

# Problèmes de représentation de la Langue des Signes Française en vue du traitement automatique

Loïc Kervajan<sup>1,2</sup>

<sup>1</sup>Université de Provence, DELIC

<sup>2</sup>France Télécom R&D/TECH/EASY/LN

loickervajan@wanadoo.fr

## Résumé

Nous proposons dans cet article une description de la Langue des Signes Française dans le but de traduire des énoncés courts du français et de les faire signer par un personnage de synthèse. Cette description pose en préalable la question de la transcription des éléments d'une langue dont le signal n'est pas linéaire. Il s'agit ensuite de repérer les différentes couches linguistiques et la forme de leurs unités constitutives en vue de la répartition des tâches informatiques : la synthèse de gestes nécessite un traitement des éléments constitutifs du geste et la génération syntaxique doit pouvoir manipuler des morphèmes.

**Mots-clés :** langue des Signes Française, description du geste, morphème, kinème, phonème, chèreme.

## Abstract

In this article, we propose a description of the French Sign Language. The aim is to translate short sentences from French and to realize them with a virtual agent. The first question is how to transcribe the elements of a language which the signal is not linear. We propose to observe the different linguistic levels and the shape of their units in order to distinguish the processing levels : gesture synthesis needs to manipulate the constitutive elements of gestures and morpheme units are required in syntactic generation.

**Keywords:** french Sign Language, description of gesture, morpheme, kineme, phoneme, chereme.

## 1. Introduction

Notre projet est le développement d'un système capable de traduire des énoncés courts du français en Langue des Signes Française (LSF) et de les faire signer par un personnage de synthèse. L'enjeu est de permettre un meilleur accès aux médias écrits, aujourd'hui largement développés sur Internet mais qui restent d'un accès difficile pour une partie de la communauté sourde pour laquelle le français reste une langue étrangère mal maîtrisée. Il est également possible d'envisager des interfaces de communication entre personnes sourdes et entendantes pour des services de proximité lors d'un accueil à un guichet ou lors d'un entretien (professionnel, médical, etc.) et des services à distance, habituellement proposés par téléphone. Le travail en cours s'appuie sur les technologies développées au sein de France Télécom R&D : la synthèse des gestes se fait grâce à l'avatar des équipes d'imagerie et le traitement linguistique grâce à TiLT (Traitement Linguistique de Textes) de l'équipe des langues naturelles. Nous commencerons par traiter du niveau morphologique avec les questions posées par la transcription des Langues Signées (LS). Ensuite, la description des éléments constitutifs du geste nous permettra de mieux identifier les couches linguistiques en vue de l'implémentation d'un lexique

informatisé prenant en compte les éventuelles modifications introduites par des règles de grammaire.

## 2. Transcrire la LSF

La première question qui se pose pour le traitement automatique des LS est liée à leur forme non linéaire. Ces langues sont habituellement opposées aux Langues Orales (LO) dont la définition renvoie ici à toute langue dont le canal est audio-vocal, indépendamment de l'existence d'un système d'écriture. Nous avons donc d'un côté les LO dont l'expression est réputée linéaire sous forme d'un signal acoustique variant sur la ligne du temps et les LS dont la mise en forme dans l'espace autorise la simultanéité d'éléments. La difficulté est donc de linéariser ces éléments pour les implémenter dans les bases de données qui nous serviront de dictionnaires lexicaux ou de grammaires.

### 2.1. Les systèmes existants

Observons trois systèmes de transcription autour de l'exemple de **[lire]** (cf. tableau 1). Le premier système de transcription est celui des **dessins**, utilisé par les dictionnaires papier. Les exemples de dessins utilisés tout au long de l'article ont été empruntés au dictionnaire d'IVT (Moody, 1983). C'est le moyen généralement utilisé pour contourner l'absence d'écriture. Mais le geste étant présenté comme un tout, les éléments constitutifs ne sont pas identifiés. De plus, les connaissances relatives aux LS sont présentées à l'aide des LO. Par ailleurs, les LS n'ont pas de dictionnaires propres, monolingues. Tous les dictionnaires de LS sont bilingues et organisés selon l'ordre alphabétique de la LO associée.


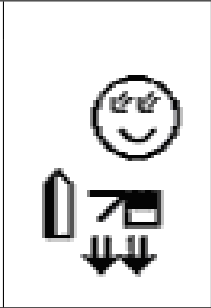

		
Dictionnaire IVT	SignWriting	HamNoSys

Tableau 1. Transcriptions de **[lire]**

Le deuxième système est **SignWriting**, créé en 1974 par (Sutton et Gleaves, 1995). Par rapport à d'autres systèmes de transcription, l'intérêt de cette proposition est qu'elle ne cherche pas à donner une description physique mais à proposer une écriture synthétique intégrant tous les éléments significatifs, tous ceux que l'on voit sur la personne en train de signer : les configurations manuelles mais aussi les gestes non manuels que l'on retrouve sur le visage et dans les attitudes corporelles. Pour atteindre ce niveau d'information, le système garde la même symbolique que la langue exprimée gestuellement et demande une capacité d'interprétation qui ne correspond pas aux contraintes du traitement automatique. Cependant, il est largement utilisé

dans les écoles spécialisées américaines et commence à s'introduire dans d'autres pays comme la France.

Enfin, **HamNoSys**, le système de notation de Hambourg (Prillwitz et Zienert, 1989), propose une codification des paramètres de formation des gestes manuels et non manuels. Très détaillé, il est difficilement utilisable par les personnes pour la prise de notes ou la relecture. En revanche, il est adapté pour la modélisation informatique du geste en vue de sa synthèse automatique puisqu'il propose une transcription linéaire du geste. De plus, l'intérêt de ce système est qu'il a été développé pour décrire le plus fidèlement et le plus économiquement possible les gestes des LS. Il semble cependant que les premières étapes de développement n'aient pas pris en compte les traits non manuels. Bien que l'effort pour les versions ultérieures a porté sur la question, cette dernière reste encore mal définie. Ce système est une source intéressante pour l'identification des unités constitutives du geste. Il met en évidence les deux niveaux de double articulation propre à toute langue et que l'on retrouve également pour les LS (voir point 3.3). Cependant, son utilisation nécessite une formation spécifique que nous pouvons éviter en choisissant d'implémenter les gestes sous un format XML. En effet, même si la gestion de tous les champs à compléter pour XML demande la mise en place d'une interface, l'avantage est l'utilisation du langage naturel pour implémenter les données.

Les trois systèmes présentés et la description informatique proposée ont donc pour vocation de représenter le geste d'un point de vue de la forme. Mais pour pouvoir mettre en relation les différents gestes au sein de la phrase, il devient nécessaire de trouver une notation permettant d'une part de les identifier simplement et d'autre part de transmettre les indications de réalisation lorsque le geste est dépendant d'informations issues de son environnement syntaxique.

## 2.2. Le système des gloses

Dans le cadre de notre développement, nous partons de la représentation sous forme de graphe sémantique. Le graphe contient des étiquettes renvoyant à des entités ou à des relations. La première étape consiste à traduire ces relations sémantiques par les relations syntaxiques de la langue, et à choisir les bonnes formes pour les étiquettes. Pour les LO, ces étiquettes deviennent alors des mots. Pour la LSF, il faut mettre en place une représentation de ces étiquettes qui pourra faire le lien entre les concepts donnés par les étiquettes sémantiques et les gestes. Cette représentation passe par l'utilisation et l'arrangement de gloses. Ces dernières sont habituellement des mots d'une langue utilisés pour expliciter les morphèmes d'une autre langue dans le but d'en comprendre le fonctionnement. Mais la glose peut ici prendre une autre dimension et représenter le geste. Par exemple, [**garçon**] (cf. figure dans le tableau 4) est une glose qui correspond au geste simple signifiant " **garçon** ".

Pour les gestes complexes, l'écriture de la glose doit être complétée d'indications dictées par les règles de grammaires. En effet, ces dernières permettent, en plus de la gestion de l'ordre des gloses, de connaître les éléments à associer aux gloses incomplètes. Par exemple, certaines relations grammaticales sont ancrées dans l'espace. Ainsi, dans le geste **A-[dire]-B** (cf. tableau 2), [**dire**] correspond à un mouvement arqué entre deux sous-espaces devant le locuteur. Ces deux espaces, nommés loci (Vercaingne-Ménard et Pinsonneault, 1996), constituent des références aux actants du verbe. " **A** " est le locus du sujet et " **B** " le locus de l'oblique. Ces deux loci sont respectivement le préfixe et le suffixe du verbe qu'il est nécessaire de placer dans l'espace avant de réaliser le verbe. C'est en ce sens que l'on peut véritablement parler d'ancrage spatial.

Certains verbes de transferts sont construits de la même manière que [**dire**] mais montrent égale-

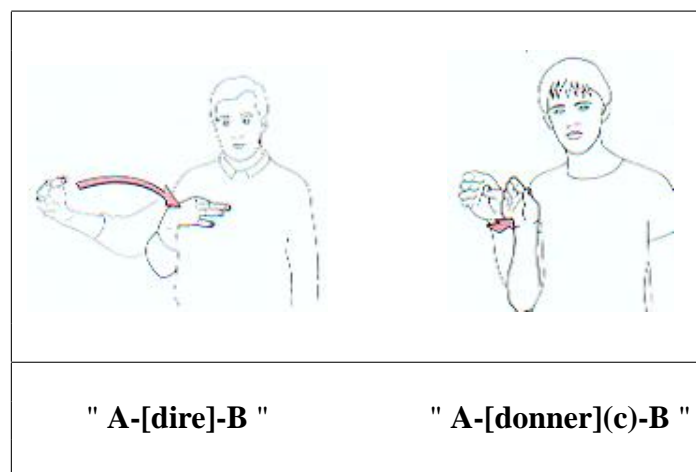


Tableau 2. Ancrage dans l'espace

ment des variations en fonction d'un autre type d'affixe : le transfixe (Kervajan *et al.*, 2005). Cet affixe donne sa forme – une configuration – particulière au verbe en fonction d'une référence à un troisième actant. Sa particularité tient au fait qu'il reste constant du début à la fin du verbe et se superpose à l'ensemble des autres paramètres constitutifs du verbe. Par exemple, **[donner]** (cf. tableau 2) admet comme **[dire]** les deux loci " **A** " et " **B** " et un transfixe renvoyant à l'objet donné. Dans le cas d'un verre, ce dernier pourra être " **c** ". Le geste complet s'écrit alors **A-[donner](c)-B**. D'un point de vue syntaxique, il semble que ces verbes décrivent, avec les affixes, un modèle flexionnel. En effet, la piste du modèle agglutinant semble à écarter dans la mesure où ces affixes ne sont pas des gestes autonomes. Par exemple, le transfixe pris seul n'a pas de valeur lexicale. Il ne s'agit que d'une configuration. De même pour les loci, ce sont des espaces qui n'ont pas de valeur lexicale en soi.

Ces deux exemples montrent que notre utilisation des gloses nous emmène au-delà de la simple notion d'étiquette. En effet, les gloses sont des morphèmes lexicaux comme **-[donner]()**-, **[garçon]** ou grammaticaux tels les affixes comme **A-**, **-B**, **c**. Choisir ce mode de transcription en gloses est intéressant parce qu'il permet de visualiser l'organisation morphologique de la langue. En plus des affixes, d'autres traits pourront également être ajoutés pour apporter des indications de variations physiques de certains éléments constitutifs du geste, comme l'amplitude ou la vitesse qui seront des éléments utilisés lors de la qualification par exemple. La difficulté vient du fait que certains traits sont communs à deux niveaux. Par exemple, l'espace est commun au niveau syntaxique – puisqu'il permet d'organiser certaines relations entre les gestes – et au niveau de la mise en forme – puisque chaque geste se réalise dans l'espace.

Les gloses nous permettent donc de manipuler de manière linéaire le niveau syntaxique. Mais indépendamment du type de glose (variable ou invariable), il est nécessaire de formaliser une description du geste qui doit nous permettre de donner les instructions au programme de synthèse. C'est ce que nous abordons dans la partie suivante.

### 3. Paramètres de formation des gestes

L'objectif est ici d'identifier le niveau d'information nécessaire à la synthèse de gestes. Partons donc de l'observation de la construction physique d'un geste. Ce dernier est constitué de paramètres de formation dont nous proposons une classification sous forme de schéma, figure 1. Il s'agit d'un instantané du geste que l'on peut nommer état. Aussi le geste est-il complet

lorsqu'il contient au moins deux états (Début et Fin) liés par un dernier paramètre : le type de trajectoire.

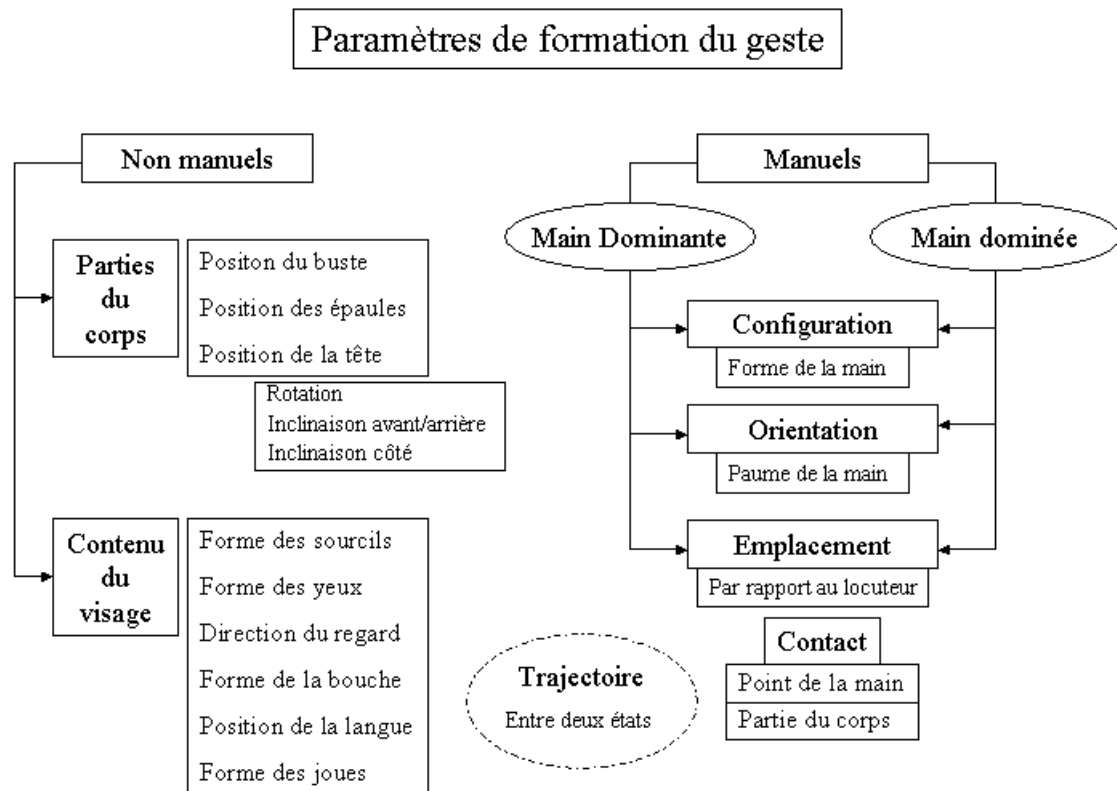


Figure 1. Paramètres de formation du geste

Il existe deux types de paramètres : les paramètres manuels et les paramètres non manuels. L'ensemble de ces paramètres constitue l'ensemble des traits qui forment le geste :

- les configurations manuelles
- les orientations manuelles
- les lieux dans l'espace
- les lieux sur le corps
- les positions des parties du corps
- les traits du visage
- les trajectoires entre deux états

Puis chacune de ces classes contient des éléments que l'on discrimine les uns des autres en fonction de leurs effets sur les gestes. Il s'agit désormais de proposer un formalisme permettant de résoudre la simultanéité des paramètres précédemment exposés.

### 3.1. Exemple à travers un geste simple [garçon]

Le tableau 3 propose, à travers l'exemple de [garçon] (voir l'image dans le tableau 4), une implémentation du geste prenant en compte la simultanéité des paramètres de formation.

<p>&lt;GESTE&gt; garçon</p> <p>&lt;DEBUT&gt;</p> <p>configuration = " un "</p> <p>orientation = horizon vers côté opposé</p> <p>emplacement = front côté dominé</p> <p>contact = oui</p> <p>bouche = " A "</p> <p>buste = rotation ( 0 ), inclinaison ( 0 )</p> <p>épaule = rotation ( 0 ), inclinaison ( 0 )</p> <p>tête = rotation ( 0 ), inclinaison ( 0 )</p> <p>sourcils = neutre</p> <p>forme des yeux = neutre</p> <p>direction du regard = neutre</p> <p>forme des joues = neutre</p> <p>&lt;/DEBUT&gt;</p> <p>&lt;MOUV&gt;</p> <p>trajectoire = rectiligne</p> <p>répétition = 0</p> <p>vitesse = 0.5 sec</p> <p>&lt;/MOUV&gt;</p> <p>&lt;/GESTE&gt;</p>	<p>&lt;FIN&gt;</p> <p>configuration = " un "</p> <p>orientation = horizon vers côté opposé</p> <p>emplacement = front côté Dominant</p> <p>contact = oui</p> <p>bouche = " O "</p> <p>buste = rotation ( 0 ), inclinaison ( 0 )</p> <p>épaule = rotation ( 0 ), inclinaison ( 0 )</p> <p>tête = rotation ( 0 ), inclinaison ( 0 )</p> <p>sourcils = neutre</p> <p>forme des yeux = neutre</p> <p>direction du regard = neutre</p> <p>forme des joues = neutre</p> <p>&lt;/FIN&gt;</p>
---	---

Tableau 3. Exemple d'implémentation d'une glose

Il s'agit là d'un geste simple, c'est-à-dire invariable. Voyons maintenant ce qu'il se passe pour des gestes nécessitant un apport extérieur pour se construire.

### 3.2. Comparaison avec un geste complexe "A-[donner](c)-B "

Le tableau 4 reprend les éléments constitutifs de deux gestes : **[garçon]** et **A-[donner](c)-B**. L'intérêt de la comparaison est de montrer la similitude apparente de deux types de gestes. Le geste **[garçon]** est construit par sept traits : une configuration, une orientation, deux formes de la bouche, deux emplacements et un type de mouvement. Le geste **A-[donner](c)-B** est, lui aussi, construit par les sept mêmes traits. Notons qu'en dépit de ce que suggère l'exemple, tous les gestes n'ont pas sept traits. En effet, la configuration et l'orientation peuvent varier et l'emplacement peut rester le même (donc pas de trajectoire).

Apparemment, **[garçon]** et **A-[donner](c)-B** relèvent de la même construction en ce qui concerne les traits. Mais en réalité, il existe une différence majeure. Dans le premier cas, il n'y a pas de lien entre les traits pris indépendamment et la signification du geste. Dans le deuxième, quatre traits portent une signification et sont donc des indications qu'il faut pouvoir maîtriser en amont de la simple description dédiée à la synthèse :

- l'arc indique la notion de transfert
- le locus " **A-** " représente la personne qui donne
- le locus " **-B** " représente la personne qui reçoit
- la configuration " **c** " est l'anaphore de l'objet donné

La mise en regard de ces deux gestes avait pour but de montrer qu'au niveau de leur réalisation,



 <p><b>[garçon]</b></p>	 <p><b>A-[donner](c)-B</b></p>
<ul style="list-style-type: none"> <li>- configuration début = " un "</li> <li>- orientation début = horizontale vers le côté opposé</li> <li>- emplacement début = front côté dominé</li> <li>- bouche début = " A "</li> <li>- configuration fin = inchangé</li> <li>- orientation fin = inchangée</li> <li>- emplacement fin = front côté Dominant</li> <li>- bouche fin = " O "</li> <li>- mouvement = rectiligne</li> </ul>	<ul style="list-style-type: none"> <li>- configuration début = " c "</li> <li>- orientation début = horizontale, la main dans l'alignement du poignet</li> <li>- emplacement début = locus A</li> <li>- bouche début = " D "</li> <li>- configuration fin = inchangé</li> <li>- orientation fin = inchangée</li> <li>- emplacement fin = locus B</li> <li>- bouche fin = " O "</li> <li>- mouvement = arc (courbure en haut)</li> </ul>

Tableau 4. Comparaison de deux gestes

les types de paramètres en jeu sont identiques. Mais en même temps, nous venons de voir que certains gestes ont un sens indépendant des paramètres (où le lieu reste un lieu, la configuration reste une configuration) alors que d'autres ont un sens lié à certains paramètres (où le lieu prend valeur de locus, la configuration prend valeur de transfixe). Dans ce dernier cas, les traits portant du sens sont de véritables morphèmes dont la forme est spécifiée par les règles de grammaire. Ces morphèmes ne deviennent paramètres de formation du geste que lors de la complétion automatique de ce dernier.

### 3.3. Retour sur la théorie linguistique

Malgré les travaux d'auteurs des siècles précédents, il faut attendre (Stokoe, 1960 ; Stokoe, 1972) pour que les LS soient observées comme des langues à part entière grâce la description selon le principe de la double articulation que (Martinet, 1960) développa pour les LO et attesta pour les LS dans l'introduction à la grammaire de (Nève-De-Mévergnies, 1996).

Le tableau 5, emprunté à (Klima et Bellugi, 1979) pour l'American Sign Language, montre les deux signes **[child]** et **[thing]** ne différant que par l'orientation de la paume. Toutes choses égales par ailleurs (configuration de la main, direction et type de mouvement, expression du visage), elle permet de distinguer deux gestes : **[enfant]** et **[chose]** et ne porte pas de sens en soi. Pour (Stokoe, 1960), l'orientation fait partie des **chérèmes**, unités linguistiques des LS qui correspondent aux phonèmes des LO. L'étymologie du terme renvoie à la main. Or il est désormais admis que le signe linguistique des langues gestuelles comprend l'ensemble des éléments corporels. Par extension, le niveau chérématique des LS correspond au niveau phonologique des



"Child" [enfant]

"Thing" [chose]

Tableau 5. Mise en évidence d'un chérème

LO. Les unités correspondant aux morphèmes sont les **kinèmes**. L'étymologie du terme renvoie au mouvement, ce qui correspond bien à l'idée première du geste comme objet non statique.

### 3.4. De la phonologie à la réalisation technique des gestes

Nous avons vu que les éléments qui se dégagent lors de la description du geste n'ont pas toujours valeur de chérèmes. Si l'on revient sur les LO, les phonèmes sont des objets abstraits. Les objets concrets sont les traits constitutifs des sons émis. Le voisement par exemple est un trait. L'explosion en est un autre, le lieu d'articulation également. Il devrait donc exister deux niveaux semblables pour les LS : un premier où l'on observe des traits et un deuxième où l'on décrit les chérèmes, réunissant un ensemble de traits.

Partons d'un exemple : [**infirmier**] (cf. figure 6). En voici deux descriptions possibles :

- À l'aide d'une phrase :

une croix réalisée avec le pouce (autres doigts fermés) sur le haut du bras.

- À l'aide de coordonnées :

Pour la croix :

Premier segment :

Position de départ

Position d'arrivée

Type de trajectoire (rectiligne)

Deuxième segment :

Position de départ

Position d'arrivée

Type de trajectoire (rectiligne)

Pour la main :

Forme : positions des doigts

Orientation de la paume de la main

Partie réalisant la croix : l'extrémité du pouce

La description " à l'aide d'une phrase " est compréhensible par l'être humain mais n'est pas exploitable en l'état par un programme. La deuxième description est suffisamment fine pour être réalisable automatiquement mais pose le problème de la masse d'informations à implémenter. La solution est d'apprendre au module de synthèse à réaliser des gestes de base qu'il est possible de moduler suivant certains paramètres : lieu, vitesse, amplitude, etc. Pour ce faire, il est nécessaire, à l'instar de ce qui est fait pour établir les phonèmes d'une LO, d'observer les gestes et de repérer, indépendamment du sens, les points communs et les différences. Dans notre exemple le geste élémentaire est la croix réalisée avec le pouce, la main en forme de " 1 " (poing fermé, pouce relevé). Le lieu de réalisation étant un élément permettant de distinguer deux gestes : sur l'épaule il s'agit de [**infirmier**], sur le front de [**infirmier**] (cf. figure 6).

Le lieu émerge donc comme un chérème. D'autres gestes sont réalisés également sur le front et sur l'épaule avec un mouvement de la main identique pour les deux lieux. On peut en déduire



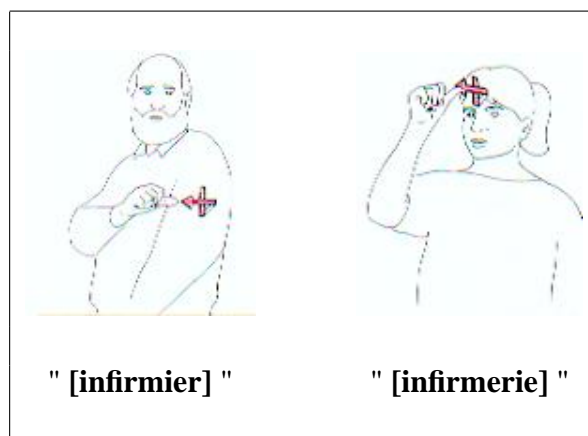


Tableau 6. Identification du chérème.

que ce qui est chérème est " **épaule** " et " **front** ". Le lieu étant un type de chérème comme la consonne est un type de phonème. Ces lieux sont à entendre comme des zones et non comme des lieux très précis : les variations de réalisation des gestes à l'intérieur de ces zones ne changent pas le geste au niveau du sens. Cette notion de zone renvoie au parallèle avec la notion de zone acoustique définie pour un phonème : ce n'est pas une réalisation précise qui définit le phonème, c'est un ensemble de réalisations proches. C'est ainsi que dans la catégorie des lieux, il est possible de définir les zones corporelles suivantes : front, joue, nez, oeil, oreille, menton, cou, épaule, bras, avant-bras, poignet, dos main, paume, doigt, pouce, poitrine, ventre, taille, jambe, dos. Aller au-delà nous entraînerait sur des questions de réalisations physiques des gestes, ce qui correspondrait aux problématiques de phonétique ou d'acoustique pour les LO. Nous laisserons ce travail au logiciel de synthèse de gestes.

## 4. Conclusion

Ainsi, il est possible de formaliser un modèle descriptif qui permet d'une part de passer des graphes sémantiques aux arbres syntaxiques par l'introduction des gloses et d'autre part de mettre en forme les gestes correspondant aux gloses, chaque glose correspondant à une entrée du dictionnaire descriptif dédié à la synthèse de gestes. De plus, notre système descriptif prend en compte d'une part un niveau morphologique – celui des kinèmes – et d'autre part le niveau phonologique – celui des chérèmes. Une des difficultés étant qu'un même paramètre de formation du geste peut, suivant les gestes, être soit un simple chérème, soit un kinème. Mais ce n'est pas une spécificité des langues gestuelles. En effet, remarquons le phonème /o/ du français qui peut, suivant les contextes être un des morphèmes suivants : *au, aulx, eau, haut, ô, oh, os*.

Après cette première étape de formalisation du modèle, il nous faut désormais entreprendre la phase d'implémentation. Nous pourrions alors vérifier nos propositions quant à la transcription du lexique que nous ne pourrions construire qu'après avoir récolté un corpus suffisamment large pour permettre des tests à grande échelle. La confrontation théorie-pratique lors de cette phase de développement nous conduira sans doute à affiner notre modèle.

## Références

KERVAJAN L., GUIMIER-DE-NEEF E. et VÉRONIS J. (2005). *Processing of French Sign Language : verbs agreement*. In S. Gibet et al. (Eds.), *Lectures Notes in Computer Science (LNCS)*, Springer-Verlag, Berlin.

- KLIMA E.-S. et BELLUGI U. (1979). *The signs of language*. Cambridge, Harvard University Press, Londres (Seconde édition).
- MARTINET A. (1960). *Éléments de linguistique générale*. Armand Colin, Paris.
- MOODY B. (1983). *Dictionnaire de LSF, Tome 1 : " Introduction à l'histoire et à la grammaire de la langue des signes "*. Ellipse - IVT, Vincennes, réédition 1998.
- NÈVE-DE-MÉVERGNIES F. X. (1996). *Essai de grammaire de la Langue des Signes Française, Fascicule CCLXXI*. Bibliothèque de la Faculté de Philosophie et Lettres, Université de Liège, Belgique.
- PRILLWITZ S. et ZIENERT H. (1989). *Hamburg Notation System for Sign Language : Development of a sign writing with computer application*. In S. Prillwitz & T. Vollhaber (Eds.) : *Current trends in European Sign Language Research*, Signum, Hamburg, Allemagne.
- STOKOE W.-C. (1960). *Sign Language Structure : an outline of the visual communication*. Studies in linguistics, Occasional Papers n°8, Buffalo, N.Y.
- STOKOE W.-C. (1972). *Semiotics and Human Sign Languages*. Mouton, The Hague.
- SUTTON V. et GLEAVES R. (1995). *SignWriter - The world's first sign language processor*. Deaf Action Committee for SignWriting, La Jolla, CA.
- VERCAINGNE-MÉNARD A. et PINSONNEAULT D. (1996). *L'établissement de la référence en LSQ : les loci spatiaux et digitaux*. In C. Dubuisson & D. Bouchard (Eds.), *Les cahiers scientifiques de l'Acfas, Spécificités de la recherche linguistique sur les langues signées*, 89, Montréal (pp. 61-74).