues des signes, dont le principe fondateur est la théorie de l'iconicité définie par Cuxac [Cuxac, 1996 & 2000], elle-même inscrite dans une sémiologie plus générale. Cette théorie peut s'appliquer à toutes les langues des signes (LS) : les LS standardisées comme la LSF, mais aussi les autres langues comme la LSI (Langue des signes internationale), les langues de petites communautés de Sourds et enfin les langues des signes primaires mises au point par des Sourds isolés.

Dans cette théorie, il est essentiellement question de l'iconicité dite d'image, c'est-à-dire du lien de ressemblance formelle entre le signe et ce à quoi il réfère, dans le monde réel. Ce type d'iconicité structure le lexique des langues des signes, et rend compte de la constitution des signes dits standard. Cependant, un autre type d'iconicité, diagrammatique, intervient de façon massive dans la structure syntaxique des énoncés en langue des signes. On note une ressemblance entre l'agencement des signes dans l'espace et le rapport des éléments dans l'espace de référence. Cette iconicité, qui n'est pas spécifique aux langues signées, trouve son origine au niveau langagier des représentations sémantico-cognitives. Elle n'est cependant plus autant perceptible dans les langues vocales [Haiman, 1985] du fait de l'obligation de linéariser l'énoncé oral. Au contraire, elle se maintient selon certaines règles en langues signées du fait de leur expression dans une modalité gestuelle spatio-temporelle.

Nous savons grâce aux travaux de Goldin-Meadow, Yau, Fusellier-Souza que des personnes sourdes vivant en milieu entendant ou vivant en petites communautés ont mis en œuvre des stratégies communicationnelles fondées sur un processus d'iconicisation de l'expérience perceptivo-pratique. Celui-ci est à l'origine de toutes les langues des signes pratiquées dans le monde, quel que soit leur statut institutionnel. De ce fait, ces langues constituent aujourd'hui un terrain d'observation privilégié des faits synchroniques concernant la sémiogénèse

(formation de signes ou de systèmes de signes par une communauté linguistique donnée) des langues.

Dans cette perspective sémiogénétique, la bifurcation postulée par Cuxac [Cuxac, 2000] détermine deux pôles qui coexistent et entre lesquels le va-et-vient est constant, dans tous types de discours. On a d'une part la grande iconicité, traces structurales d'une visée illustrative (ou iconocisatrice) de l'expérience vécue, qui donne à voir tout en disant, d'autre part les signes standard, sans visée illustrative. C'est par ces deux grands axes qu'une grammaire de la LS peut être établie.

Il est étonnant d'observer que les langues des signes naturelles utilisent des structures autant syntaxiques que sémantiques présentant un grand degré d'iconicité. Ainsi, on peut exprimer des idées complexes sans recourir au lexique standard. Cependant, il est intéressant d'analyser à quel moment du discours les signes standard sont nécessaires et inversement, et dans quels cas la grande iconicité est obligatoire (par exemple dans le cas d'énoncés absurdes qui ne peuvent être traduits que par un transfert personnel) ; on cite souvent l'exemple de Cuxac [1996] "le chocolat mange le garçon".

Nous tentons de démontrer que ces structures de grande iconicité, trop souvent considérées comme de la pantomime, sont bien des éléments linguistiques et constituent même la démonstration la plus subtile et la plus convaincante de ce qu'est une langue. Elles sont quasiment identiques d'une langue des signes à l'autre, ce qui permet une intercompréhension rapide entre des signeurs de nationalités différentes.

Toutefois, aucune étude systématique n'ayant encore été menée sur ces structures, des questions fondamentales demeurent.

Les grammaires cognitives apportent un cadre d'analyse pertinent à tous les niveaux d'analyse de la LS. Ces grammaires offrent effectivement des points de vue intéressants sur une langue iconique et spatiale. Elles nous permettent d'envisager l'iconicité pour les langues orales donc a fortiori pour les langues gestuelles. Elles envisagent le passage du cognitif au linguistique en trois niveaux, cognitif, langagier et de la langue. Aux deux premiers niveaux se réalisent des opérations de représentation, cognitives et langagières communes à toutes les langues. Chaque langue détermine son propre système, à travers son lexique et sa syntaxe, mais les opérations langagières sont toujours du même type. Une langue gestuelle peut s'inscrire dans un tel modèle. Enfin, et surtout, ce modèle recourt à un vocabulaire abstrait topologique et spatial pour formaliser les opérations langagières, ainsi qu'à des schématisations qui rappellent fortement l'agencement des signes dans l'espace de signation. On peut mettre en évidence en LSF une construction de l'espace du même ordre que les constructions envisagées par les grammaires cognitives : repérages, orientation de la relation prédicative, positionnement de l'espace par rapport au signeur.

### ***1.2.1.2 Sur l'apport du traitement d'image***

Les recherches menées en France sur la LS tant en linguistique qu'en informatique, ont ceci en commun qu'elles exploitent des données réelles de terrain en s'appuyant sur des corpus d'images vidéo. Cette démarche, nécessaire à l'analyse de la LS, est néanmoins délicate à mener, car elle nécessite la réalisation d'une transcription (linguistique) ou étiquetage (informatique) des séquences vidéo, en tenant compte de divers aspects :

- la quadrimensionnalité du support ;
- la prise en compte simultanée de paramètres sémiotiques tels que les gestes, le regard, la mimique faciale ainsi que les mouvements du visage et du corps ;
- la difficulté du maniement de systèmes de transcription tenant compte de tous ces effets de sens ;

- l'inadaptation des rares logiciels d'aide à la transcription (quand ils sont accessibles) de documents vidéos.

Cette transcription rend le travail d'analyse extrêmement long et fastidieux.

Un domaine de recherche particulièrement intéressant concerne l'élaboration de systèmes informatiques dédiés à la LS, dont le but est de faire de l'analyse, de la reconnaissance ou de la génération automatique. Dans le cadre de notre étude, nous utilisons des caméras vidéo, qui permettent de capter les mouvements des mains, mais aussi du buste, les expressions faciales ainsi que le regard. Les séquences d'images ainsi enregistrées servent de support aux études informatiques notamment dans le domaine du traitement d'images.

Pour les chercheurs spécialisés en traitement d'image, l'analyse de séquences vidéo de signeurs en LS présente un intérêt particulier :

- C'est un des rares cas où l'image véhicule un sens explicite, où elle est produite avec une intention de communication et où les mouvements sont produits en respectant les règles d'une grammaire (contrairement aux scènes habituellement traitées en vision par ordinateur). De ce fait, on peut exploiter cet aspect sémantique et cette base linguistique pour introduire de nouveaux mécanismes d'analyse, de représentation et de reconnaissance (choix des primitives visuelles, analyse multi-niveaux, prédiction pilotant l'analyse, etc.).
- De plus, l'étude des relations spatiales et la conception d'outils d'interaction portant sur l'organisation spatiale, sont des éléments essentiels qui peuvent être ré-exploités dans le cadre de systèmes d'Interaction Homme-Machine (IHM), ou dans celui d'IHM de système d'interrogation d'images ou de construction de scènes.

### 1.2.2 Objectifs de l'action

Nous croyons qu'une recherche croisée de la linguistique et de l'informatique peut nous apporter des réponses intéressantes concernant le fonctionnement interne des langues des signes. La technologie informatique peut formuler une modélisation de paramètres qui pourrait valider l'hypothèse de la bifurcation fonctionnelle prédite par C. Cuxac. Concernant la recherche sur l'espace en LSF, le recours à l'analyse d'image peut s'intégrer de deux manières : d'une part afin de mettre à l'épreuve les paramètres formels dégagés, qui ne sont pas de type phonologique mais morphémique ; d'autre part, afin d'objectiver la construction de l'espace telle qu'elle est réalisée dans un énoncé, par la mise en relief des portions d'espace pertinences, et des relations entre les espaces topologiques.

Pour réaliser cette étude nous avons dans un premier temps confronté nos approches de l'analyse de corpus vidéo sur la LS. Nous avons ensuite défini les objectifs suivants :

- Extraire des primitives signifiantes de séquences vidéo utiles pour la reconnaissance et l'interprétation tant au niveau linguistique qu'au niveau informatique ;
- Trouver et quantifier des structures dans une séquence d'image (événements périodiques, configurations...) ;
- Mettre au point un système de transcription manuel à partir de vidéo numérique adapté aux besoins des linguistes et des informaticiens en tenant compte des différentes approches.

#### **Rappel des enjeux et objectifs fixés à l'origine**

L'apport croisé des recherches en linguistique et en informatique doit permettre de créer de nouveaux outils pour l'analyse du fonctionnement interne des langues des signes (LS) et leur traitement automatique.

Les modèles informatiques de reconnaissance des formes et de traitement d'image permettent d'extraire d'un corpus vidéo des paramètres pertinents pour la compréhension des LS. Ils contribuent à une validation et à une formalisation de modèles linguistiques concernant l'iconicité et l'espace en LS.

L'iconicité des messages fonctionne comme analyseur cognitif. On sera amené :

- à affiner l'inventaire des structures dites de "grande iconicité" définies par Cuxac (2000);
- à modéliser la manière dont peu à peu, un large vocabulaire standard a pu émerger du processus d'iconicisation initial;
- à vérifier, à partir de productions de locuteurs sourds en LSF, ainsi qu'à partir de créations gestuelles des sourds isolés, certaines des hypothèses des grammaires cognitives, en particulier l'expression complémentaire des primitives spatiales et non spatiales issues des représentations sémantico-cognitives dans les énoncés en LS.

Concernant la recherche sur l'espace en LSF, le recours à l'analyse d'image peut s'intégrer de deux manières :

- mettre à l'épreuve les paramètres formels dégagés, qui ne sont pas de type phonologique mais morphémique ;
- mettre en évidence la construction de l'espace telle qu'elle est réalisée dans un énoncé, par la mise en relief des portions d'espace pertinences, et des relations entre les espaces topologiques.

Les modèles linguistiques découlant de cette étude, serviront de base à l'élaboration de systèmes informatiques pour le traitement et la reconnaissance de langues des signes à partir d'images captées par caméra vidéo.

Nous avons dans un premier temps confronté nos approches sur les différentes démarches d'analyse d'un corpus vidéo de LS. Nous avons ensuite défini les objectifs suivants :

- extraire des primitives signifiantes de séquences vidéo utiles pour la reconnaissance et l'interprétation tant au niveau linguistique qu'au niveau informatique ;
- trouver et quantifier des structures dans une séquence d'image (événements périodiques, configurations...);
- mettre au point un outil de transcription automatisé à partir de vidéo numérique adapté aux besoins des linguistes et des informaticiens en tenant compte des différentes approches.

Dans un deuxième temps, notre objectif était de trouver un mode de transcription commun aux linguistes et aux informaticiens qui permettrait :

- de passer entre les niveaux (du formel au signifié) ;
- de guider, cibler voire anticiper les traitements d'images (prédiction-vérification).