

---

# Problématique des chercheurs en traitement automatique des langues des signes

**Christian Cuxac\*** — **Patrice Dalle\*\***

*\*Université Paris 8 – UMR 7023  
2, rue de la Liberté  
F-93256 Saint-Denis cedex  
ccuxac@univ-paris8.fr*

*\*\*IRIT Université de Toulouse – UMR 5505  
118, route de Narbonne  
F-31062 Toulouse cedex 9  
dalle@irit.fr*

---

*RÉSUMÉ. Les chercheurs en traitement automatique de la langue des signes française (LSF) sont amenés à étudier et à traiter une langue encore incomplètement décrite, sans forme écrite et à partir de corpus fragmentaires. Ils doivent donc opérer des réductions à tous les niveaux de représentation de la langue, contraindre les conditions d'acquisition des discours en LSF et limiter le champ de leurs applications. Il est important de bien expliciter et justifier ces réductions et d'en connaître les implications. Nous présentons les caractéristiques de la LSF et de son évolution et indiquons les différentes approches permettant de réduire la complexité des traitements tout en conservant l'essentiel des spécificités de la LSF.*

*ABSTRACT. French Sign Language (LSF) automatic processing research has to deal with uncomplete corpora of a language which is not yet fully described and has no written form. So, the researchers must carry out reductions at every representational level, constrain the conditions of acquisition of LSF discourses and limit the field of their applications. It is important that they should explain and justify those reductions and know what such reductions imply. In this paper, we will present some of the characteristics of LSF organisation and evolution as well as the different approaches that allow to reduce the complexity of the treatments while preserving most of LSF peculiarity.*

*Mots-clés : iconicisation, création lexicale, analyse d'images, génération, contraintes.*

*Key words: iconization, lexical creation, image analysis, generation, constraints.*

---

## 1. Introduction

Ce numéro thématique de la revue TAL est le premier consacré aux langues des signes (LS). Aussi, pour les lecteurs peu familiers de ces langues, il nous a paru utile de décrire le contexte dans lequel s'effectuent les recherches en traitement automatique des langues de signes (TALS). Si les premières recherches linguistiques sur les LS ne sont pas récentes, elles se sont amplifiées depuis les travaux de Stokoe (dont le dictionnaire sur la LS américaine a été publié en 1965), mais leurs descriptions sont encore très incomplètes et, suivant les points de vue, très différentes. Les recherches en TALS sont, quant à elles, beaucoup plus récentes et encore peu développées. Les LS n'étant officiellement reconnues que très récemment (depuis 2005 en France), elles évoluent encore beaucoup, au fur et à mesure qu'elles investissent les différents espaces socioculturels. Cette évolution est d'autant plus libre que ces langues n'ont pas de formes écrites qui pourraient les stabiliser et avoir un effet normatif. Cette absence d'écriture impose aux chercheurs en TALS de travailler à partir de données vidéo. Il est maintenant aisé d'obtenir des vidéos de qualité avec des caméras standard, mais ceci n'est vrai que depuis une décennie. D'autre part, des contraintes, comme le respect du droit à l'image, la faible densité de la population sourde, la difficulté de trouver des collaborateurs sourds intéressés par ces travaux ou suffisamment formés pour y être associés, ont jusqu'à présent freiné la constitution de corpus élaborés suivant des critères rigoureux et suffisamment riches et complets. Enfin les outils d'analyse robustes et performants qui permettraient d'exploiter ces corpus de manière automatique font encore défaut. Il en est de même pour les outils de génération et de restitution d'énoncés en LS.

Le chercheur en TALS doit donc modéliser, représenter et traiter une langue encore incomplètement décrite, peu stabilisée, sans forme écrite, à partir de corpus fragmentaires et avec des outils insuffisants. Il devra donc opérer des réductions à tous les niveaux de représentation de la langue, se placer dans des conditions opératoires maîtrisées et viser des applications au champ bien délimité. Pour que la langue étudiée ne soit pas une version de laboratoire peu conforme à la réalité et pour que ces applications ne constituent pas des réalisations *ad hoc* non transférables, il est important de bien expliciter et justifier ces réductions et d'en connaître les implications.

## 2. Compilation des données vs théorie(s) linguistique(s)

Le traitement automatique d'une langue relève en quelque sorte d'une problématique analogue à celle de la carte et du territoire : on ne peut tout consigner sauf à dupliquer l'ensemble des données attestées sur le terrain. Ce n'est pas tant l'aspect surhumain d'une telle tâche qui pose problème c'est en premier lieu son

caractère trivial dans la mesure où, n'étant pas portées par un regard métalinguistique, toutes les données s'équivalent. La description d'une langue, pour être pertinente, se construit à partir d'un cadre théorique, non à partir d'une accumulation de données. Cela, qui paraît aujourd'hui d'une certaine banalité, a été exposé par Chomsky il y a une cinquantaine d'années comme problème fondamental et résolu de la manière suivante : la carte doit être une formalisation à même de générer toutes les possibilités du territoire et rien qu'elles. Cette manière de poser le problème est devenue incontournable pour tout ce qui relève du traitement automatique des langues, quelles que soient les orientations théoriques des chercheurs, que la carte propose un ensemble de règles ou de contraintes proscriptives, et quand bien même on ne partagerait pas les prises de position innéistes et/ou modularistes des grammaires génératives.

Le fait que toute description présuppose une formalisation n'annule pas pour autant la confrontation du modèle formel à l'épreuve des faits attestés sur le terrain.

On a pu objecter que le recours à l'intuition du locuteur chercheur permettait de se passer du patient recueil de données. Que le chercheur ait des intuitions est indéniable, mais, en mettant en jeu l'acceptabilité d'un énoncé ou d'une phrase, elles sont tout au plus d'ordre épilinguistique, et restent sans fiabilité pour ce qui est du niveau de conscience métalinguistique. Ainsi, pour ne prendre que l'exemple de l'identification des unités minimales, quel francophone lettré et même un tant soit peu formé à la linguistique – on peut en faire le test avec des étudiants en L3 de sciences du langage – a l'intuition métalinguistique qu'en français /ε/ est le morphème du pluriel ? De plus, sans expérience de terrain les intuitions du chercheur ne peuvent que concerner au mieux les usages proches de son propre idiolecte.

Car la langue n'est pas un ensemble homogène : toutes les données ne sont pas identiques en fréquence et ne sont pas partagées par l'ensemble des membres de la communauté linguistique. Aussi, pour ne pas retomber dans le cercle épistémologiquement vicieux d'une compilation exhaustive comme seule garante de l'objectivité du chercheur, il devient nécessaire pour ce dernier de procéder à des choix et de donner des ordres de priorité.

Ce qui revient à dire que toute formalisation est – même si l'on s'en défend –, assortie d'un point de vue normatif-prescriptif qui s'ignore la plupart du temps comme tel. Et, bien évidemment, c'est là que le bât blesse. Toutefois, et indépendamment du cadre théorique adopté par le chercheur qui relève en dernière instance de sa subjectivité, une hiérarchisation objective des données est possible à partir de la procédure de dégagement de ce que l'on appelle la « norme descriptive ». On donnera un exemple relatif au système vocalique du français. Le terrain révèle une importante diversité des usages : dialectes à trois ou quatre degrés d'aperture, quelques locuteurs réalisant même des oppositions systématiques de longueur (mettre/maître, lit/lie, bout/boue)... À Paris, les jeunes générations n'attestent plus l'opposition /a/ d'avant, /a/ d'arrière, alors que les plus âgés la

maintiennent. Le nord de la France ignore l'opposition arrondi/rétracté des nasales d'avant de troisième degré d'aperture, le sud la maintient, etc. En accumulant l'ensemble des données recueillies sur le terrain de l'hexagone on obtient ce qu'il est convenu d'appeler le système plein qui consiste en l'addition de toutes les caractéristiques des structurations présentant le plus fort taux de complexité, c'est-à-dire les plus marquées. À partir de cette collecte, se dégage objectivement une norme en ne prenant en compte que ce qui est commun à l'ensemble des locuteurs et qui se définit comme le minimum nécessaire, mais suffisant, à une intercompréhension non problématique. La hiérarchisation des données opérée par le chercheur devra donc s'effectuer prioritairement en respectant l'intégralité et l'intégrité de la norme, en sachant toutefois que celle-ci peut fluctuer en fonction de la modalité (orale ou écrite), des registres et des genres discursifs. Ainsi, le passé simple dont la connaissance excède la norme de l'oral est indispensable à la norme du français écrit, alors que ce n'est pas le cas de l'imparfait du subjonctif.

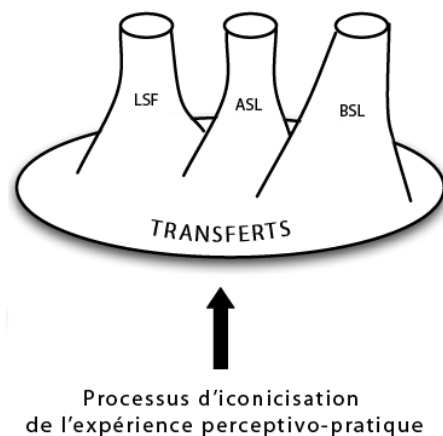
Il faut alors que la machine fabriquée par le chercheur puisse établir un lien entre système plein et système minimal afin que, dans l'idéal d'une communication homme machine, cette dernière ait la compétence du système plein en compréhension et de la norme minimale en production, et que les deux soient, comme dans la communication humaine, interprétables mutuellement.

### **3. La LSF à l'épreuve du TAL**

Le modèle descriptif présenté dans la partie qui va suivre accorde un statut authentiquement linguistique aux occurrences fortement iconiques que l'on trouve massivement en LSF – en particulier dans les activités narratives et descriptives – ainsi que dans différentes autres langues des signes, comme par exemple en LIS et en ASL (Antinoro Pizzuto, Rossini, Sallandre et Wilkinson, 2006; Vermeerbergen, 2006).

Sa longue interdiction en milieu scolaire, le peu d'ouvrages consacrés à sa description, le manque d'informations sur ses différentes variantes dialectales font de la langue des signes française un objet d'autant plus fragile qu'il n'existe pas de norme écrite permettant de la fixer (Garcia, 2007). Bien plus qu'avec des langues très documentées, normées, voire surnormées comme l'anglais ou le français, la formalisation élaborée par le chercheur en TALS(F) est exposée à un double risque : celui, pragmatique, de rater la communication avec le public attendu, et celui, même s'il s'en défend, qui engage sa responsabilité d'exercer une influence négative sur l'évolution de cette langue. D'autre part, la LSF, après une longue stigmatisation, connaît actuellement une période propice à une importante créativité lexicale, suite à son retour progressif dans les établissements scolaires. Il est donc impératif que le chercheur en plus d'avoir une vision précise du système plein soit au fait des forces actuellement en jeu dans l'évolution de cette langue. On les rappellera brièvement dans la partie qui va suivre.

Il est à présent clairement établi que toutes les LS pratiquées aujourd'hui dans le monde ont eu pour point de départ des *homesigns* (Goldin Meadow, 1991), productions linguistiques fondées sur une iconicisation de l'expérience perceptivo-pratique (Cuxac, 1996 et 2000), que les jeunes enfants sourds antérieurement à toute rencontre avec une communauté de sourds créent dans une perspective communicationnelle avec leur entourage entendant (Yau, 1992 ; Fusellier-Souza, 2001 et 2004). Par la suite, au hasard de rencontres, ces personnes sourdes, devenues adultes, se sont spontanément constituées en microcommunautés avant même que des expériences de regroupement institutionnel (scolarisation) aient lieu (Jirou, à paraître). L'institutionnalisation va systématiser et renforcer à grande échelle un phénomène déjà attesté dans les langues des signes microcommunautaires, phénomène qui consiste en une bifurcation de visées sémiologiques. D'une part les produits que sont les *homesigns* subissent une déperdition de leur iconicité initiale, deviennent moins coûteux articulatoirement et voient leur forme s'adapter économiquement à la réception visuelle du message (Cuxac, 2004 ; Bonnal, 2005), d'autre part, le processus d'iconicisation de l'expérience, toujours activable, donne lieu à des productions très fortement iconiques de plus en plus sophistiquées au fur et à mesure du développement et de la maturation psychomotrice et cognitive des locuteurs sourds. Les premiers, dont la forme s'est économiquement stabilisée, sont les éléments d'un lexique standard entrant dans la visée de dire sans intention de donner à voir, les seconds, appelés transferts, sont des structures complexes utilisées dans le cadre d'une visée sémiologiquement marquée qui consiste à dire tout en donnant à voir. Ces structures, largement décrites (Cuxac, 2000 ; Sallandre, 2003 ; Fusellier-Souza, 2004) constituent le point d'aboutissement extrême du processus d'iconicisation. Loin d'être assimilables à une pantomime holistique, elles relèvent au contraire d'une organisation compositionnelle morphémique très complexe (Cuxac, 2004). En effet, bien qu'étant une forme minimale de réalisation (à l'égal d'un signe standardisé), un transfert, lorsqu'il est traduit dans une langue vocale comme le français, correspond à un énoncé complet, voire, comme c'est le cas dans un double transfert, à plusieurs énoncés multiactanciels où chaque élément qui le compose (direction du regard, expression du visage, posture corporelle, configuration et orientation de chacune des mains, mouvement de la ou des mains et emplacement de départ et d'arrivée du mouvement) peut être porteur de sens. La simultanéité de la réalisation de ces éléments compositionnels et leur interdépendance ravivent donc la dialectique saussurienne du paradigme et du syntagme : les éléments sont commutables au niveau du paradigme et, en ce sens, ont un statut morphémique, leur fonction au niveau du syntagme dépend de leur encadrement contextuel, à savoir de l'ensemble des autres composants articulés simultanément. Leur ancrage cognitif dans la perception action et leurs caractéristiques iconiques font que des sourds pratiquant des langues des signes différentes ont recours à ces structures de transfert lorsqu'ils désirent communiquer entre eux (Monteillard, 2001). Le schéma suivant (figure 1) illustre la façon dont s'est opérée la sémiogenèse des langues des signes. Il est censé rendre compte des points qu'elles ont en commun.

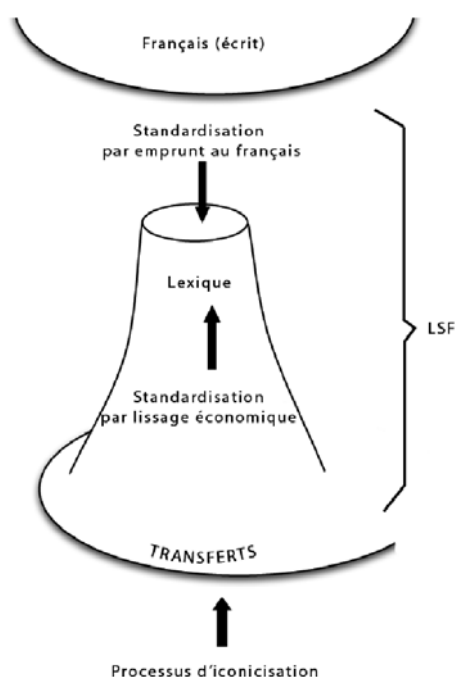


**Figure 1.** *Sémiogénèse des langues des signes*

La partie constituant la base issue du processus d'iconicisation renvoie aux structures de transfert, à peu de chose près, partagées par l'ensemble des langues des signes, alors que les colonnes montantes concernent chaque LS particulière (LSF, ASL, etc.), et caractérisent les signes standard qui leur sont spécifiques. La partie intermédiaire entre la base (transferts) et le purement spécifique (lexique standard) rend compte des points de recouvrement partiel entre les LS différentes. Ce phénomène est dû à une organisation grammaticale à iconicité diagrammatique (cf. Cuxac et Sallandre, 2007), qui est une manière économique de traiter les effets de sens en reliant spatialement les signes standard les uns par rapport aux autres.

Par ailleurs, la LSF est une langue dominée en contact permanent avec le français (surtout écrit) qui n'est pas sans influencer la structuration de son lexique standard. Le schéma suivant (figure 2) fait voir les deux forces complémentaires agissant en sens contraire à l'œuvre dans le processus de création lexicale : standardisation du lexique issu du processus d'iconicisation et des structures de transfert par lissage économique, et standardisation par emprunt à la langue dominante.

Enfin, une partie du lexique standard de la LSF a pour source des éléments empruntés à la gestuelle coverbale de l'environnement entendant (Delaporte, 2002 ; Wilcox, 2004 ; Cuxac à paraître). Ceci est particulièrement vrai pour les emblèmes (« chut ! ») et certains stéréotypes gestuels culturalisés (se gratter la tête comme indice de perplexité).



**Figure 2.** *Processus de création lexicale*

Concernant les emprunts, le chercheur devra être particulièrement sensible au fait que s'intégrant dans un système autonome ayant sa propre logique organisationnelle, ils y sont comme instantanément recyclés. Ainsi même lorsqu'il s'agit d'un emprunt de la totalité du signifiant comme, par exemple, P U R, épelé en dactylologie, il ne faut pas croire pour autant que c'est un décalque des valeurs signifiées du mot français « pur ». De même, la reprise signifiante de l'injonction emblématique « chut ! » peut être investie en fonction du contexte d'une valeur substantivale et signifier « le silence », ou verbale et signifier « se taire », « ne rien dire ».

Bien sûr, il ne saurait être question de proposer en application un traitement automatique couvrant la totalité de la langue des signes. Des réductions à tous les niveaux (grammaire, lexique) seront forcément opérées en fonction des tâches et des domaines sémantiques qui intéressent le chercheur.

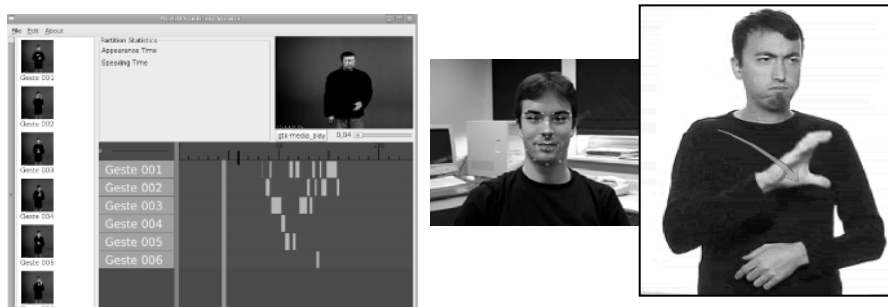
Si, par exemple, pour implémenter des dialogues à valeur générique (« où se trouve [...] ? » « à quelle heure [...] ? »), les signes standard peuvent suffire, les descriptions spécifiques de lieux, ou bien les communications en contexte international (aéroports) ne peuvent se passer d'avoir recours aux transferts.

## 4. Le TALS

### 4.1. Les applications du TALS

On retrouve en traitement automatique de la langue des signes (TALS) les mêmes familles d'applications qu'en traitement automatique des langues possédant une forme orale et une forme écrite (TAL), mais avec des degrés de développement assez différents.

Les recherches linguistiques sur la LSF étant relativement récentes, elles manquent encore d'outils adaptés à son analyse. Aussi, à un premier niveau, les applications sont-elles destinées aux linguistes eux-mêmes : programmes informatiques d'analyse de corpus ou d'illustration du fonctionnement de la LSF (segmentation, structuration, mesures et caractérisations, construction de l'espace de signation, éditeur pour l'annotation d'énoncés en LSF, validation de génération d'énoncés par signeur virtuel...) (Braffort et Dalle, 2007) (Lefebvre-Albaret et Dalle, à paraître).



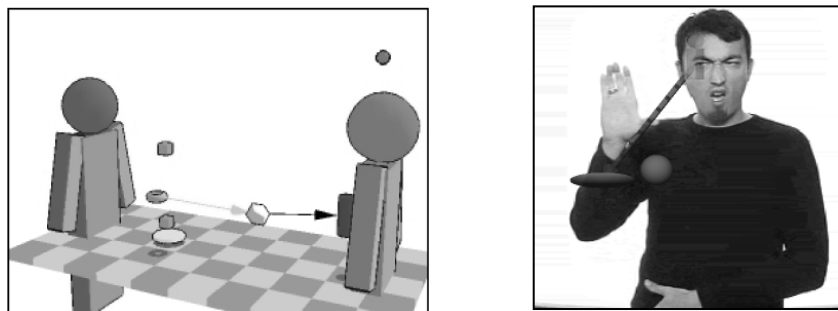
**Figure 3.** Outil d'annotation de vidéo (Ancolin) incluant du traitement d'image (segmentation, détection, mesures), un module d'analyse des expressions du visage, et un module de suivi et de visualisation des mouvements des mains

Un deuxième niveau concerne des applications qui sont encore de nature linguistique et portent sur la LSF : fabrication et utilisation de dictionnaires,



indexation et recherche de signes dans un corpus, logiciels d'aide à l'enseignement ou à l'apprentissage de la LSF. Il y a encore peu de logiciels opérationnels.

□



**Figure 4.** *Modélisation et construction semi-automatique de l'espace de signation*

Le troisième niveau concerne les activités en langue des signes ; ces applications peuvent couvrir tous les domaines, mais beaucoup sont encore à l'état de maquette : compréhension d'énoncés en LSF, recherches par le contenu, requêtes en LSF, messagerie en LSF, production de messages par signeur virtuel, dialogues en LSF, aide à la traduction français-LSF...

On voit que ces applications concernent aussi bien l'analyse que la synthèse, la tendance des projets de recherche en cours étant de coupler et de faire coopérer ces deux modes. Dans tous les cas, elles doivent s'appuyer sur une modélisation de la LSF, une formalisation et sa représentation informatique.

La complexité et la variabilité des informations à traiter et des modèles à élaborer, selon ces deux aspects d'analyse et de synthèse, rendent le problème insoluble, actuellement, dans le cas général couvrant la totalité de la LSF et acceptant tous les types d'usage.

Pour pouvoir proposer des formalisations et construire des applications opérationnelles, on va donc être amené à simplifier les modèles, à introduire de fortes contraintes sur les données, à apporter au système des connaissances *a priori* et à envisager un traitement non automatique sollicitant l'intervention de l'utilisateur.

Afin de limiter la complexité des énoncés à traiter, on va introduire des restrictions de nature linguistique et des contraintes au niveau des conditions d'énonciation.

#### **4.2. Modélisation informatique de la LSF : quelles restrictions ?**

La modélisation de la LSF sera partielle en restreignant les structures grammaticales prises en compte. La limitation portera souvent sur les structures de transferts, notamment les structures de transferts personnels, très difficiles à prendre en compte totalement. Quand on sait que dans certains types d'énoncés (récits), ces structures peuvent être majoritaires, on mesure la perte de représentativité des énoncés traités. Les signes autorisés dans un énoncé peuvent être restreints à un champ sémantique donné, afin de réduire la richesse et la variété lexicale. Enfin le nombre d'entités différentes dans un même énoncé peut être réduit pour simplifier la structure de l'énoncé.

D'autre part, des contraintes peuvent être introduites sur les situations d'énonciation. Un même message pouvant être signé de différentes façons, on va limiter ces variations dues aux locuteurs (manière de signer, morphologie...) et aux conditions de production de l'énoncé. Pour cela, on peut limiter le nombre de locuteurs présents simultanément (dialogues) ou le nombre de locuteurs différents connus du système (analyse) ou modélisés dans le système (synthèse), ce qui permet d'apprendre ou d'introduire des connaissances *a priori* sur ces locuteurs. Enfin, on peut imposer aux locuteurs des contraintes sur la manière de signer (vitesse, amplitude, rythme, etc.), de façon à réduire la variabilité des styles ou limiter l'effet des coarticulations.

#### **4.3. Analyse et génération des énoncés en LSF : réduction de la complexité**

La LSF n'ayant pas de forme écrite, les énoncés seront traités ou générés sous forme d'enregistrements vidéo. On peut intervenir sur la complexité des images ou introduire des connaissances sur leurs conditions de production.

Afin de réduire la complexité de l'apparence des locuteurs pour faciliter en analyse le suivi des composants corporels et la segmentation, il est possible d'agir sur différents paramètres de la scène. On peut contraindre la position des locuteurs (fixe ou variable) ainsi que leur attitude (observation toujours de face, par exemple). Certains artefacts seront évités (lunette, barbe, cheveux masquant partiellement le visage...). Des restrictions seront imposées sur les vêtements, leur type (manches longues, pas de vêtement amples) et leur couleur (différente de la couleur de la peau et de la couleur du fond de la scène). Les caractéristiques du fond auront un impact sur la qualité d'estimation de la silhouette. Par exemple, sa complexité (uni/texturé) et sa variabilité (fixe/évolutive) peuvent être limitées.

En synthèse, on peut également réduire la complexité de la structure des signeurs virtuels (nombre de facettes composant leur « peau » ou nombre de segments composant leur « squelette ») afin de simplifier ou d'alléger les calculs.

La vidéo étant obtenue par un système de vision, en analyse, on adoptera un protocole d'acquisition rigoureux, d'une part en maîtrisant, si possible, les conditions d'acquisition pour limiter les variations – éclairage contrôlé, caméras de qualité (stabilité, résolution, nombre d'images par secondes) –, et, d'autre part, en enrichissant l'information par l'usage de plusieurs caméras (stéréovision) et de plusieurs vues (plan américain + gros plan sur le visage + vue de dessus et/ou de profil). Enfin, en obtenant des informations sur les paramètres du système de vision (position, focale) qui seront exploitées lors du traitement des images pour retrouver des informations sur le locuteur.

Toutes ces remarques sur la nécessité de contrôler les caractéristiques du locuteur ou les paramètres d'acquisition peuvent sembler superflues car relevant du bon sens. Cependant les traiteurs d'images savent bien que dans la détection et la reconnaissance d'un objet par un système de vision, la difficulté provient autant de la complexité et de la variabilité du contexte que des caractéristiques de l'objet étudié lui-même. On est alors devant le dilemme suivant : soit nous n'intervenons pas sur ces caractéristiques et alors les parties de corpus que l'on sera capable actuellement de traiter seront très réduites, soit on contraint certaines caractéristiques, il s'agira alors de scènes fabriquées pour les besoins de l'étude, et il faudra veiller à ce que la production des énoncés ne soit pas biaisée par ces conditions artificielles.

En synthèse, on pourra s'appuyer sur certaines connaissances des propriétés biomécaniques du corps humain (amplitude des flexions des articulations, des mouvements des éléments du visage tels que les sourcils, dynamique des mouvements des bras, du buste...). La quantification de ces données nécessite des études en amont à partir de corpus vidéo ou 3D obtenus à l'aide de systèmes de capture de mouvements.

#### **4.4. Applications : introduction de contraintes**

On peut enfin introduire des contraintes sur l'application et son mode opératoire.

Le traitement en différé permet d'utiliser des algorithmes élaborés que l'on ne sait pas exécuter en temps réel.

L'introduction de l'opérateur humain dans la chaîne de traitement permet de faciliter la réalisation de certaines fonctions ou de suppléer des modélisations incomplètes. En analyse, les interactions peuvent intervenir dans une étape de prétraitement en sélectionnant manuellement certains événements dans la vidéo, en initialisant la localisation des composants corporels ou en effectuant une présegmentation. Elles peuvent aussi intervenir après le traitement pour valider ou corriger des résultats ou pour identifier des erreurs. En synthèse, on peut créer des animations particulièrement complexes à générer automatiquement, à l'aide de techniques d'animation couramment employées dans le domaine des films

d'animation, telles que la capture de mouvements à l'aide de dispositifs dédiés, ou la rotoscopie, qui consiste à copier une animation à partir d'un modèle vidéo à l'aide de logiciels spécialisés.

Il est enfin possible d'inclure avant le traitement, une phase d'apprentissage supervisé du fond, des personnages et de leur manière de signer (traitement spécifique à un locuteur).

#### **4.5. Bilan**

On voit que les deux premiers types de restrictions portent sur la langue elle-même et vont contraindre les types d'énoncés qui peuvent être traités ou générés. C'est à ce niveau qu'il est important pour le chercheur d'être conscient des limitations qu'il introduit et des effets que cela peut avoir sur le développement de la LSF et de ses usages.

Les deux types de restrictions suivants concernent davantage le traitement d'images et influent moins sur la LSF. Cependant, vouloir maîtriser, par exemple, les conditions d'acquisition vidéo ou 3D, ou vouloir quantifier certains phénomènes tels que les mouvements de sourcils implique d'enregistrer des énoncés en LSF produits spécifiquement pour l'étude, ce qui va limiter la variété et la richesse des entrées et risque de produire une langue de laboratoire.

Enfin la dernière contrainte est purement conjoncturelle : elle porte sur l'application et non sur les données et permet de mettre en œuvre, grâce à l'intervention de l'utilisateur, des applications que nous ne sommes pas encore capable d'automatiser, mais sans limiter la langue utilisée.

Quels que soient le niveau et la nature des restrictions introduites, la conception du système doit réaliser un compromis entre :

- la capacité du système à résoudre les problèmes liés aux objectifs de l'application (en analyse : pourcentage de reconnaissance, degré d'automatisation, taux d'erreurs de segmentation temporelle ; en synthèse : capacité d'expression, aspect naturel des gestes...);

- la perte de généralité liée à l'introduction de contraintes et à la simplification des problèmes. Il faut en particulier veiller à ce que le système réalisé soit une version spécialisée d'un système plus générique (que l'on ne sait pas encore ou qu'il n'est pas nécessaire de construire) et non un système *ad hoc* non généralisable qui limiterait la LSF à une forme ou à des capacités d'expression réduites. Cela veut dire, en particulier, que l'information *a priori* introduite dans un programme d'analyse d'une vidéo en LSF, pour améliorer ses performances, doit pouvoir être modifiée pour changer de contexte.

## 5. Les recherches en TALS

Les remarques précédentes pourraient s'appliquer au traitement de toute langue, mais elles prennent ici une acuité particulière pour une double raison.

### 5.1. *Spécificités des langues des signes*

Les LS ont des caractéristiques qui les différencient fortement des langues vocales (Cuxac, 2000 ; Vermeerbergen, 2006 ; Pizzuto, Pietrandrea et Simone, 2007) : elles font appel au canal visuo-gestuel permettant d'exprimer des productions fortement iconiques dans une visée illustrative et elles utilisent l'espace comme substrat pour organiser l'énoncé, en y situant ou en référant les éléments du discours et en mettant en jeu pour cela plusieurs composants corporels en parallèle. Si ces caractéristiques sont réellement présentes dans les corpus étudiés, leur prise en compte fera émerger des modélisations originales et des approches nouvelles d'analyse : une remise en cause de la compositionnalité de bas niveau (plutôt morphémique que phonétique), une modélisation des signes plutôt de formes géométriques, des méthodes d'analyse cherchant à s'affranchir des structures linéaires et des décompositions en arbres, la prise en compte d'une articulation forme-sens à tous les niveaux de l'analyse, etc.

### 5.2. *Absence de forme écrite de la LSF*

L'absence actuelle de forme écrite de la LSF a plusieurs conséquences sur les axes de recherche qui sont explorés en TALS :

- l'existence d'un thème de recherche actif sur les formes graphiques de la LSF (Garcia, 2006). Si des formalismes de notation existent et sont largement utilisés, les bases d'une écriture sont encore en gestation et nécessitent d'explicitier le statut structural des paramètres et leurs relations ;
- une forte activité sur la constitution de corpus de vidéos en LSF, seuls moyens d'entrée dans l'analyse de la langue, et le développement de recherches pour l'exploitation et la représentation de ces corpus ;
- la nécessité de traiter simultanément l'objet linguistique et sa forme externe, généralement un enregistrement vidéo.

En synthèse comme en analyse, le traitement informatique de la LSF comporte toujours une double problématique :

- en synthèse, il faut savoir générer des énoncés syntaxiquement et sémantiquement corrects et être capable de les exprimer en animant un signeur virtuel de manière acceptable par un locuteur de la langue (Braffort, 2007) ;
- en analyse, on devra résoudre un problème de vision par ordinateur (détection, suivi, caractérisation et reconnaissance des éléments constituant les images vidéo : le fond, le personnage et ses composants corporels) et interpréter les configurations

et les mouvements de ces composants en termes langagiers pour comprendre le message (Dalle, 2006).

On peut vouloir aborder séparément ces deux domaines que sont le traitement de la langue et celui de l'image. Il devient donc nécessaire d'introduire une représentation intermédiaire. On est alors fortement tenté d'utiliser un formalisme basé sur le français ou sur une représentation linéaire à partir de ce que l'on appelle « glose », mais l'on risque ainsi de perdre une bonne part de la spécificité des langues des signes (Pizzuto, Rossini et Russo, 2006 ; Antinoro Pizzuto, Chiari et Rossini, à paraître).

## 6. Conclusion

Nous avons présenté dans cet article le contexte dans lequel se trouve le chercheur lorsqu'il entreprend de traiter par ordinateur la LSF. En réalité, nous avons surtout abordé les questions qui se posent en amont et en aval de ce qui relève classiquement du TAL dans les langues vocales.

D'un côté, ce que le chercheur doit savoir du fonctionnement de la LSF et de son évolution actuelle. Notre propos ne prétend pas avoir couvert tous les aspects des LS, mais il souligne un des aspects les plus originaux des langues des signes, qui ne peut donc pas être ignoré.

De l'autre côté, les caractéristiques des données que le chercheur doit prendre en compte ou qu'il doit générer, et les difficultés pour trouver le bon compromis entre une complexité abordable par les méthodes actuelles d'analyse ou de génération et le respect d'une langue ; cette langue ne peut être abordée que sous sa forme parlée, puisqu'elle n'a pas d'écriture, *via* des données vidéo dont le processus d'acquisition ne doit pas biaiser les productions.

Lorsque ces deux aspects ont été pris en compte, se pose alors la question du choix des formalismes de représentation et des outils d'analyses syntaxique et sémantique. Les LS étant des langues à part entière, on devrait pouvoir leur appliquer des méthodes et des hypothèses semblables à celles des autres langues. Mais ces méthodes et les outils qui en découlent ont été élaborés pour des langues qui sont soumises à la contrainte de linéarité propre au canal audiophonatoire des langues vocales, des langues qui n'utilisent pas l'espace 3D pour structurer le discours et qui n'intègrent pas, comme les LS, un processus d'iconicisation de l'expérience et les structures de transfert très complexes qu'il produit. Peut-on alors adapter les cadres théoriques et les méthodes existants et à quel prix ? Sinon, quels nouveaux types d'approches faut-il développer et seront-ils à leur tour adaptés aux langues vocales ? On trouvera dans ce numéro des tentatives de réponses selon chacune de ces deux voies.

## 7. Bibliographie

- Antinoro Pizzuto E., Chiari I. & Rossini P., The representation issue and its multifaceted aspects in constructing sign Language corpora : questions, answers, further problems. in *3rd Workshop on the Representation and Processing of Sign Languages* ). LREC 2008, Marrakech, May 27-June 1, 2008, (à paraître).
- Antinoro Pizzuto E., Rossini P., Sallandre M-A, Wilkinson E., Deixis, anaphora and Highly Iconic Structures : Cross-linguistic evidence on American (ASL), French (LSF) and Italian (LIS) Signed Languages. In *TISLR9, 9<sup>th</sup> Theoretical Issues in Sign Language Research Conference*, Florianopolis, Brazil, December 2006,
- BONNAL F., Sémio-genèse de la langue des signes française : étude critique des signes attestés sur support papier depuis le 18<sup>ème</sup> siècle et nouvelles perspectives de dictionnaires, Thèse de doctorat, Université Toulouse Le Mirail, 2005.
- Braffort A., Dalle P., « Sign language applications : preliminary modeling » *Universal Access in the Information Society* Springer-Verlag, numéro spécial *Emerging Technologies for Deaf Accessibility in the Information Society*, vol. hors-série, 2007.
- Cuxac C., Fonctions et structures de l'iconicité des langues des signes, thèse de doctorat d'État, université René Descartes, Paris V, 1996.
- Cuxac C., *La langue des signes française (LSF) ; les voies de l'iconicité, Faits de Langues* 15/16, Ophrys, Paris, 2000.
- Cuxac C., « Phonétique de la LSF : une formalisation problématique », *Actes du Colloque « Linguistique de la LSF : recherches actuelles »*, Université de Lille 3, 2004, p. 23-24.
- Cuxac C., « Langues des signes et gestuelle co-verbale : pour un programme commun de recherches », *Cahiers de Linguistique Analogique*. à paraître.
- Cuxac C. et Sallandre M-A., "Iconicity and arbitrariness in French Sign Language : highly iconic structures, degenerated iconicity and diagrammatic iconicity", E. Pizzuto, P. Pietrandrea & R. Simone (eds), *Verbal and Signed Languages, Comparing Structures, Constructs and Methodologies*, Mouton de Gruyter, Berlin-New York, 2007, p. 13-33.
- Dalle P. "High level models for sign language analysis by a visio system" *Proceedings of the Second Workshop on the Representation and Processing of Sign Languages : 5th International Conference on Language Resources and Evaluation LREC 2006*, 28 mai 2006, Gênes, Italie
- Delaporte Y., *Les Sourds, c'est comme ça. Ethnologie de la surdimutité*, Éditions de la Maison des sciences de l'homme, Paris, 2002.
- Fusellier-Souza I., « La création gestuelle des individus sourds isolés : de l'édification conceptuelle et linguistique à la sémiogénèse des langues des signes », *AILE* 15, 2001, p. 61-95.
- Fusellier-Souza I. Étude de langues des signes primaires pratiquées par des sourds brésiliens, thèse de doctorat, université Paris 8, 2004.
- Garcia B., "The methodological, linguistic and semiological Bases for the elaboration of a written form of LSF (French Sign Language)", C. Vettori (ed). *Proceedings of the Second*

*Workshop on the Representation and Processing of Sign Languages : Lexicographic Matters and Didactic Scenarios, 5th International Conference on Language Resources and Evaluation - LREC-2006.* ILC-CNR, Gênes, Italie.

Goldin-Meadow S., "When does gesture become language ? A study of gesture used as a primary communication system by deaf children of hearing parents", K. Gibson & R. Ingold (eds.), *Tools, language and cognition in human evolution*, Cambridge University Press, Cambridge, Mass., 1991, p. 63-85.

Jirou G., « Description d'une langue des signes informelle en dehors du milieu institutionnel. Analyse lexicale du parler de Mbour (Sénégal) », *Cahiers de Linguistique Analogique*. à paraître.

Lefebvre-Albaret F., Dalle P., « Une approche de segmentation de la langue des signes française », *TALN 2008 Avignon*, 9-13 juin à paraître

Monteillard N., La langue des signes internationale. Aperçu historique et préliminaires à une description, *AILE* 15, 2001, p. 97-112.

Pizzuto E., Pietrandrea P., & Simone, R. (eds.). *Verbal and Signed Languages - Comparing structures, constructs and methodologies*. Mouton De Gruyter, Berlin / New York, 2007.

Pizzuto E., Rossini P., Russo T., "Representing signed languages in written forms : questions that need to be posed", C. Vettori (ed). *Proceedings of the Second Workshop on the Representation and Processing of Sign Languages : Lexicographic Matters and Didactic Scenarios, 5th International Conference on Language Resources and Evaluation - LREC-2006.* ILC-CNR, Gênes, Italie, pp.1-6.

Sallandre M-A., Les unités du discours en langue des signes française. Tentative de catégorisation dans le cadre d'une grammaire de l'iconicité, Thèse de Doctorat en sciences du langage, Université Paris 8, 2003.

Vermeerbergen M. (2006). "Past and current trends in sign language research" *Language & Communication*, 26, 168-192.

Wilcox S., "Gesture and language : Cross-linguistic and historical data from signed languages", *Gesture*, 2004, vol. 4, n° 1, pp. 43-73.

Yau S. C., « Création gestuelle et début du langage — Création de langues gestuelles chez les sourds isolés », *Langages Croisés*, Hong Kong, 1992.