
Vers un système de traduction automatique : français/langue des signes française (TLF)

Loïc Kervajan* — Emilie Guimier De Neef** — Gaspard Breton**

* *Equipe DELIC*

29, avenue Robert-Schuman, F-13621 Aix-en-Provence Cedex 1
loickervajan@wanadoo.fr

** *France Télécom Recherches et Développement*

2, avenue Pierre-Marzin, F-22300 Lannion
{emilie.guimierdeneef,gaspard.breton}@orange-ftgroup.com

RÉSUMÉ. Notre travail se situe dans le cadre de la traduction automatique du français vers la langue des signes française (LSF) avec synthèse de gestes au moyen d'un signeur virtuel. Nous présentons tout d'abord quelques éléments descriptifs et théoriques de la LSF. Après avoir situé notre travail, nous proposons une formalisation de la partie dite standard du lexique de la LSF et de quelques phénomènes morphosyntaxiques. Nous poursuivons par la présentation du système de traduction automatique à pivot interlingue « TiLT » développé à France Télécom R&D et l'adaptation de son module de génération à la LSF. Nous terminons par présenter la technologie avatar également développée à France Télécom R&D et sa prise en charge de la LSF.

ABSTRACT. We present a work in the framework of machine translation from French to French Sign Language (FSL) with synthesis of gestures through a virtual agent. We first give some descriptive and theoretical elements about the FSL. Then we propose a formalization of the standard part of the FSL lexicon as well as propositions concerning some morpho-syntactic phenomena. Then we present the machine translation system based on interlingua representation system "TiLT" developed at France Telecom R&D and the adaptation of its generation module to FSL. We end with the avatar technology also developed at France Telecom R & D and its FSL treatment.

MOTS-CLÉS : traduction automatique, LSF, modélisation lexicale.

KEYWORDS: machine translation, FSL, lexicon modelisation.

1. Introduction¹

La traduction automatique de textes écrits et la synthèse de gestes par signeur virtuel sont des technologies qui peuvent participer au développement de services favorisant l'accessibilité des personnes sourdes pratiquant la LSF à l'information. L'objectif de nos travaux est de proposer des solutions alternatives à la lecture dans deux types de situations, principalement :

- des situations où la disponibilité systématique d'un interprète humain est difficile (accueil du public en général : tourisme, administrations, hôpitaux, boutiques, etc.) et où des solutions de type bornes interactives peuvent déjà être en place ;
- des situations où l'information est fréquemment mise à jour et ne peut être préalablement enregistrée sous forme de vidéos (presse, météo, catalogues de produits, pages Web, etc.).

Nous présentons dans cet article un prototype de traduction automatique français/LSF basé sur la technologie TiLT² de traduction automatique et une technologie d'avatar signeur, tous deux développés dans les laboratoires de France Télécom Recherche et Développement.

Après avoir exposé les éléments descriptifs et théoriques sur lesquels nous nous appuyons, nous présentons le cadre de notre travail et notre modélisation du lexique et de la syntaxe de la LSF. Le système de traduction automatique TiLT est ensuite présenté ainsi que la brique de génération de la LSF et son interface avec le système de synthèse de gestes. Nous terminons par quelques perspectives sur la suite de ce travail.

2. Quelques éléments descriptifs et théoriques

2.1. *Éléments descriptifs*³

Les signes de la LSF se caractérisent par cinq paramètres (Moody, 1998) : la forme des mains dite « configuration », l'orientation de la paume, le lieu d'articulation, le mouvement et les gestes non manuels (expressions faciales principalement). Cet ensemble de paramètres constitue les articulateurs qui, associés les uns aux autres, permettent la création de signes distinctifs.

L'espace est utilisé pour la structuration du discours : les actants, par exemple, sont positionnés dans l'espace situé devant le locuteur – l'espace de signation – et

1. Nous remercions Maxime Amblard pour sa relecture et ses conseils.

2. TiLT est une plateforme logicielle pour le traitement automatique des langues.

3. Cette partie est loin de décrire exhaustivement les caractéristiques de la LSF mais se borne à une présentation des points pertinents pour les modélisations proposées au paragraphe 4.

référencés par pointage au long du discours. Le sous-espace de signation utilisé pour marquer la référence est traditionnellement appelé locus⁴. Les actions sont instanciées dans cet espace de référence :

AVOCAT[a] JUGE[b] [a]TELEPHONER[b] *l'avocat téléphone au juge*

C'est la réutilisation des loci (a et b) par le verbe qui permet la mise en relation actancielle : le signe du verbe TELEPHONER débute sur le locus référençant l'avocat [a] et finit sur le locus référençant le juge [b]. C'est par ce mécanisme que l'avocat s'interprète comme étant l'agent du prédicat tandis que le juge en est le patient.

Enfin, une configuration manuelle peut être utilisée comme élément de référence à un signe. La valeur anaphorique de la configuration lui donne alors le statut de proforme. Selon le contexte, un même référent pourra être repris par différentes proformes. Par exemple, un être humain (au singulier) pourra être représenté par les proformes suivantes :

- « G »⁵ pour les situations de déplacement, rencontre, positionnements ;
- « V » pour les situations de nage, de chute ;
- « L » pour la marche, etc.

2.2. *Éléments théoriques*

(Stokoe, 1960) décrit pour la première fois les langues des signes selon le principe de la double articulation proposé par (Martinet, 1955) pour les langues vocales. La formation des énoncés en langue des signes s'articule sur deux niveaux : un niveau morphologique et un niveau phonologique. Les morphèmes, unités minimales de sens, permettent, par leur combinatoire, la production d'un nombre infini d'énoncés. Ils peuvent être une partie d'un signe (équivalent de l'affixe et/ou des morphèmes dépendants que l'on trouve dans les langues vocales) ou le signe lui-même. Le second niveau correspond à celui de la phonologie des langues vocales et permet l'identification d'unités minimales non significatives mais distinctives, à partir desquelles il sera possible d'opposer des paires minimales de morphèmes. Par exemple MAL-AU-CŒUR s'oppose à CONTENT par la seule expression faciale, les éléments manuels étant strictement identiques (mouvement circulaire, main à plat, sur le ventre), INFIRMERIE et INFIRMIER ont en commun une croix réalisée avec le pouce mais ancrée sur des lieux différents du corps : le premier sur le front le deuxième sur l'épaule. Cette approche est à la base de nombreuses propositions de transcription – ou écritures – des langues des signes notamment le système de notation phonétique HamNoSys développé à Hambourg (Prillwitz et al., 1989) ou encore SignWriting (Sutton & Gleaves, 1995). Elle est également à la source des

4. Dans les exemples cités, nous notons les loci par des lettres entre crochets.

5. Dans tout l'article, les lettres majuscules entre guillemets pour un classificateur renvoient aux formes de la dactylogogie, cette dernière étant la transcription des lettres de l'alphabet latin.

travaux de nombreux linguistes des langues des signes à travers le monde, notamment aux États-Unis avec (Liddell, 2003) et (Brentari, 1998), au Canada avec (Miller, 1996) qui a présenté sa thèse de doctorat sur la phonologie de la langue des signes québécoise (LSQ) et avec la grammaire descriptive de la LSQ proposée par (Dubuisson *et al.*, 1999). Cette approche a également des échos en France que l'on retrouve dans les travaux de (Bonucci, 1998), (Kervajan, 2006), etc.⁶

3. Corpus et périmètre de l'étude

Pour cette première expérience de traduction français/LSF, nous avons travaillé dans le domaine médical à partir d'un corpus réel d'interactions entre un pharmacien et un client acquis expérimentalement. Ce corpus est composé de dialogues dont nous avons extrait une centaine d'énoncés simples récurrents tels que :

<i>Mon enfant tousse</i>	<i>Je pense qu'il est malade</i>
<i>Donnez-moi du sirop</i>	<i>Je voudrais de l'aspirine</i>
<i>Hier j'étais malade</i>	<i>Combien coûte ce sirop ?</i>
...	

Ces énoncés ont été soumis pour traduction en LSF à deux informateurs sourds, locuteurs natifs de cette langue et à un interprète professionnel pour l'observation linguistique. Les modélisations proposées dans la suite de ce travail sont directement issues de ces observations. S'agissant d'énoncés simples et factuels et d'une tâche de génération⁷, le travail s'est principalement fait autour des signes standardisés de la LSF⁸ et autour des contraintes d'ordre entre les gestes. La modélisation du lexique, de la morphologie et de la syntaxe de la LSF que nous proposons pose ses bases dans une perspective structurale.

4. Modélisation de la LSF

4.1. Principes de la modélisation lexicale en TAL

Pour générer ou analyser automatiquement un énoncé d'une langue, disposer d'un recensement et d'une caractérisation des mots de cette langue est nécessaire. En traitement automatique des langues, ces connaissances sont généralement stockées dans un dictionnaire ou un lexique. Ces lexiques s'organisent autour d'entrées lexicales, c'est-à-dire de formes de base (autrement nommées lemmes)

6. Cette approche structurale est réinterprétée par (Cuxac, 2000) qui propose un modèle faisant de l'iconicité le principe organisateur des langues signées.

7. Par opposition à une tâche de reconnaissance de la LSF.

8. C'est-à-dire des signes invariables mais aussi des signes dont la réalisation dépend de l'environnement syntaxique proche comme nous le détaillons à partir du paragraphe 4.

auxquelles sont associées des informations morphologiques, syntaxiques et sémantiques (voir par exemple Bouillon *et al.*, 1998).

Le paradigme flexionnel du mot constitue l'information morphologique minimale associée à chaque entrée lexicale. Ce paradigme peut être vu comme un codage du « *processus par lequel les mots reçoivent des affixes qui expriment des informations, sans changer leur catégorie syntaxique* » (Bouillon, 1998, p. 13). Les paradigmes flexionnels permettent non seulement d'identifier les formes existantes dans une langue mais aussi de les opposer entre elles au moyen de descripteurs. Pour l'entrée lexicale *chapeau*, par exemple, les descripteurs NOMBRE/SINGULIER et NOMBRE/PLURIEL distinguent *chapeau* de *chapeaux* :

Lemme	Flexion	Descripteurs
<i>Chapeau</i>	<i>Chapeau</i>	GENRE/MASCULIN NOMBRE/SINGULIER
	<i>Chapeaux</i>	GENRE/MASCULIN NOMBRE/PLURIEL

Tableau 1. *Information flexionnelle du mot chapeau*

Les informations syntaxiques d'une entrée donnée sont principalement la catégorie grammaticale et les cadres de sous-catégorisation. La catégorie grammaticale (NOM, VERBE, ADJECTIF, etc.) est une étiquette qui permet de rassembler et/ou distinguer les mots qui partagent un ensemble de propriétés syntaxiques. Par exemple, les verbes du français à la différence des noms, peuvent être précédés de pronoms sujets : *il mangeait* vs **il chapeau*. La catégorie peut être précisée par une sous-catégorie grammaticale (COMMUN ou PROPRE pour les noms, SUJET ou OBJET pour les pronoms, etc.) pour affiner la partition. Les cadres de sous-catégorisation encodent la structure argumentale des entrées lexicales, c'est-à-dire le nombre des compléments régis et leur type. En français, un verbe intransitif comme *récidiver* s'oppose à un verbe transitif direct comme *lire* ou ditransitif comme *parler*.

Un verbe polysémique comme *voler* sera associé à au moins deux entrées lexicales pour distinguer le sens de *dérober* du sens de *voler dans les airs*. Ces entrées seront liées à deux prédicats sémantiques à la valence et au type différents :

Lemme	Informations syntaxiques	Prédicat sémantique
<i>Voler</i>	DITRANSITIF	ACTION.STEAL(<i>agent, theme, patient</i>)
<i>Voler</i>	INTRANSITIF	ACTION.FLY(<i>agent</i>)

Tableau 2. *Information sémantique de voler*

Dans la suite, nous détaillons la modélisation de la morphologie proposée pour la LSF et le système de transcription sous forme de gloses mis en place. Nous exposons également quelques éléments de syntaxe. Les aspects sémantiques de la modélisation, non spécifiques à la LSF, ne sont pas exposés ici.

4.2. Les entrées verbales en LSF

On oppose traditionnellement en langues des signes deux classes de verbes : les verbes directionnels et les verbes non directionnels. Les verbes directionnels, au contraire des verbes non directionnels, ont pour particularité de voir leur point de départ et d'arrivée modifié en fonction des actants, à l'instar du verbe TELEPHONER (cf. 2.1.2).

Nous avons proposé dans (Kervajan *et al.*, 2006) d'intégrer la variation en configuration manuelle, c'est-à-dire en proforme, à cette typologie. Par exemple, un verbe comme DONNER varie non seulement en locus de départ et en locus d'arrivée en fonction du donneur et du receveur, mais également en proforme en fonction du type d'objet donné :

– ZOE[a] TOM[b] LIVRE{conf/bco⁹}¹⁰ [a]DONNER(bco)¹¹[b]
Zoe donne un livre à Tom

La proforme donne sa forme – sa configuration – au verbe en fonction d'une référence à un actant. Sa particularité tient au fait qu'elle reste constante du début à la fin du verbe et se superpose à l'ensemble des autres paramètres constitutifs du signe. La proforme peut donc être vue comme un morphème marquant l'accord entre le verbe et l'un de ses arguments. Nous avons proposé de dénommer transfixe ces proformes qui se superposent au signe verbal. Une représentation des mécanismes d'accord pour un verbe comme DONNER est proposée figure 1 où la flèche représente la contrainte d'accord entre le signe LIVRE et le signe DONNER.

9. « bc » signifie « bec de canard » ; « bco » signifie « bec de canard ouvert ».

« conf » signifie « configuration manuelle ».

10. L'utilisation des accolades renvoie à la notion d'ensemble : un nom peut être repris par différentes proformes en fonction du contexte. Par souci de clarté nous ne donnons que la forme utilisée par le verbe.

11. L'utilisation des parenthèses indique la réalisation effective de la proforme dans le verbe.

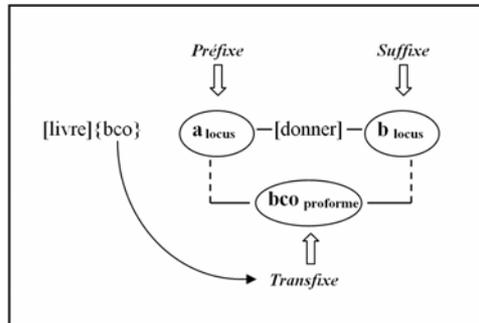


Figure 1. *L'accord verbal selon (Kervajan et al., 2006)*

La combinaison des mécanismes de variation en proforme et en loci donne finalement plusieurs classes de verbes pour la LSF dont nous détaillons ci-dessous les choix de modélisation dans notre lexique.

4.2.1. *Verbes invariables*

Les verbes invariables sont les verbes dont les paramètres de lieux de réalisation et de configuration sont fixes. Avec ces verbes, les actants sont spécifiés par des gestes autonomes distincts du geste verbal.

– TU PENSER *Tu penses*

La modélisation de ces verbes dans notre lexique se réduit à une glose associée à une étiquette catégorielle. Le paradigme flexionnel ne comporte qu'une forme :

Lemme	Flexion	Descripteurs
Penser	Penser	VERBE

Tableau 3. *Modèle flexionnel du verbe PENSER*

4.2.2. *Verbes variables en locus*

Certains verbes admettent des variations des paramètres de lieux de réalisation de début et/ou de fin de signe. Ces verbes sont variables en loci. La variation peut s'effectuer sur un locus ou deux loci. Seul un locus est modifié dans le cas d'un verbe comme ARROSER :

(1) PLANTE[c] ARROSER[c] *J'arrose une plante*

tandis que les loci de départ et d'arrivée sont modifiés pour un signe comme DIRE :

(2) AVOCAT[a] JUGE[b] [a]DIRE[b] *L'avocat dit au juge*

On remarque également que cette modification en locus peut intervenir en début et/ou en fin de geste comme pour le verbe DIRE ou bien du début jusqu'à la fin de la réalisation du signe comme dans le cas d'ARROSER.

D'un point de vue thématique, le locus variable peut correspondre à l'agent du prédicat, au destinataire ou au thème. L'agent et le destinataire sont affectés dans le cas du verbe DIRE alors que c'est uniquement le thème pour le verbe ARROSER. En conséquence, selon les verbes, les loci renvoient à des pronoms référencés conventionnellement (correspondant en français à « je », « tu » et « il ») ou à des pronoms référencés anaphoriquement (correspondant en français à « celui-ci », « celui-là », « ce », etc.).

Dans notre lexique, nous avons choisi d'enrichir la transcription pour visualiser ces variations. Elles prennent la forme de un à trois suffixes séparés de la racine par une apostrophe, la racine étant la partie du signe qui ne varie pas. Ils sont conventionnellement¹² organisés de la façon suivante : (1) lieu de départ, (2) lieu d'arrivée, (3) proforme. Les lieux de départ et d'arrivée s'opposent simplement deux à deux pour la construction d'un signe en ce qu'ils sont pour le premier, le départ et pour le second, l'arrivée. Pour la construction d'énoncés simples, il n'est pas nécessaire de gérer plus de deux loci cooccurrents¹³.

Par exemple, pour traduire « je te dis que (...) » et « il lui dit (...) », en tant que phrases finies, c'est-à-dire sans énonciation des référents personnels, nous utilisons la convention suivante autour des pronoms : le locuteur est l'agent, l'interlocuteur est en face de l'agent virtuel, comme s'il s'adressait à un public derrière une caméra. Nous utilisons les pronoms standard « je » et « tu ». Par extension, les tiers anaphoriques sont à droite et à gauche du signeur. Nous les avons nommés « il1 » et « il2 ». Notre transcription LSF donne : « dire'je'tu » et « dire'il1'il2 » ou « dire'il2'il1 ».

Lorsque les actants sont préalablement référencés, comme dans les exemples (1) et (2) précédemment vus, nous opposons trois espaces nommés a, b et c. Les deux premiers sont utilisés pour une opposition deux à deux (exemple (2)) et le dernier une localisation centrée s'il n'y a pas d'opposition (exemple (1)). Les transcriptions sont respectivement¹⁴ : « avocat'a juge'b dire'a'b » et « plante'c arroser'c ».

12. Cette convention est purement interne et technique. Elle n'a pas vocation à être diffusée comme modèle de transcription linguistique.

13. En effet, la problématique de la cooccurrence d'un nombre important de loci, référant chacun à une entité différenciée, relève de l'organisation globale du discours. Nous ne dépassons pas, pour l'heure, la capacité à gérer les oppositions relatives à un seul verbe, donc à deux loci cooccurrents. Si un troisième référent apparaît, l'espace de signation doit être reconstruit en fonction des nouvelles oppositions mises en jeu par le nouveau verbe concerné.

14. Il est également possible d'opérer la référenciation par pointeur (manuel ou visuel) : « avocat pointeur'a juge pointeur'b dire'a'b » et « plante pointeur'c arroser'c ».

Une partie des trente entrées du modèle flexionnel d'un verbe comme DIRE avec ses descripteurs et des traductions françaises littérales est donné dans le tableau 4.

Flexion	Descripteurs	Traduction littérale
dire'je'a	VERBE-AGENT/JE-DEST/A	'je dis à celui-ci'
dire'je'b	VERBE-AGENT/JE-DEST/B	'je dis à celui-là'
dire'je'je	VERBE-AGENT/JE-DEST/JE	'je me dis'
dire'je'tu	VERBE-AGENT/JE-DEST/TU	'je te dis'
dire'je'il1	VERBE-AGENT/JE-DEST/IL1	'je lui dis'
dire'je'il2	VERBE-AGENT/JE-DEST/IL2	'je lui dis'
dire'tu'a	VERBE-AGENT/TU-DEST/A	'tu dis à celui-ci'
...
dire'il1'a	VERBE-AGENT/IL1-DEST/A	'il dit à celui-ci'
...

Tableau 4. *Modèle flexionnel partiel du verbe DIRE*

4.2.3. Verbes variables en proforme

Certains verbes admettent des variations en proforme où le choix de la proforme peut être dicté par un agent :

- VOITURE{conf/mp¹⁵} AVANCER(mph) *La voiture avance*
- HOMME{conf/G} SE_DEPLACER(G) *L'homme se déplace*
- HOMME{conf/L} MARCHER(L) *L'homme marche*

Le choix de la proforme peut également être dicté par un thème¹⁶ :

- PIECE_DE_MONNAIE{conf/pince} SAISIR(pince) *Je saisis la pièce de monnaie*

La variation peut également être en loci et en proforme. Donnons l'exemple le plus complexe à travers le verbe DONNER qui peut à la fois modifier sa proforme et les lieux de réalisation de début et de fin :

- TOM[b] ZOE[a] LIVRE{conf/bco} [a]DONNER(bco)[b] *Zoe donne un livre à Tom*

Le modèle complet pour les verbes comme DONNER sera la combinatoire de l'ensemble du modèle des verbes comme DIRE avec l'ensemble des proformes qu'il peut accepter. Nous avons fait le choix d'un modèle flexionnel prédictible, basé pour le verbe DONNER sur un certain nombre de proformes dont la liste doit être connue et soigneusement énumérée. Le système prévoit qu'en cas d'absence d'une proforme

15. « mph » signifie « main plate horizontale ».

16. Le thème est le terme générique utilisé par le système automatique correspondant en français à l'objet direct.

attendue pour marquer un accord, c'est la proforme neutre qui est utilisée. Le tableau 5 donne un aperçu partiel de ce modèle, construit sur le paradigme du lieu de départ « je » et du lieu d'arrivé « tu » avec quatre exemples de proformes possibles qui n'épuisent pas la liste des proformes autorisées avec le verbe DONNER.

Flexion	Descripteurs	Traduction littérale
donner'je'tu'bco	VERBE-AGENT/JE-DEST/TU-PROF_THEME/BCO	'je te donne quelque chose de type « bco »'
donner'je'tu'bc	VERBE-AGENT/JE-DEST/TU-PROF_THEME/BC	'je te donne quelque chose de type « bc »'
donner'je'tu'pince	VERBE-AGENT/JE-DEST/TU-PROF_THEME/PINCE	'je te donne quelque chose de type « pince »'
donner'je'tu'c	VERBE-AGENT/JE-DEST/TU-PROF_THEME/C	'je te donne quelque chose de type « C »'
...

Tableau 5. *Modèle flexionnel partiel du verbe DONNER*

4.3. Les entrées nominales en LSF

Au contraire des verbes, la configuration manuelle fait partie intégrante des signes nominaux : une différence de configuration définit deux entrées lexicales spécifiques (par exemple, MARSEILLE et BEURRE se réalisent de la même manière sauf pour la configuration de la main dominante : en M pour le premier, en N pour le second).

La mise en relation des prédicats et des arguments, comme nous l'avons vu au paragraphe précédent, passe par l'utilisation des loci qui jouent le rôle de référents spatiaux. L'activation des loci peut se faire par transfert, c'est-à-dire en délocalisant un signe par rapport à son espace de réalisation neutre. Cette délocalisation passe par un déplacement des épaules ou par simple déplacement des mains sur le locus visé. L'activation peut également se faire par ajout d'un pointeur manuel ou visuel vers le locus visé. Enfin, cette localisation peut également se faire par positionnement d'une proforme représentant le nom sur le locus visé. Le tableau 6 reprend ces éléments du point de vue de la formalisation que nous en avons faite.

Localisation par déplacement (épaules ou mains)	avocat'a
Localisation par pointeur (manuel ou visuel)	avocat pointeur'a
Localisation par utilisation d'une proforme	avocat G'a

Tableau 6. *Mise en relation d'un locus (a) avec le nom qu'il référence*

La modélisation lexicale retenue prend donc en compte les différents loci possibles¹⁷ pour les noms. Les proformes de rappel disponibles pour chaque nom et activées lors de l'accord sont encodées sous forme de descripteurs. Le tableau 7 présente le modèle flexionnel du nom AVOCAT, pour lequel un sous-ensemble de la liste des proformes est encodé : « G » pour les verbes comme SE_DEPLACER, « N » pour le verbe TOMBER, etc.

Flexion	Descripteurs	Traduction littérale
avocat	NOM-LOCUS/neutre-PROF/G/N/S/...	'un avocat'
avocat'a	NOM-LOCUS/A-PROF/G/N/S/...	'l'avocat qui est sur « a »'
avocat'b	NOM-LOCUS/B-PROF/G/N/S/...	'l'avocat qui est sur « b »'
avocat'c	NOM-LOCUS/C-PROF/G/N/S/...	'l'avocat qui est sur « c »'

Tableau 7. *Modèle flexionnel du nom AVOCAT*

Pour les pointeurs manuels¹⁸, la modélisation retenue est présentée dans le tableau 8.

Flexion	Descripteurs	Traduction littérale
pointeur'a	DET-LOCUS/A	'là (sur « a »)'
pointeur'b	DET-LOCUS/B	'là (sur « b »)'
pointeur'c	DET-LOCUS/C	'là (sur « c »)'

Tableau 8. *Modèle flexionnel du pointeur*

Par extension, le modèle flexionnel d'une proforme utilisée comme déterminant est présenté tableau 9.

Flexion	Descripteurs	Traduction littérale
G'a	DET-LOCUS/A-PROF/g	'lui (sur « a »)'
G'b	DET-LOCUS/B-PROF/g	'lui (sur « b »)'
G'c	DET-LOCUS/C-PROF/g	'lui (sur « c »)'

Tableau 9. *Modèle flexionnel du déterminant proforme G*

17. Rappel : dans notre modèle trois loci sont utilisés : a, b et c. À la variation sur ces loci se rajoute la forme de base, appelée neutre.

18. Nous traiterons le déplacement (transfert) sur un locus particulier au moment où nous traiterons ce phénomène pour les verbes.

4.4. *Éléments de syntaxe*

L'ordre des gestes relevé dans les langues des signes est sujet à controverse. Même si de nombreuses combinaisons sont possibles, avec des conséquences au niveau du sens, un certain nombre d'entre elles est préféré. Nous nous basons sur les propositions de (Moody, 1998, p. 77 et 77) pour générer l'ordre des signes : la structure globale de l'énoncé s'organise suivant les principes de l'ordre thème/rhème (Moody, 1998, p. 123). Nous retenons également que le temps et le lieu de l'action précèdent l'action, le temps étant énoncé avant le lieu, et que le temps de durée suit l'action. La qualification suit l'objet qualifié. Par ailleurs, un certain nombre de règles se dégagent également à la lecture des sous-sections précédentes : 1) la détermination par un pointeur suit l'objet qualifié ; 2) les actants nécessaires à la construction dynamique d'un verbe (loci et/ou proforme) précèdent le verbe.

L'ensemble de ces contraintes est modélisable dans une grammaire formelle utilisable dans une chaîne de TAL. Ces contraintes, ainsi que les différentes formes d'accord exposées, ont été implémentées dans le formalisme des grammaires de dépendance utilisé à France Télécom R&D.

5. Traduire automatiquement du français en LSF

5.1. *Présentation du système de traduction automatique TiLT*

Le logiciel de traduction automatique TiLT est un système avec pivot sémantique interlingue basé sur des règles (Iheddadene, 2006). La phase d'analyse se fait de façon classique : après une étape de segmentation, les informations morphologiques, syntaxiques et sémantiques de chacun des mots de l'énoncé source sont recherchées dans un lexique. Une grammaire de dépendance est ensuite utilisée pour construire le(s) arbre(s) représentant la structure syntaxique de l'énoncé. L'arbre obtenu pour l'analyse de la phrase « j'ai un peu de fièvre » est donné figure 2.

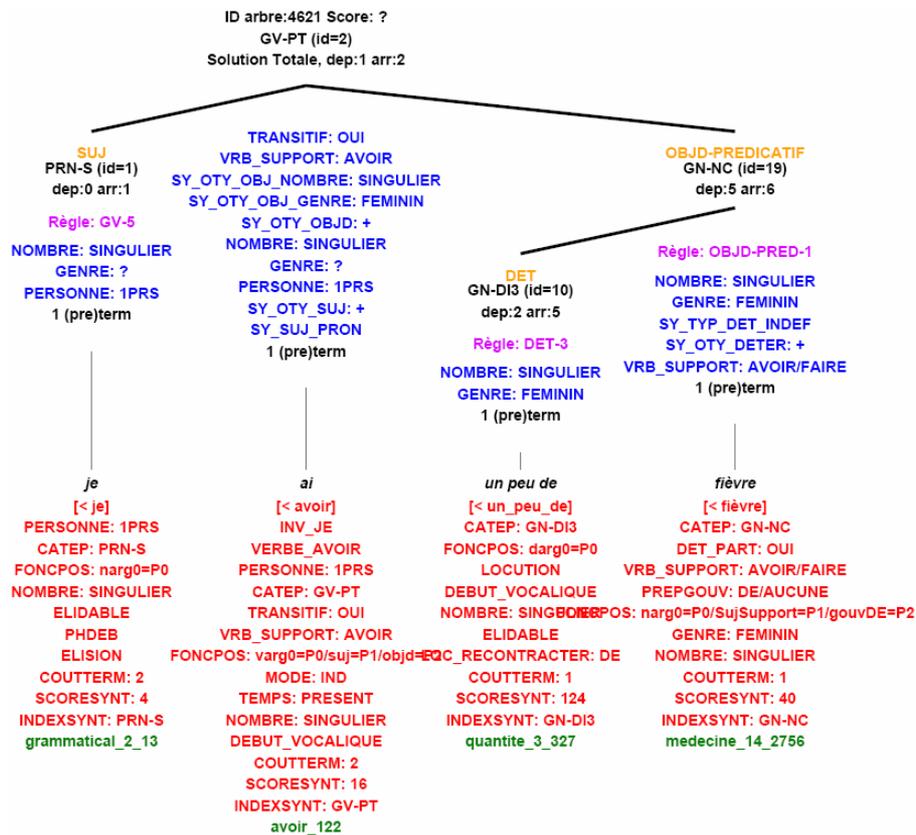


Figure 2. Exemple d'arbre en dépendance pour la phrase « j'ai un peu de fièvre »

L'analyse syntaxique se projette ensuite en un graphe sémantique (figure 3) dont le formalisme est inspiré des graphes conceptuels (Sowa, 1976). Ce graphe est composé d'individus, de prédicats et de relations. Les individus représentent les actants et les situations d'un énoncé ; ils sont décrits par les relations qu'ils entretiennent avec les prédicats. Par exemple, (x6018) est identifiable comme le locuteur par sa relation entity avec le prédicat [ACTANT.locuteur], (u6028) est la fièvre et (x6017) est la situation d'avoir de la fièvre. Le graphe comporte des prédicats lexicaux représentant le sens des unités lexicales de l'énoncé source et des prédicats grammaticaux qui transfèrent le matériel morphologique sémantiquement pertinent pour la traduction comme le nombre, la temporalité, etc.

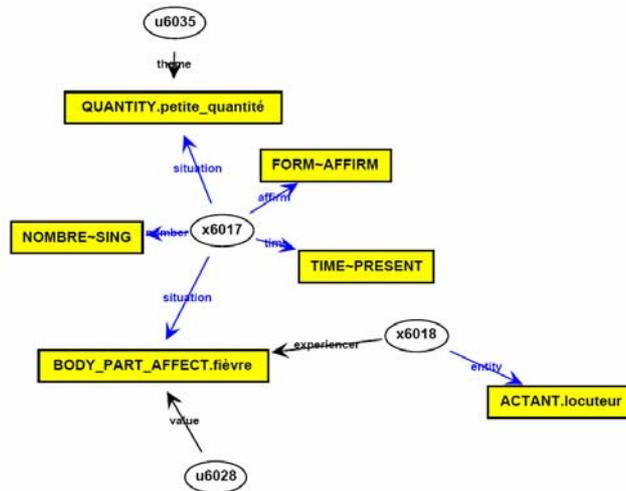


Figure 3. Exemple de graphe sémantique pour la phrase « j'ai un peu de fièvre »

Comme dans tout système de traduction automatique avec pivot sémantique, ces graphes sémantiques jouent efficacement leur rôle de pivot interlingue en cas d'absence d'isomorphie syntaxique entre langue source et langue cible, par exemple, dans le cas (3) où un énoncé impersonnel français est sans équivalent en anglais ou comme dans le cas (4) où une construction à verbe support française devient une construction attributive anglaise. Ces différences syntaxiques sont en effet gommées dans le graphe sémantique. Ainsi, le graphe projeté à partir de l'énoncé source permet-il de générer directement l'énoncé cible.

- (3) a. Il me faudrait de l'aspirine.
 b. I need aspirin.
- (4) a. J'ai des nausées.
 b. I feel sick.

En l'absence d'isomorphie au niveau sémantique entre langue source et langue cible, comme dans les exemples (5), un système de règles de transfert permet de modifier le graphe source pour permettre une génération correcte en langue cible (cf. Park et al., 2007).

- (5) a. J'ai mal au dos.
 b. My back hurts.

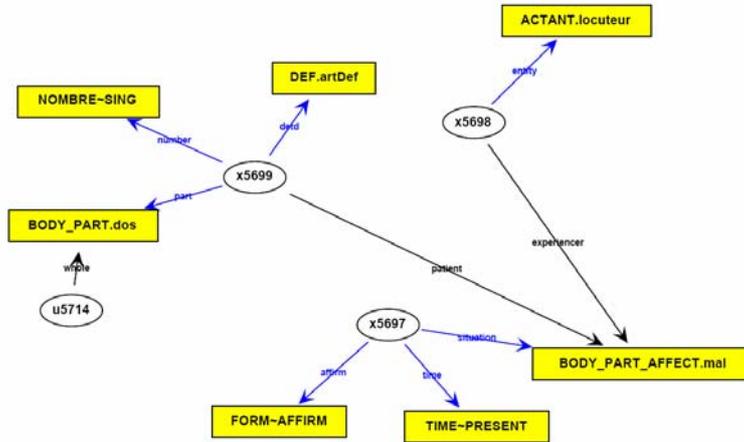


Figure 4. Graphe sémantique initial de l'énoncé « j'ai mal au dos »

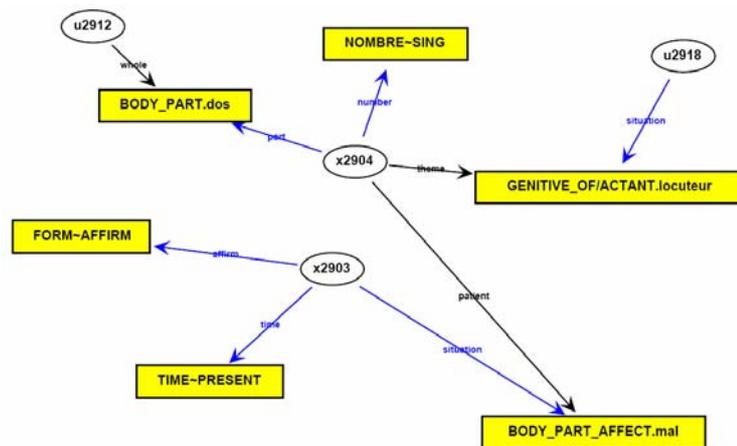


Figure 5. Graphe sémantique transformé de l'énoncé « j'ai mal au dos » pour traduire « my back hurts »

Dans la figure 5, le prédicat [ACTANT.locuteur] a été remplacé par [GENITIVE_OF/ACTANT.locuteur] pour permettre la génération du déterminant possessif *my* au lieu du pronom sujet *I*.

Une fois le graphe sémantique obtenu, la phase de génération enchaîne différentes étapes pour produire une ou plusieurs traductions de l'énoncé source :

- lexémisation : recherche des lexicalisations des prédicats du graphe sémantique dans le lexique cible ;
- arborisation : transformation du graphe sémantique sous forme arborescente ;
- validation syntaxique : filtrage des arborisations en fonction de la grammaire de la langue cible ;
- linéarisation : extraction des traductions et mise en forme éventuelle (gestion des élisions, majuscules, etc.).

Le prototype de traducteur automatique TiLT a été mis au point sur le domaine réduit de l'interaction entre un pharmacien anglophone et un client francophone. Il traduit donc du français vers l'anglais et inversement.

5.2. Génération en LSF

La génération de la LSF avec le système de traduction automatique TiLT se fait à partir du graphe sémantique pivot en utilisant les ressources linguistiques (lexique, grammaire) propres à la LSF. On garantit ainsi l'indépendance complète vis-à-vis de l'énoncé source. La séquence de gestes générée correspond bien à une séquence de LSF et non de français signé. Pour générer du français signé, une approche de première génération est suffisante : l'ordre des mots du français est préservé, seul le lexique est modifié. La traduction automatique de l'exemple « donnez-moi du dentifrice » génère le graphe sémantique figure 6.

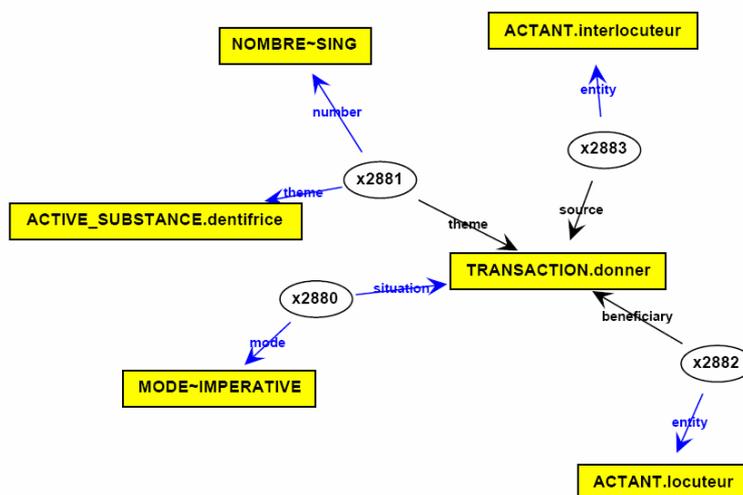


Figure 6. Graphe sémantique de l'énoncé « donnez-moi du dentifrice »

TRANSACTION.donner est la représentation prédicative du verbe *donner*. Sont liés à ce prédicat quatre individus représentant le bénéficiaire [ACTANT.locuteur], l'interlocuteur [ACTANT.interlocuteur], la situation [MODE~IMPERATIVE] et le dentifrice [ACTIVE_SUBSTANCE.dentifrice]. On peut noter que l'interlocuteur émerge de la flexion de *donnez* et ne correspond pas à une unité lexicale autonome dans la phrase source.

Le verbe DONNER en LSF se fléchit en agent, en patient et en proforme. En génération, les prédicats [ACTANT.locuteur] et [ACTANT.interlocuteur] se grammaticalisent donc sur le verbe pour restreindre le paradigme des formes possibles pour le verbe DONNER. La règle d'attachement de la relation THEME, donnée ci-dessous, contraint les verbes de ce type à s'accorder en proforme (c'est-à-dire en configuration) avec leur thème. Cette contrainte sélectionne la proforme « bco » disponible pour un signe comme DENTIFRICE et compatible avec le verbe DONNER. Le schéma de cette règle précise également que le constituant THEME se place avant le verbe (>>). La traduction de l'énoncé cible en LSF donne : « dentifrice donner't'je'bco ».

```
RègleAttachement GV-PROF-CC THEME
Schéma GV-PT-ABC >> GN-NC|GN-NP
ConditionsPrincipal (SY_THEME_GV/!)
AutresConditions((P/CONFIGURATION U D/CONFIGURATION_VERB)
(P += SY_THEME_GV))
```

Figure 7. Règle de grammaire pour l'attachement du THEME

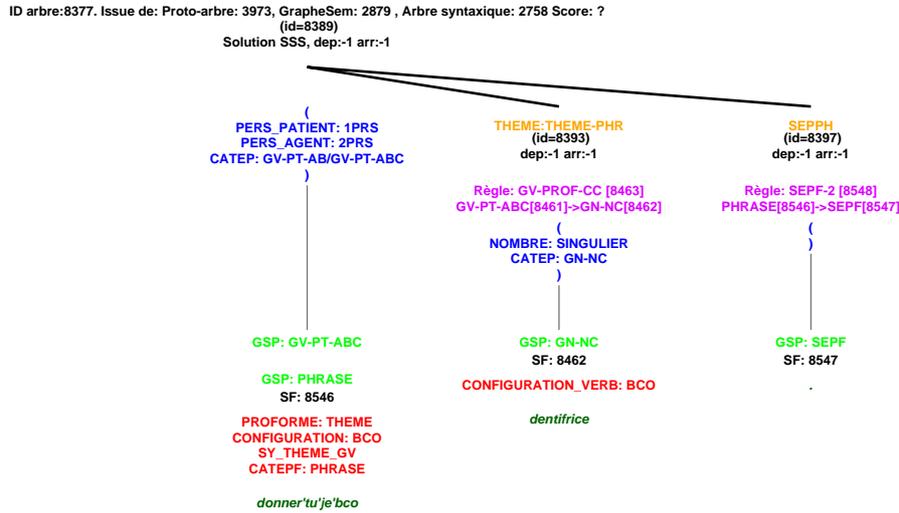


Figure 8. Arbre syntaxique généré pour « dentifrice donner'tu'je'bco »

L'arbre syntaxique (figure 8) généré permet de linéariser correctement l'énoncé en LSF.

5.3. Format d'interfaçage

Nous avons utilisé, section 4, une transcription nous permettant d'identifier la racine d'un signe et ses éléments flexionnels manuels. Pour être lisible par le moteur de synthèse de gestes, cette transcription est convertie au format XML. Pour un exemple comme « je te donne un livre », le moteur de traduction génère en LSF « livre donner'je'tu'bco ». La sortie XML correspondante est la suivante :

```
<SYNTHESE_SIGNES>
  <PARA COD="ISO-8859-1" ISO="fr" LAN="francais" FOR="Fr">
    <PHR TXT="je te donne u livre.">
      <TEXTE_SOURCE LAN="francais">je te donne un livre.</TEXTE_SOURCE>
      <TEXTE_CIBLE LAN="sgn-FR">livre donner'je'tu'bco.</TEXTE_CIBLE>
      <SIGNE>
        <ID="livre"/>
        <TYP=NOTVAR/>
      </SIGNE>
      <SIGNE>
        <ID="donner"/>
        <TYP=VAR/>
        <TRAITS=""/>
        <LIEU_DEP="1PRS"/>
        <LIEU_ARR="2PRS"/>
        <CONFIGURATION="bco"/>
      </SIGNE>
    </PHR>
  </PARA>
</SYNTHESE_SIGNES>
```

Pour réaliser « livre », considéré comme invariable (<TYP=NOTVAR/>), le module de synthèse de gestes pourra récupérer le signe directement dans sa base de données et le reproduire à l'identique.

« donner'je'tu'bco » comporte trois éléments renseignés : CONFIGURATION="bco", LIEU_DEP="1PRS" LIEU_ARR="2PRS" les lieux de début et de fin permettant au module de synthèse de récupérer le geste donner avec les bons loci de départ et d'arrivée et la configuration manuelle adaptée.

6. Le moteur de synthèse de gestes

Une fois la séquence à signer traduite et transcrite au format XML, le module de synthèse de gestes interprète cette transcription pour constituer l'animation finale. Ce processus est décrit dans la figure 9. Pour effectuer l'interprétation, il faut d'abord traduire la séquence dans une structure qui permette de faire le lien avec les paramètres d'animation. Une fois ces paramètres obtenus, il suffit de les faire jouer par le système d'animation pour que le personnage produise la séquence de signes. Notre approche est similaire à celle de (Lee & Kunii, 1993) et de (Lebourque & Gibet, 1999) dans le sens où notre objectif est de développer un procédé qui permette de traduire le langage naturel¹⁹ en langue des signes. Cependant, comme exposé au paragraphe 5, nous utilisons ici une approche interlingue et non un système de transfert comme (Lee & Kunii, 1993).

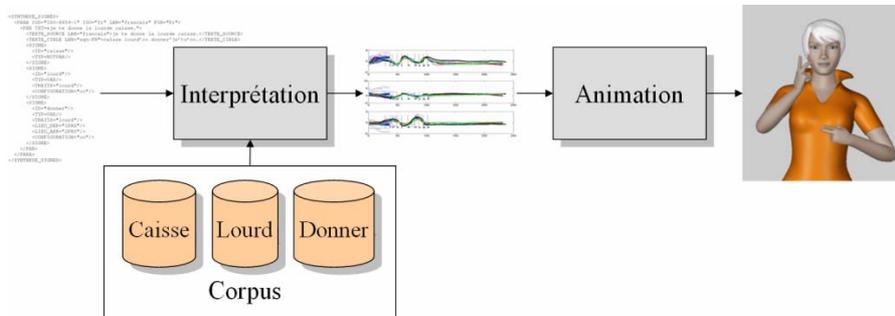


Figure 9. Schéma global

De plus, nous ne nous attachons pas à réaliser un langage de description de gestes comme (Prillwitz *et al.*, 1989 ; Lee & Kunii, 1993 ; Elliott *et al.*, 2004) ou encore (Losson & Vannobel, 1998). Dans cette étude, nous partons du principe que nous disposons d'un corpus qui contient l'intégralité du lexique de signes dont nous avons besoin. L'interprétation consiste alors seulement à composer ou à concaténer des fragments du corpus. Les gestes unitaires contenus dans le corpus sont le résultat d'un travail d'infographie et sont stockés sous la forme de fichiers BVH (BioVision Hierarchical data).

19. Langage naturel s'oppose ici à langage machine. Dans le cadre de notre développement, il s'agit du français.

6.1. Structure de la séquence

Le système d'interprétation reçoit en entrée la description de la séquence au format XML. La séquence est lue en mémoire et est stockée sous la forme d'un arbre.

Chaque signe est constitué d'un identifiant ID et d'un type TYP. L'identifiant sert à faire le lien avec la séquence correspondante dans le corpus. Le type, NOTVAR ou VAR, indique que le geste est invariant ou non. Un geste invariant doit être réalisé tel quel sans aucune modification sous peine d'en changer la signification. À l'inverse, un geste pouvant varier sera adapté en fonction du contexte.

Au niveau corporel, le langage de représentation travaille sur deux canaux différents. Le premier canal représente la configuration de la main dans son intégralité, y compris l'orientation du poignet. Le deuxième canal code seulement la position du poignet. L'attribut CONFIGURATION affecte directement le premier canal et indique que, pour ce geste, la configuration par défaut doit être remplacée. Le deuxième canal est modifié par l'attribut TRAIT par *Motion Blending* (Bruderlin & Williams, 1995). L'animation faciale fait l'objet d'un troisième canal que nous ne décrivons pas ici. Dans ce premier prototype, la direction du regard n'est pas prise en compte. C'est un manque qu'il nous faudra pallier par la suite.

6.2. Instanciation des trajectoires

L'instanciation est le procédé qui construit les trajectoires des paramètres d'animation. Elle transforme la structure discrète que constitue l'arbre construit précédemment en une séquence continue de paramètres. Ce procédé insère les transitions manquantes, réalise la superposition spatiale de séquences, et finalement concatène les paramètres d'animation (figure 10).

L'instanciation commence par la lecture du fichier XML. Chaque geste est inséré dans l'arbre sous le nœud racine *Séquence*. Les transitions *Début* et *Fin* sont ensuite ajoutées, ainsi que les transitions intermédiaires entre les gestes.

Le système parcourt l'arbre une première fois pour instancier les gestes ainsi que les postures initiale et finale, puis une seconde fois pour instancier les transitions. Lors du premier passage, les gestes invariants du corpus sont instanciés tels quels sur les deux canaux. Pour les gestes variables, les positions de début et de fin sont déterminées suivant les attributs. La trajectoire est calculée par *Motion Warping* (Witkin & Popovic, 1995), technique qui permet d'infléchir des mouvements préenregistrés sous contrainte, puis la configuration est modifiée s'il y a lieu. Enfin, le mouvement résultant est altéré par *Motion Blending* ou « mélange de mouvements », suivant le trait. Lors du second passage, les transitions sont

calculées, à partir des paramètres finaux du geste précédant et initiaux du geste suivant, très classiquement par l'intermédiaire de *splines* cubiques.

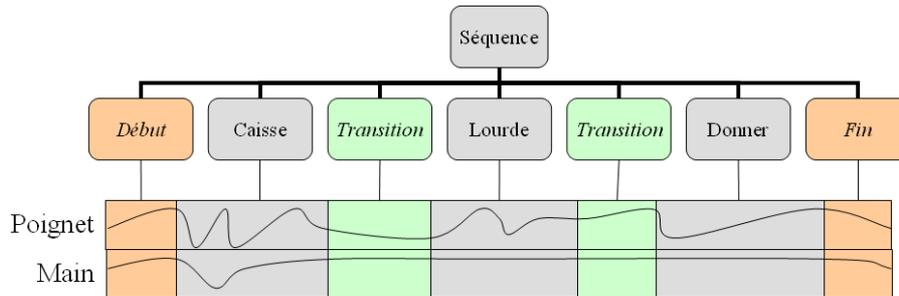


Figure 10. *Instanciation de la séquence*

Dans le calcul des paramètres d'animation, la position du poignet est fixée par les segments du corpus ; en revanche la position du coude est calculée par cinématique inverse (Tolani *et al.*, 2000). La configuration de la main est calculée en cinématique directe et les configurations intermédiaires sont réalisées par interpolation de quaternions²⁰. Cette méthode s'est avérée efficace avec les séquences sur lesquelles nous travaillons mais nécessiterait une validation à plus large échelle.

6.3. Animation

L'animation de l'agent est réalisée de façon tout à fait classique par l'utilisation d'un système *Skin and Bones*. Les *Bones* sont un ensemble de segments qui forment le squelette de l'humanoïde. L'animation est obtenue en modifiant les valeurs angulaires des articulations entre les segments, paramètres que nous avons calculés précédemment. La figure 11 montre une vue du système de *Bones* que nous utilisons sur l'agent *Diane* développé à France Télécom en partenariat avec l'entreprise *Polymorph Software*. Les sommets du maillage formant la peau, *Skin*, sont ensuite animés en suivant les mouvements des *Bones* selon un ensemble complexe de pondérations.

20. Un quaternion est un quadruplet de nombres réels, le premier élément étant un « scalaire » et les trois éléments restants formant un « vecteur » ou « imaginaire pur ». Ils rendent possible la représentation des rotations et notamment une chaîne de rotations, en trois dimensions.



Figure 11. *Structure de l'humanoïde*

Enfin, la figure 12 donne trois vignettes représentant l'exemple du signal obtenu pour l'énoncé en français « il me donne un livre », traduit par « livre donner'il'me'bco ».



Figure 12. *Visualisation de « livre donner'il'me'bco »*

7. Conclusion et perspectives

Nous avons proposé de modéliser les variations des signes que l'on observe en LSF autour de la notion de flexion pour ce qui concerne les variations de configurations et de loci. Les conséquences de cette modélisation sur les entrées nominales et verbales de notre lexique de la LSF pour le TAL ont été exposées ainsi que le mécanisme de l'accord pour garantir la propagation de ces contraintes en syntaxe.

La question de la modification verbale ou nominale – variation des éléments non manuels tels que l’amplitude, la répétition, etc., modification des traits du visage pour marquer l’affirmation, la question, etc. – n’a pas été abordée dans ce travail et constitue l’un des axes de recherche pour sa poursuite. Cette réflexion intègrera également les gestes non manuels, en particulier le regard, dont la prise en compte est indispensable. Sur cette question, des études sur corpus devront être menées pour poursuivre les travaux sur la description linguistique de la LSF.

Le prototype de traduction automatique et la brique de synthèse de gestes sont aujourd’hui intégrés dans une maquette avec une couverture de l’ordre de quatre-vingts gestes. Une évaluation de la qualité de restitution n’a pas encore été faite, mais les premiers retours des personnes sourdes sollicitées laissent à penser que le signeur virtuel est compris et bien ressenti.

8. Bibliographie

- Bonucci A., Analyse phonologique et indexation figurative pour une base de données d’entrées lexicales de la Langue des Signes Française. Thèse de doctorat, Université Lyon 2, 1998.
- Bouillon P., Vandooren F., Da Sylva L., Jacqmin L., Lehmann S., Russel G., Viegas E., *Traitement automatique des langues naturelles*, Duculot, 1998.
- Brentari, D., *A prosodic model of sign language phonology*. Bradford Book, 1998.
- Bruderlin A., Williams L., « Motion signal processing », Proceedings of the 22nd annual conference on Computer graphics and interactive techniques SIGGRAPH ‘95, New York, NY, USA : ACM Press, 1995, p. 97-104.
- Cuxac C., *La LSF, les voies de l’iconicité*. Ophrys, Paris, 2000.
- Dubuisson C., Lelièvre L., Miller C., Grammaire descriptive de la langue des signes québécoise. Tome 1 : Le comportement manuel et le comportement non manuel, 2^e éd. rev. et augm. Montréal : Université du Québec à Montréal, 1997.
- Elliott R., Glauert J. R. W., Jennings V., Kennaway J., « An overview of the sigml notation and sigml signing software system », In Fourth International Conference on Language Resources and Evaluation, LREC 2004, p. 98-104.
- Iheddadene M., Traduction Automatique : étude et réalisation d’un module de génération à partir d’une représentation sémantique interlingue. Thèse de Doctorat, Université de Provence, France, 2006.
- Kervajan L., Guimier de Neef E., Véronis J., « Processing of French Sign Language : verbs agreement », In S. Gibet et al. (Eds.), Lectures Notes in Computer Science (LNCS), Springer-Verlag, Berlin, 2006.
- Kervajan L., « Problèmes de représentation de la Langue des Signes Française en vue du Traitement Automatique », Actes de la conférence TALN-RECITAL 2006, Leuven, Belgique.

- Lebourque T., Gibet S., « A complete system for the specification and the generation of sign language gestures ». *Lecture Notes in Artificial Intelligence, LNAI 1739*, in *Gesture-Based Communication in Human-Computer Interaction*, 1999.
- Lee J., Kunii T., « Computer animated visual translation from natural language to sign language ». *Journal of Visualization and Computer Animation*, 4(2), 1993, p. 63-78.
- Liddell S., *Grammar, Gesture, and Meaning in American Sign Language*, Cambridge Univ. Press, UK, 2003.
- Losson O., Vannobel J., « Sign language formal description and synthesis », *International Journal of Virtual Reality*, 3(4), 1998, p. 27-34.
- Martinet A., *Économie des changements phonétiques : traité de phonologie diachronique*, Berne : Francke, 1955.
- Miller, C., *Phonologie de la langue des signes québécoise. Structure simultanée et axe temporel*, thèse de doctorat, Université du Québec à Montréal, Canada, 1997.
- Moody B., *Dictionnaire de LSF, Tome 1 : Introduction à l'histoire et à la grammaire de la langue des signes*, Ellipse - IVT, Vincennes, France, 1998.
- Park J., Maillebau E., Guimier de Neef E., Vinesse J., Heinecke J., « Evaluating an Interlingual Semantic Representation - application to FTRD's MMT-inspired system », *Third International Conference on Meaning Text Theory (MTT)*, Austria, 2007.
- Prillwitz S. L., Leven R., Zienert H., Zienert R., Hanke T., Henning J., « Hamnosys. version 2.0 ; hamburg notation system for sign languages. An introductory guide », *International Studies on Sign Language and Communication of the Deaf*, 1989.
- Sowa J. F., « Conceptual Graphs for a Data Base Interface », *IBM. Journal of Research and Development* 20(4), 1976, p. 336-357.
- Stokoe W. C., « Sign Language Structure. Studies in Linguistics ». *Occasional Papers n° 8*. Buffalo, NY, University of Buffalo Press, 1960.
- Sutton V., Gleaves R., *SignWriter - The world's first sign language processor*. Deaf Action Committee for SignWriting, La Jolla, CA, 1995.
- Tolani D., Goswami A., Badler N. I., « Real-time inverse kinematic techniques for anthropomorphic limbs ». *Graphical Models*, 62(5), 2000, p. 353-388.
- Witkin A., Popovic Z., « Motion warping ». *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques SIGGRAPH '95*, New York, NY, USA : ACM Press, 1995, p. 105-108.