

---

# Formalisme de description des gestes de la langue des signes française pour la génération du mouvement de signeurs virtuels

Sylvie Gibet\* et Alexis Héloir\*\*

\*IRISA

Campus de Beaulieu  
F - 35 042 Rennes Cedex  
sylvie.gibet@irisa.fr

\*\*Laboratoire VALORIA

Campus de Tohannic – Bât. Yves Coppens  
F-56017Vannes cedex

---

*RÉSUMÉ.* Cet article présente un modèle de génération de gestes de la langue des signes française qui s'appuie sur une approche de modélisation semi-formelle des gestes, et sur un formalisme de spécification capable de traduire un énoncé en un flux continu de données pour le contrôle du mouvement d'un signeur virtuel. Ce modèle bénéficie à la fois d'une connaissance des structures linguistiques propres à cette langue, et de résultats d'analyse issus de mouvements capturés.

*ABSTRACT.* This paper presents a model for generating French sign language gestures, which is based on both a semi-formal modelling approach, and on a specification formalism yielding to the translation of an utterance into a continuous data flow for the control of a virtual character. This approach benefits from knowledge of structural linguistics proper to sign language, and results of motion capture analysis.

*MOTS-CLÉS :* Langue des signes française, signeur virtuel, animation, modélisation, spécification, génération, analyse segmentale, synthèse, iconicité.

*KEYWORDS:* French sign language, virtual signer, animation, modelling, specification, generation, segmental analysis, synthesis, iconicity.

---

## 1. Introduction

Nous proposons dans cet article un modèle de description et de spécification des gestes de la langue des signes en vue de la génération de séquences gestuelles expressives par un humanoïde signeur. Ce travail est original dans le sens où il concilie une approche linguistique théorique et expérimentale qui vise à analyser les mécanismes de modulation des gestes de la langue des signes française (LSF), et une approche de spécification de haut niveau traduite en commandes gestuelles pour l'animation d'agent virtuel.

Il serait réducteur de considérer le problème de génération d'une séquence gestuelle en langue des signes française d'un point de vue uniquement lexical., ou comme une simple concaténation de primitives de mouvements. Une telle conception conduirait à un système de génération capable de produire au mieux, des énoncés de français signé<sup>1</sup> et dans bien des cas, des séquences inintelligibles. Il est donc impossible de faire l'économie des caractéristiques fondamentales des langues des signes (désormais LS). Une appréhension de ces caractéristiques requiert la prise en compte de résultats d'études sémiotiques des LS. Le système de génération doit ainsi intégrer les mécanismes qui structurent et modulent ces langues. Ces mécanismes sont décrits en particulier dans le cadre de la théorie de l'iconicité proposée par (Cuxac, 2000) qui est définie par le lien de ressemblance entre ce qui est dit et la réalité décrite. Une modification, même subtile, de l'énoncé se traduit souvent par une modification sensible des mouvements produits, à la fois d'un point de vue de la sélection et de l'organisation des signes, que des procédés flexionnels appliqués aux signes.

Par ailleurs, la génération informatisée de gestes requiert la mise en œuvre de méthodes capables de traduire un énoncé exprimé en langue des signes française en une séquence de commandes gestuelles interprétables par un moteur d'animation. Les difficultés sont de plusieurs ordres. Le système d'animation doit être capable, à partir d'une spécification de haut niveau, de produire un flux continu de mouvements qui sont perçus comme étant naturels et compréhensibles. Or, s'il apparaît nécessaire pour les usagers que les agents virtuels soient crédibles, la naturalité des mouvements reste encore difficile à qualifier. La plupart des approches développées pour l'animation réaliste de personnages reposent sur l'utilisation de mouvements préenregistrés. L'inconvénient de ce type d'approche est lié au caractère figé des mouvements capturés, qui ne se prête pas à la génération de nouvelles séquences de gestes. Il faudrait être capable d'enregistrer et de traiter de très gros corpus de données, associés à un grand volume de données de mouvements, pour explorer une multitude de situations linguistiques différentes dans des contextes diversifiés. À cette variabilité linguistique se superpose une variabilité gestuelle qui se traduit par une modification des articulateurs impliqués dans les gestes, ainsi que des modifications de la géométrie et la cinématique des mouvements.

---

1. Le français signé emprunte à la LSF son vocabulaire imagé, les signes, mais la construction du discours reste calquée sur la syntaxe du français parlé.

Il est important dans ce cadre de mettre en correspondance les modifications linguistiques et celles des mouvements engendrés, de façon à préserver la cohérence des structures des signes ainsi que la naturalité des mouvements produits. Opérer directement des transformations sur les mouvements capturés peut conduire à des séquences gestuelles dont la structure est altérée, et qui deviennent par conséquent inintelligibles. Cette variabilité des signes, agissant au niveau de la structure et du signal constitue un défi majeur au problème de la génération automatique de mouvements.

Dans cet article nous abordons le problème de la génération de gestes de la LSF, en analysant, à la lumière d'études sémiotiques, la relation entre la structure des signes et la production d'un flux continu de mouvement. La génération est réalisée grâce à une spécification hiérarchisée allant de la caractérisation d'un énoncé à la production paramétrée de postures de mouvement. Partant des structures de grande iconicité proposées par (Cuxac, 2000) (transferts de taille ou de forme, transferts situationnels ou personnels), nous nous intéressons aux mécanismes de flexion qui modulent les signes en agissant directement au niveau *phonologique*. La représentation phonologique (terme emprunté aux langues orales) sera utilisée dans le cadre de cet article pour décrire dans les langues des signes la façon dont s'organisent les unités gestuelles atomiques afin de former des signes et des phrases signées. Les unités gestuelles atomiques sont également appelées paramètres ou éléments constitutifs des signes. Pour décrire les gestes nous adoptons une méthodologie à la fois théorique et expérimentale. La partie théorique s'appuie sur une modélisation des flexions appliquées aux paramètres des gestes de la LSF. La partie expérimentale concerne l'annotation des gestes prenant en compte l'aspect multidimensionnel des gestes et le découpage en unités signifiantes et non signifiantes. Nous montrons ensuite comment cette spécification des gestes à plusieurs niveaux peut être directement interprétée par un moteur d'animation dirigé par les buts.

La seconde section explore les principaux travaux dédiés à la description et la génération de gestes des LS. La section 3 décrit l'approche que nous avons adoptée pour la génération des gestes. La section 4 présente une modélisation des gestes de la LSF qui s'appuie sur une représentation du discours et des entités qui le constituent. Un modèle segmental ainsi qu'un modèle de spécification morphémique des signes sont décrits dans la section 5. Quelques résultats significatifs illustrent la synthèse de gestes de la LSF dans la section 6. La section 7 présente un bilan et ouvre quelques pistes de recherches futures.

## **2. Principaux travaux relatifs à la description et la génération de gestes**

Représenter et engendrer les gestes des langues des signes (LS) est un travail multidisciplinaire qui requiert une compréhension fine des mécanismes linguistiques des LS, mais également des compétences relatives à la modélisation du mouvement afin

de définir des systèmes permettant le passage de la spécification des gestes à leur génération, ceci dans le but d'animer des humanoïdes virtuels expressifs.

Cette section ne prétend pas faire un état de l'art exhaustif des modèles linguistiques propres aux LS, mais propose un tour d'horizon rapide des différents systèmes de notation et de transcription des gestes des LS, ainsi que des langages informatiques permettant leur description.

### ***2.1. Systèmes de notation et de transcription des gestes des langues des signes***

Les tentatives pour représenter et décrire les gestes humains ont conduit à l'élaboration de systèmes de notation et de transcription gestuelle, ainsi qu'au développement de langages informatiques dédiés à l'animation d'agents virtuels de communication. Tous ces systèmes visent à décrire des mouvements plus ou moins codifiés et structurés suivant qu'ils s'appuient ou non sur des unités gestuelles bien identifiées, et des règles syntaxiques ou sémantiques. Les objectifs de ces représentations sont multiples : elles permettent de fixer des phénomènes gestuels ayant un déroulement dans le temps, de les explorer, de les transmettre et de les faire évoluer. Elles favorisent en particulier le stockage, l'échange et la communication d'informations à distance.

Différents systèmes de taxonomie et de classification ont déjà été proposés dans le cadre de l'étude des gestes des LS, et de manière plus générale des gestes coverbaux accompagnant la parole. L'identification des types de gestes se base sur des études dédiées à la détermination des phases constitutives des gestes (Kendon, 1993 ; McNeill, 1992). Des études plus récentes dédiées aux gestes expressifs s'attachent à la segmentation et l'annotation de ces gestes. Ces études visent à caractériser la structure spatiale et temporelle de phrases signées, permettant ainsi le passage d'une description segmentale vers la synthèse du mouvement (Kipp, 2004).

Certains travaux sur la phonologie du signe ont donné lieu à différents types de représentations. Les travaux de (Stokoe, 1972) ont abouti à la description de l'ASL (American Sign Language) sous la forme d'une combinaison d'unités de mouvements élémentaires, appelées *cherems*. La notation consiste en un nombre limité de symboles appartenant à trois classes distinctes, chacune de ces classes étant associée à l'un des paramètres constituant le signe : l'emplacement du signe (*tabula ou TAB*), la forme de la main (*designator ou DEZ*), et le mouvement (*signation ou SIG*). L'une des hypothèses de base repose sur le fait que deux signes distincts peuvent être différenciés lorsque l'un seulement des paramètres constitutifs est modifié (paires minimales). En d'autres termes, la variation de ces paramètres au cours d'un signe n'est pas considérée comme étant significative. Un dictionnaire de l'ASL a été constitué à partir de cette représentation, en respectant l'ordre sur les paramètres TAB, DEZ, SIG. Les mouvements peuvent être exécutés en séquence ou en parallèle. Poursuivant les travaux de Stokoe, d'autres paramètres ont été identifiés, qui participent à la formation et la distinction des signes. En particulier l'orientation de la main et quatre zones principales

d'articulation du signe ont été définies. (Klima et Bellugi, 1979) définissent par ailleurs l'orientation de la paume de la main et un certain nombre de zones de contact de la main avec le corps (joue, bras, etc.).

D'autres analyses linguistiques de la langue des signes française (LSF) s'appuient sur une *théorie de l'Iconicité* (Cuxac, 2000), qui place « l'icône » ou la ressemblance au centre de la théorie des signes. Cuxac propose ainsi une alternative à l'approche phonologique de Stokoe, en intégrant dans les structures des signes la dimension de leur iconicité. Il assimile certains paramètres constitutifs d'un signe à des morphèmes. Les morphèmes sont définis ici comme étant les plus petites unités de signes porteuses de sens. Ces morphèmes peuvent être associés à un ou plusieurs paramètres identifiés ci-dessus.

Un certain nombre de travaux s'intéressent à la notation et à la transcription informatisée des signes. L'un des systèmes de notation les plus exhaustifs est « HamNoSys », acronyme d'Hamburg Notation System (Prillwitz *et al.*, 1989). Ce système propose un codage alphabétique des signes constitué d'un nombre important de symboles (autour de 200) et permet de décrire de façon internationale la plupart des signes.

Le système SignWriting (Sutton, 1998), qui s'inspire d'une écriture issue de la chorégraphie, propose un système de représentation graphique des signes. Une étude de l'informatisation du formalisme SignWriting au niveau du système d'exploitation est présentée dans (Aznar et Dalle, 2004).

La plupart de ces systèmes de notation et de transcription ont conduit à l'élaboration de dictionnaires adaptés à la recherche d'items lexicaux (Moody, 1993). Cependant, ces formalismes restent limités quant à leur capacité à représenter précisément l'espace de signation, ainsi que les aspects dynamiques et parallélisés des signes.

Par ailleurs, les notations décrites ci-dessus, bien adaptées à des fins d'études linguistiques, sont purement scripturales et ne permettent pas de représenter tous les aspects variables des langues des signes, qu'ils proviennent de variations syntaxico-sémantiques, ou de flexions des signes liées à l'iconicité. Enfin, ils ne permettent pas une traduction aisée de la description linguistique des signes vers la génération de commandes gestuelles destinées à contrôler un personnage virtuel. En effet, il est difficile de décrire des séquences gestuelles temporelles et de leur appliquer des transformations dans le temps, tout en préservant la cohérence structurelle de l'ensemble.

## ***2.2. De la spécification à la génération de mouvements***

Le développement de langages informatiques pour la spécification et la génération automatique de signes dans les LS a vu le jour plus récemment. En 1998, les travaux de Lebourque ont ouvert la voie aux systèmes de spécification et de génération de gestes de la LSF (Lebourque, 1998). Le langage de Lebourque repose sur une description

textuelle des éléments constitutifs des signes, appelés *gestèmes*, couplée à des méthodes de génération du geste basées sur des paradigmes sensorimoteurs (Gibet et Marteau, 1994). Les gestèmes sont des unités de mouvement porteuses ou non de sens. Ils sont directement reliés aux paramètres qui permettent de décrire les signes, à savoir l'emplacement des signes, la posture et le mouvement des mains ainsi que leur orientation. Certains gestèmes permettent d'assurer les transitions entre les signes. En 2000, (Losson, 2000) a proposé une description des gestes de la LSF qui repose sur le modèle « Movement-Hold » de (Liddell et Johnson, 1989).

Plus récemment, dans le cadre des projets européens ViSiCast et eSign, la notation HamNoSys a été transcrite dans le langage de spécification SigML, et testée pour décrire les langues des signes anglaise, néerlandaise et allemande (Kennaway, 2004). Ce langage textuel est utilisé pour contrôler des avatars signeurs. C'est l'un des langages de description de gestes des LS le plus abouti aujourd'hui.

La plupart des systèmes centrés sur l'animation d'avatars signeurs ont conduit à proposer des approches de synthèse pure, directement inspirées de la description en éléments de base des gestes des LS (appelés ici morphèmes ou gestèmes). En conséquence, les séquences générées à partir de ces modèles peuvent difficilement dépasser la construction de gestes par agencement morphémique ou lexical. Ces systèmes se limitent le plus souvent à la prise en compte des signes standard stabilisés et sont difficilement paramétrables. Par ailleurs, les séquences gestuelles générées sont difficilement réalistes. D'autres travaux se sont appuyés sur des données capturées pour produire des animations plus réalistes. Ces derniers travaux se contentent de passer d'une annotation déterministe vers des commandes gestuelles relativement figées (lignes temporelles prédéfinies). Les méthodes de synthèse associées sont pour la plupart des méthodes de traitement du signal qui transforment les gestes enregistrés (interpolation), sans forcément s'intéresser aux aspects sémantiques des mouvements.

Ce n'est que très récemment que les recherches sur la génération de gestes par des agents virtuels signeurs se sont intéressées à certains mécanismes fondamentaux qui structurent les LS, autorisant ainsi une plus grande variabilité dans la production de ces gestes ainsi qu'une adaptation au contexte du discours. En particulier, les travaux de Huenerfauth prennent en compte les aspects directionnels des verbes ainsi que les relations spatiales entre les entités des énoncés en ASL (Huenerfauth, 2006). (Lenseigne et Dalle, 2006) ont proposé une modélisation de l'espace de signation (là où la personne signe) et des relations sémantiques entre les éléments du discours qui s'expriment de façon géométrique dans cet espace. Ce modèle a été appliqué à la description d'énoncés narratifs en LSF et de situations de dialogues. D'autres travaux ont pour objectif la création d'un modèle de spécification de gestes de la LSF basé sur une grammaire spatiale et une représentation géométrique du lexique (Filhol et Braffort, 2006).

Dans le cadre de nos recherches, nous tentons d'établir un pont entre des représentations linguistiques et gestuelles. Nous n'avons pas cherché à étudier les structures syntaxiques et sémantiques sous-jacentes à un ensemble de phrases signées

en LSF, mais nous nous sommes intéressés aux modulations subies par les signes dans un contexte discursif, modulations qui peuvent être attribuées à des mécanismes sémiotiques, et qui influencent les mouvements engendrés. Dans ce contexte, nous proposons une approche originale qui allie étroitement les problématiques liées à l'analyse et à la génération de gestes de la LSF. L'analyse répertorie un certain nombre de mécanismes linguistiques responsables de la flexion des signes et les interprète sous la forme de paramètres de spécification des gestes. L'analyse permet également d'identifier des invariants au niveau des mouvements capturés. La génération permet de gérer le passage de la spécification des gestes vers le moteur d'animation.

### 3. Approche

Les langues des signes remettent en question les frontières habituelles des théories linguistiques associées aux langues orales. Ceci est principalement dû au fait qu'elles utilisent le canal visuo-gestuel, contrairement aux langues orales qui utilisent le canal audio-oral. Cette spécificité des LS est probablement à l'origine de l'utilisation de mécanismes liés au caractère iconique des signes. C'est pourquoi les travaux de mise en évidence des structures de grande iconicité par (Christian Cuxac, 2000) sont placés au centre de nos analyses. Le processus *d'iconicisation* est défini par Cuxac comme étant le processus par lequel le locuteur va rendre iconique l'expérience vécue, imaginée ou exécutée. L'iconicité peut être caractérisée par le lien de ressemblance plus ou moins étroit entre les entités du monde réel, le référent et le signe qui s'y rapporte. Dans une perspective sémiogénétique, Cuxac définit deux façons de signer, qui coexistent en permanence lors d'activités discursives. Il y a d'une part la grande iconicité, qui consiste à « dire en montrant » d'autre part les signes standard, sans visée illustrative.

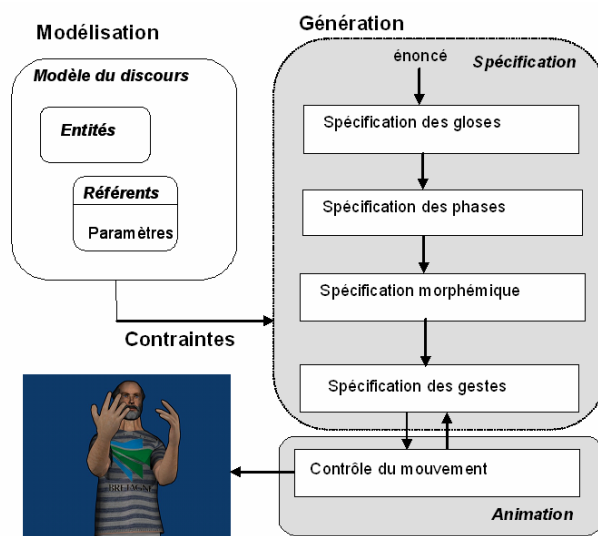
Plutôt que de considérer ces structures de grande iconicité comme de la pantomime, Cuxac les identifie comme des éléments linguistiques fondamentaux, et les met au centre d'une théorie linguistique de l'iconicité (Cuxac, 2000), (Sallandre, 2003). Ces structures constituent des invariants dans les différentes langues des signes, favorisant ainsi la compréhension mutuelle de signeurs de différentes nationalités.

Nous verrons à la section 4 quelques mécanismes de base de modulation des signes manuels, et comment ces mécanismes affectent directement les attributs spatiaux et temporels des signes. Nous verrons également comment les structures de grande iconicité contribuent activement à ces phénomènes modulateurs.

Comme il existe une certaine dualité entre la représentation des signes (cognitive, syntaxique et sémantique) et la représentation des mouvements mis en jeu lors de l'exécution des signes, nous nous intéressons dans cette étude à la façon dont les paramètres linguistiques vont agir sur les éléments constitutifs des signes.

L'approche que nous avons adoptée sépare les niveaux modélisation et génération. Elle est présentée à la figure 1.

La modélisation des gestes de la LSF est abordée à travers un processus sémiotique qui permet de sélectionner et d'agencer un ensemble d'entités qui représentent des concepts appréhendés par l'agent signeur. Ces entités sont elles-mêmes associées à des référents qui possèdent un certain nombre d'attributs spatiaux ou temporels. Les attributs spatiaux caractérisent le lieu d'articulation du mouvement, les localisations dans l'espace de signation, ou bien la forme d'une trajectoire. Ils peuvent également exprimer des relations de dépendance de groupes d'articulateurs. Les attributs temporels se traduisent par des relations d'ordonnement de signes ou d'unités morphémiques (telle unité se déroule avant telle autre ou en parallèle avec une autre). Ce niveau de modélisation reste relativement qualitatif. Il peut être associé à un ensemble de valeurs sémantiques auxquelles seront affectées des valeurs numériques lors de l'exécution du mouvement.



**Figure 1.** Architecture du système de modélisation/génération des geste de la LSF

La génération des gestes de la LSF nécessite quant à elle de mettre en place plusieurs niveaux de spécification et un modèle d'animation. Dans le cadre du modèle de spécification, nous partons d'une description d'un énoncé en LSF sous la forme d'une séquence de gloses, lesquelles sont ensuite décrites en une spécification segmentale structurale mettant en œuvre un découpage spatial et temporel. Le découpage spatial caractérise la modélisation multicanale explicitée à la section 5, chaque canal étant associé à un groupe d'articulateurs spécifique. Le découpage temporel détermine les phases des mouvements, lesquelles peuvent être signifiantes ou non signifiantes. Enfin, les phases signifiantes sont associées à des séquences de



gestèmes<sup>2</sup>, chaque gestème correspondant à une primitive de mouvement paramétrée (Lebourque et Gibet, 1999), (Gibet *et al.*, 2001). Cette spécification est ensuite traduite en une commande gestuelle quantitative qui contrôle directement les chaînes articulées de l'agent virtuel.

Le passage de la modélisation à la spécification est réalisé par l'identification d'un ensemble de contraintes spatiales et temporelles qui expriment les flexions des signes et l'instanciation de paramètres inhérents à la spécification des gestes. Les aspects propres à l'enchaînement des signes, la variabilité temporelle qui caractérise une certaine forme de « prosodie », ainsi que la mise en place de pauses dans les mouvements sont directement traités par le modèle de contrôle du mouvement.

En ce qui concerne la spécification de la commande gestuelle, nous nous appuyons sur l'analyse d'une base de données de mouvements capturés (Héloir *et al.*, 2006b), contrairement aux travaux antérieurs décrits dans (Lebourque, 1998) qui n'abordaient qu'une problématique de synthèse pure et l'articulation de la spécification de bas niveau vers la génération. L'analyse des données enregistrées a mis en évidence des invariants temporels qui caractérisent d'une part la coordination entre les différents canaux et d'autre part la répartition relative des phases transitoires et des phases porteuses de sens (Héloir et Gibet, 2007). Elle a permis en outre de définir un niveau de représentation segmentale associé à une spécification morphémique qui est décrit dans la section 5. Pour la spécification morphémique, nous avons utilisé la grammaire définie dans (Gibet *et al.*, 2001), que nous avons complétée et adaptée (Héloir, à paraître).

La génération ainsi décrite a été implémentée pour synthétiser des gestes de la langue des signes française appartenant aux deux corpus définis en annexe 1. Le moteur d'animation de l'agent virtuel est guidé par des trajectoires capturées appartenant au corpus, et exploite les mécanismes de cinématique inverse et de coarticulation définis dans (Gibet et Marteau, 1994). Une analyse spécifique de la variabilité temporelle des gestes, effectuée sur la même séquence de gestes avec des styles différents a conduit à proposer un module de synthèse « stylisée » des gestes qui permet de moduler les gestes produits par notre système d'animation, favorisant ainsi la plausibilité et l'expressivité de l'agent virtuel signeur (Héloir *et al.*, 2006a).

#### 4. Modélisation des gestes de la langue des signes française

En nous appuyant sur un modèle du discours tel qu'il a été défini dans (Héloir et Kervajan, 2007), nous décrivons dans cette section la notion d'entité qui constitue l'élément de base de ce modèle, avec le langage semi-formel de modélisation *UML*<sup>3</sup>. Ce métalangage permet de représenter à un certain niveau d'abstraction les relations entre les éléments descriptifs des entités. Il peut s'adapter à différents formalismes exprimant la grammaire de la LSF. Nous décrivons ensuite les principaux mécanismes de

---

2. Un gestème est la plus petite unité de geste participant à la construction d'un signe

3. UML signifie « langage de modélisation unifié », de l'anglais *Unified Modeling Language*.

modulation des signes. Dans le cadre de la visée iconique, nous décrivons, à titre d'exemple, quelques processus de transferts identifiés par Cuxac, et montrons comment ces processus sont susceptibles d'incorporer des éléments de modulation des signes et d'affecter les entités dans le contexte du discours (Cuxac, 2000).

#### 4.1. Notion d'entité et sa représentation

Une entité représente un concept qui est appréhendé par le signeur virtuel. Elle peut représenter une personne, un objet ou une notion plus abstraite dont il est question dans le cadre du discours. Une entité dans le discours est liée à un référent. Nous nous appuyons sur la notion de référent proposée par (Lenseigne et Dalle, 2006), qui représente une unité sémantique du discours pouvant être spécialisée en *date*, *lieu*, *action* ou *objet*. Un référent revêt une *fonction* dans le discours. La représentation gestuelle d'une entité peut se faire hors visée iconique à l'aide d'un signe standard qui la représente (par exemple, un signeur utilisera le signe standard « arbre » pour représenter un arbre). Ce signe standard ou *stabilisé* représente, par convention, l'entité. En visée iconique, la représentation d'une entité ne passe plus par les signes standard mais par une description gestuelle des propriétés pertinentes de l'entité selon sa fonction dans le discours.

Par exemple, lorsqu'un locuteur représente un chien en visée iconique, il va s'attacher à représenter le chien évoqué « tel qu'il est », en faisant référence à ses caractéristiques remarquables. Si le chien dont il est question est un boxer à la démarche pataude, le locuteur s'appliquera à représenter la physionomie et la démarche du boxer. Un locuteur représentant un chien est illustré à la figure 2. Ainsi, en visée iconique, le locuteur représente le référent en montrant des traits saillants qui le caractérisent. En visée non iconique, le locuteur fait référence au signifié, « le chien » à l'aide d'un signe standard.



**Figure 2.** *Signe du boxer ; il s'agit du signe [CHIEN], représenté de façon iconique : ce chien a une démarche « pataude »*

Le diagramme de la figure 3 offre, à l'aide du langage UML, une représentation des relations entre l'entité, son référent et les fonctions associées. Les fonctions peuvent couvrir trois types : agent, localisation spatiale ou temporelle. Elles permettent de

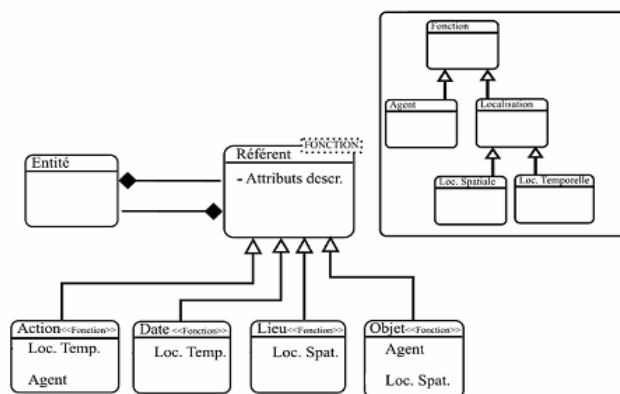
construire des spécialisations d'actions, de date, de lieu ou d'objet, héritant de la classe *réfèrent*. Ce dernier référence une entité et la qualifie à l'aide d'attributs descriptifs.

Hors visée illustrative, une entité peut être représentée à l'aide d'un signe standard. Lorsqu'il n'existe pas de signe standard pour représenter une entité (par exemple pour introduire une personne entendante ne disposant pas de nom signé), celle-ci est épelée à l'aide de la dactylogogie, après avoir précisé éventuellement son type (personne, lieu, etc.).

#### 4.2. Modulations des signes

Les langues des signes structurent la phrase en positionnant dans *l'espace de signation* des entités définies par leur fonction (actants, objets, etc.). Le contexte (lieu, date, etc.) est ainsi précisé en suivant un certain ordonnancement.

Parmi les phénomènes modulateurs des signes, nous avons considéré uniquement ceux qui agissent sur les paramètres manuels des signes au niveau géométrique et dynamique. Nous n'avons pas traité les flexions non manuelles et en particulier les expressions faciales dont les fonctions syntaxiques et lexicales revêtent une grande importance dans les LS. Nous nous sommes intéressés plus spécifiquement aux mécanismes de localisation, de verbes directionnels, aux caractéristiques propres à la forme et à l'amplitude des mouvements, ainsi qu'aux modulations de la dynamique des signes.



**Figure 3.** Diagramme de classe représentant l'entité et son référent ; les flèches  $\rightarrow$  représentent le mécanisme d'héritage : une action, une date, un lieu ou un objet héritent des attributs descriptifs du référent associé ; les liens  $\blacklozenge$  symbolisent l'association entité/référent. Fonction est une classe générique qui peut être spécialisée (agent, localisation) ; cette classe est passée en argument de la classe référent.

Pour exprimer les positions et orientations relatives entre entités signifiées, les langues des signes ont recours aux positions et configurations manuelles. Le signeur utilise ainsi *l'espace* en positionnant les entités, animées ou non, présentes dans sa narration. Plusieurs techniques permettent de préciser une localisation : le décalage du signe dans l'espace, la désignation de l'endroit par pointage de l'index, l'orientation du regard, de la tête ou du buste pendant l'exécution du signe, ou l'utilisation de signe possédant une fonction intrinsèque de localisation. Dans le langage de spécification défini dans (Gibet *et al.*, 2001), nous avons défini un emplacement à partir des définitions *location* et *body\_location*, elles-mêmes définies à partir de valeurs de *direction* et de *distance* (cf. annexe 2). Au niveau des entités, la localisation peut être précisée grâce à l'attribut *location* associé au référent du signe.

Les verbes dans les langues des signes peuvent être séparés en deux groupes : les verbes directionnels, qui s'accordent avec leur sujet et objet, et ceux qui ne peuvent pas incorporer de pronoms personnels. Ces derniers sont toujours réalisés de la même façon, les références pronominales étant réalisées par geste déictique. Les verbes directionnels, quant à eux, sont tels que les positions et orientations initiale et finale dépendent de l'agent et du patient par rapport auxquels le signe est fléchi. Un tel mécanisme peut être intégré à la primitive de mouvement inhérente au signe. Par exemple, le verbe *donner* est représenté dans le cadre de la grammaire définie dans (Gibet *et al.*, 2001) par :

« *Pointing(loc\_agent, loc\_patient)* », où *loc\_agent* et *loc\_patient* représentent les localisations respectives des agents et patients, et *Pointing* correspond à la primitive de mouvement de pointage propre à l'action de donner.

La forme et l'amplitude des mouvements peuvent être gérées par des mécanismes similaires à ceux utilisés dans les processus de transfert de taille ou de forme pour les signes à visée iconique (cf. section 4.3.1).

Enfin, les flexions propres à la dynamique des signes traduisent la manière ou la durée. Par exemple, la répétition rapide du verbe aller signifie *aller souvent*. Ces flexions sont directement traitées par le modèle de contrôle sensorimoteur, capable de prendre en compte un changement de vitesse en début et fin de mouvement.

#### **4.3. Structures de grande iconicité**

En visée illustrative, le signeur s'attache à représenter les aspects importants du référent de l'entité dans le contexte du discours. Ces aspects importants peuvent être des caractéristiques marquantes et représentatives du référent (caractéristiques physiques, forme globale, taille), des relations spatiales et/ou topologiques entre les référents, ou encore des traits relatifs au comportement, au caractère, à la culture ou à l'état mental d'un actant de l'énoncé. Selon le cas, ces aspects seront représentés à l'aide de l'une des trois structures de grande iconicité identifiées par (Cuxac, 2000) : *transferts de taille et/ou de forme*, *transferts de situation* et *transferts personnels*. Nous évoquons ici les

deux premières structures, les transferts personnels mettant en œuvre des processus plus complexes nécessitant la plupart du temps la modélisation des mimiques faciales.

#### 4.3.1. *Transfert de taille ou de forme*

La structure de transfert de taille et de forme permet de représenter la taille ou la forme, partielle ou globale, de lieux, d'objets ou de personnages. Les gestes permettant de figurer ces tailles ou ces formes dans l'espace de signation se composent :

- d'une configuration de la main (ou des mains) indiquant une forme de base : une oreille de chien sera par exemple représentée par un cône, esquissé par le pouce, l'index et le majeur ;
- d'un mouvement et d'une orientation de la main (ou des mains) signifiant le déploiement de la forme dans l'espace : selon la physionomie du chien, la forme du cône représentant l'oreille sera différente ;
- d'un emplacement de départ du déploiement : les oreilles du chien sont représentées sur la tête au-dessus du front de l'agent ;
- d'une mimique faciale

Un inventaire de configurations manuelles décrivant les formes illustrées a été répertorié par (Cuxac, 2000). Les formes décrites par les configurations manuelles peuvent représenter, entre autres, l'épaisseur (plus ou moins), la forme (sphérique, carrée, rectangulaire) ou un agencement (rayonnant, aligné, grillagé, en éventail). Lorsque plusieurs transferts de formes sont réalisés pour caractériser un objet complexe ou un personnage, on parle de *descripteurs*. Les descripteurs peuvent être préférés au lexique standard dans la mesure où ils offrent une représentation plus pertinente de l'objet dans son contexte. Par exemple, dans un récit tiré du corpus de Cuxac, mettant en scène deux chiens de morphologies très différentes (un boxer et un chien-loup), les deux protagonistes du récit sont décrits à l'aide de transferts de formes sans que le signe standard [CHIEN] ne soit utilisé (Cuxac, 2000).

Selon la visée illustrative, les transferts de taille ou de forme décrivent des attributs significatifs d'un référent faisant partie de l'énoncé. Par exemple, les attributs des deux chiens protagonistes du récit de Cuxac sont caractérisés par le museau, les oreilles, les bajoues du boxer ainsi que ses pattes (figure 2).

De manière générique, on peut spécifier la liste d'attributs des référents associés à une entité du discours, chaque attribut comportant les descripteurs d'emplacement (*location*), de forme (*form*) et de lieu d'action (*plan*). La taille est implicitement spécifiée au sein de l'attribut *form* :

*Referent::Attribut listOfAttributs*  
*Attribut* = { *location*  
                   *form*  
                   *plan* }

Pour plus de précision, se référer à l'annexe 2.

#### 4.3.2. Transfert de situation

Par un *transfert de situation*, le locuteur vise à reproduire dans son espace de signation les scènes qui figurent le déplacement spatial d'un actant par rapport à un locatif stable. L'action effectuée par la main dominante s'inscrit dans un sous-ensemble restreint de configurations que l'on appelle *proformes*. Les énoncés de transferts situationnels se réalisent de la façon suivante :

- configuration du locatif par la main dominée : par exemple, pour une piste de décollage, on aura la main dominée à plat et le bras horizontal ;
- configuration de l'agent par la main dominante, comme pour représenter un avion : le pouce et l'auriculaire de la main dominante écartés suggèrent les ailes ;
- mouvement effectué par la main dominante, représentant la trajectoire de l'avion relativement à la piste ;
- emplacement d'arrivée, où la main dominante est « en l'air », bras tendu.

Lorsqu'un signeur représente un avion en train de décoller, la main dominée est animée d'un mouvement de va-et-vient sous la main dominante pour figurer le sol qui défile sous les ailes de l'avion. La relation qui lie l'actant et le locatif est de type trajectoire. Cette relation s'applique sur l'actant et sa localisation et est paramétrée par une trajectoire et un profil de vitesse. La figure 4 représente l'élaboration d'un *transfert de situation* selon notre modèle du discours. L'exemple employé est le *transfert de situation* représentant un avion qui décolle.

En *transfert de situation*, la réalisation de la trajectoire a une fonction hautement illustrative. La manière dont l'agent réalise le mouvement offre, en effet, une description du déplacement réalisé par l'actant (forme de la trajectoire, profil de vitesse) et des relations spatiales qui le lient avec l'objet référé (Braffort et Lejeune, 2006).



**Figure 4.** Spécification des relations spatiales en transfert de situation ; à gauche : l'avion (main dominante) sur la piste d'atterrissage (main dominée) ; à droite : mouvement de la main dominante représentant l'avion en train de décoller

Dans l'exemple proposé, chaque objet (l'avion et la piste) peuvent être représentés par son référent, sous la forme d'un ensemble d'attributs.

Objet <<Agent>> avion Référent :: Attributs listOfAttributs Forme :: » Aile »	Objet <<Localisation>> piste Référent :: Attributs listOfAttributs Emplacement :: zone_neutre Forme :: » Plat » Plan :: Y
---	---

La relation spatiale entre l'actant et le locatif peut être représentée en début et fin de mouvement par la spécification de deux gestes réalisés en séquence :

Geste 1 : « L'avion est sur la piste »

*Pointing(main\_dominée, zone\_neutre)*

déplacement de la main dominée dans la zone neutre

*Handshape(main\_dominée, piste)*

la main dominée est dans la configuration « Plat » (piste)

*Pointing(main\_dominante, contact, main\_dominée)*

la main dominante est sur la main dominée

*Handshape(main\_dominante, avion)*

la main dominante est dans la configuration « Aile » (avion)

Geste 2 : « L'avion décolle »

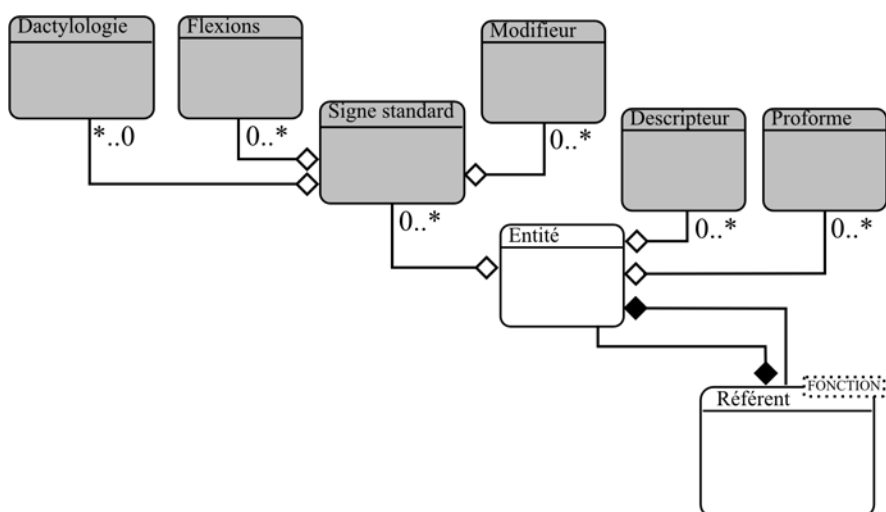
*Straight\_line(main\_dominante, up-left)*

mouvement de la main dominante en haut à gauche

De manière plus générique, les relations spatiales sont réalisées à travers le mécanisme de référencement spatial précisé dans (Gibet et al., 2001), complété par la spécification morphémique décrite dans la section 5.

#### 4.4. Modèle du discours

Les entités constitutives du discours peuvent être sélectionnées et agencées lors d'un processus sémantique et sémiotique qui n'est pas détaillé ici. Un modèle global du discours exprimé sous la forme d'un diagramme de classe a été proposé par (Héloir et Kervajan, 2007). Ce modèle prend en compte le contexte du discours, et les différents types de référents gestuels composant une entité, comme l'illustre la figure 5.



**Figure 5.** Relation entre l'entité, son référent et les signes associés

Une entité peut être représentée par son signe standard, mais peut également être référencée dans l'énoncé du discours par une forme spécifique de la main appelée *proforme*, ou par un descripteur spécifique qui permet de l'identifier. L'espace de signation est en général exploité dans le cadre de ce processus d'identification.

#### 5. Spécification des gestes de la langue des signes

Afin d'aborder le problème de la génération du mouvement, nous proposons dans cette section un modèle segmental ainsi qu'un modèle de spécification des gestes de la langue des signes. Le modèle segmental doit permettre d'identifier les différents articulateurs impliqués dans le mouvement, ainsi que les unités segmentales qui caractérisent les phases du mouvement significantes ou non. Le modèle de spécification est capable de décrire de manière concise et exhaustive les mouvements élémentaires participant à la construction des référents des entités intervenant dans un énoncé.



### 5.1. Organisation segmentale des gestes de la langue des signes

Le point de départ du modèle de génération est une suite de référents, accompagnés de leurs flexions éventuelles et des relations qui lient les entités auxquelles les référents se rapportent. Ce modèle doit également être capable de spécifier des gestes de transition ne participant pas directement à la sémantique du message produit. En ce sens, nous nous appuyons sur les travaux réalisés dans le domaine de l'étude des gestes de communication (Kendon, 1993). Ces travaux se sont basés sur l'étude des séquences gestuelles dans leur globalité, à l'opposé des études dédiées aux langues de signes qui se sont concentrées sur le caractère lexical des signes, mettant en évidence l'opposition entre les parties dynamiques et les parties statiques intervenant lors de l'exécution des signes standard (Liddell et Johnson, 1990). En considérant la production de signes comme un flux continu d'informations, il est possible d'identifier des phases de mouvement qualitativement différentes. (Kita *et al.*, 1998) proposent une organisation que nous reprenons et adaptons pour la génération de gestes de la LS. (Vogler, 2003) considère également des phases spécifiques de préparation lors de processus de reconnaissance de gestes de l'ASL. Nous distinguons ainsi, parmi les phases non porteuses de sens, les phases de *préparation* et de *rétraction*, et parmi les phases porteuses de sens celles de *tenue* (H) et de *mouvement* (M) qui constituent les signes.

#### 5.1.1. Phases non porteuses de sens

Les phases de *préparation* permettent de placer dans la position initiale d'un geste les articulateurs mis en jeu dans la production de ce geste. Bien que ces phases aient été identifiées par des travaux dédiés à l'étude des gestes coverbaux, elles n'ont, à notre connaissance, été que peu décrites par la littérature consacrée à la langue des signes. L'introduction d'une phase spécifique liée à la préparation avant l'exécution d'un signe limite l'influence de la coarticulation<sup>4</sup> sur le signe à réaliser et facilite sa compréhension. La figure 6 illustre une phase de préparation entre deux signes standard.

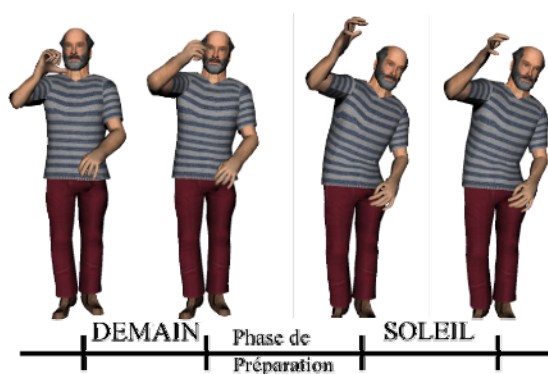
La spécification des phases de préparation est directe et s'appuie sur la (ou les) dernière(s) posture(s) de la phase précédant la préparation et sur la (ou les) première(s) posture(s) de la phase suivant la préparation.

Les phases de *rétraction* permettent de placer les articulateurs mis en jeu dans la production d'un geste dans une position neutre, ou de repos. Au sein d'une phrase gestuelle en langue des signes, cette phase de rétraction entre deux signes peut être fortement diminuée, à la fois d'un point de vue spatial et temporel. On parle alors de

---

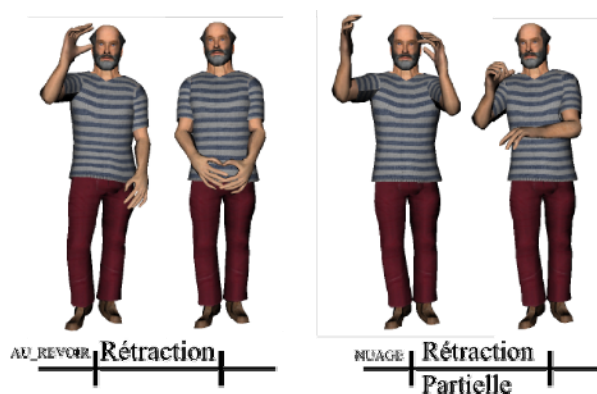
4. La coarticulation définie ici est similaire à celle définie en parole : elle correspond au pré positionnement des articulateurs en vue d'exécuter le mouvement à venir.

*rétraction partielle*. Elle peut dans certains cas être confondue avec la phase de préparation qui préfigure le signe suivant. La figure 7 illustre une rétraction partielle intervenant entre deux signes au sein d'une phrase et une rétraction complète à la fin d'une phrase.



**Figure 6.** Illustration d'une phase de préparation : « *Demain, soleil (il y aura)* »

La spécification des phases de rétraction s'appuie sur la (ou les) dernière(s) posture(s) du signe précédent et sur une posture neutre, ou de repos, qui est atteinte si la phase de rétraction est exécutée entièrement. Il n'est pas nécessaire que ces mécanismes de transition entre signes soient explicitement spécifiés. Ils peuvent être réalisés directement au niveau du moteur de génération, par un procédé de coarticulation qui est détaillé dans (Gibet et *al.*, 2001).



**Figure 7.** à gauche : phase de rétraction totale en fin de séquence (*AU REVOIR*) : les mains reviennent à une position neutre ; à droite :) phase de rétraction partielle (*NUAGES*) : on observe une légère pause avant d'amorcer le signe suivant

### 5.1.2. Phases porteuses de sens

Les phases gestuelles ayant une signification permettent la reconnaissance des signes ou syntagmes gestuels au cours de leur exécution. Ces phases sont caractérisées par des unités gestuelles de base qui peuvent être identifiées : à la fois des primitives dynamiques qui évoluent au cours du temps (phases de mouvement M) et des primitives statiques qui sont invariantes pendant la durée du signe (phases de tenue H). La décomposition des phases signifiantes en séquences de mouvements et de tenues s'inspire de la modélisation de Liddell, reprise par Vogler au niveau de la reconnaissance des gestes de la LS (Vogler et Metaxas, 2004). Ces phases peuvent être réalisées par des groupes d'articulateurs différents, ce qui nous conduira au niveau de la synthèse à identifier des groupes d'articulations associés à des contrôleurs différenciés : on distingue en particulier les configurations manuelles, les mouvements manuels, ou les expressions faciales. Au niveau qualitatif, ces phases se traduisent par une accentuation prononcée, caractérisée par des profils de vitesse et d'accélération spécifiques (Kita *et al.*, 1998).

Les phases de mouvement (M) se distinguent des phases de préparation et de rétraction par leur intention. En effet, elles sont associées à une modification de la dynamique des signes, qui peut être une trajectoire à suivre dans l'espace, un mouvement de pronation/supination<sup>5</sup> ou un changement de configuration manuelle à respecter. Parmi ces mouvements, on distingue les mouvements simples, qui ne présentent qu'un changement de configuration unique, et les mouvements complexes, qui présentent plusieurs changements successifs.

Les phases de tenues (T) sont également couplées à une intention : en effet, certains signes ne présentent aucun mouvement. Ce fait a été identifié par des travaux consacrés à la représentation et la reconnaissance de la langue des signes (Liddell et Johnson, 1989 ; Vogler et Metaxas, 2004). Il est à noter par ailleurs que la plupart des phases de tenue sont réalisées par des postures clés de la main.

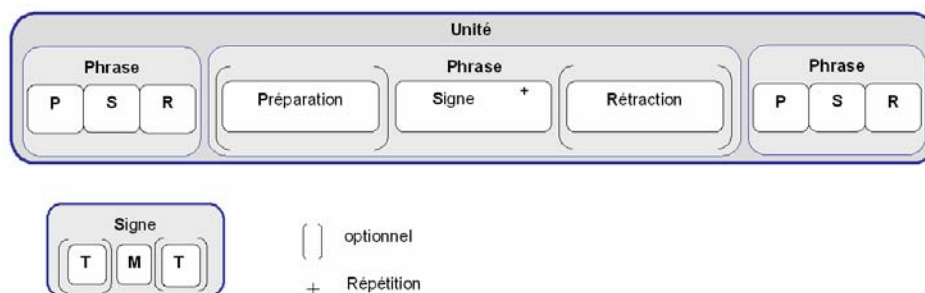
### 5.1.3. Organisation entre phases gestuelles

L'agencement des phases gestuelles de la LS peut être représentée par le schéma de la figure 8. Une unité gestuelle (*Unité*) est constituée par une suite de phrases (*Phrase*). Chaque phrase est composée d'une séquence de signes (*Signe*) précédée éventuellement d'une phase de préparation (*Préparation*) et suivie d'une phase de rétraction partielle ou totale (*Rétraction*). Chaque signe est constitué de séquences de phases mouvement (M) ou tenue (T).

---

5. Le membre supérieur est en pronation lorsque les paumes sont tournées vers l'arrière et le radius se trouve alors devant le cubitus ; la supination correspond aux paumes tournées vers l'avant.

Cette organisation peut être représentée à l'aide de la grammaire de réécriture hors contexte :

$$\begin{aligned} \langle \text{Unité} \rangle &:= \langle \text{Phrase} \rangle \{ \langle \text{Phrase} \rangle \}^6 \\ \langle \text{Phrase} \rangle &:= [ \langle \text{Préparation} \rangle ] \{ \langle \text{Signe} \rangle \} [ \langle \text{Rétraction} \rangle ] \\ \langle \text{Signe} \rangle &:= [ \langle \text{Tenué} \rangle ] \{ \langle \text{Mouvement} \rangle \} \langle \text{Tenué} \rangle \end{aligned}$$


**Figure 8.** Organisation de l'unité gestuelle et décomposition en phases

## 5.2. Spécification morphémique

Les modèles de spécification existants reposent sur une description morphémique des gestes des LS, qui convient, dans une certaine mesure, à la génération d'assemblages lexicaux. Ces modèles se concentrent sur la description du signe et, à l'exception de Losson qui spécifie des opérateurs syntaxiques modifiant les signes standard, ne prennent pas en compte les mécanismes de flexions opérés au niveau des phrases, des signes et des éléments constituants des signes. L'apport d'un modèle de spécification des signes est de plusieurs ordres. D'une part, ce modèle fournit une description non ambiguë des gestes, tout en répondant aux objectifs de concision et de compacité ; en particulier il doit comporter un jeu de descripteurs le plus réduit possible tout en garantissant des descriptions précises. D'autre part, il doit prendre en compte les aspects modulatoires des signes spécifiés en amont de la description morphémique. Ces modulations concernant aussi bien l'organisation spatiale des signes, que des modulations temporelles induites par les modificateurs spécifiés à un niveau d'abstraction supérieur. Elles sont l'une des manifestations des *structures de grande iconicité* présentées à la section 4.

À des fins de synthèse, nous proposons une spécification du geste qui soit capable de rendre compte des opérations de bas niveau articulaire qui interviennent dans la génération des gestes expressifs. Le modèle théorique que nous proposons s'appuie sur

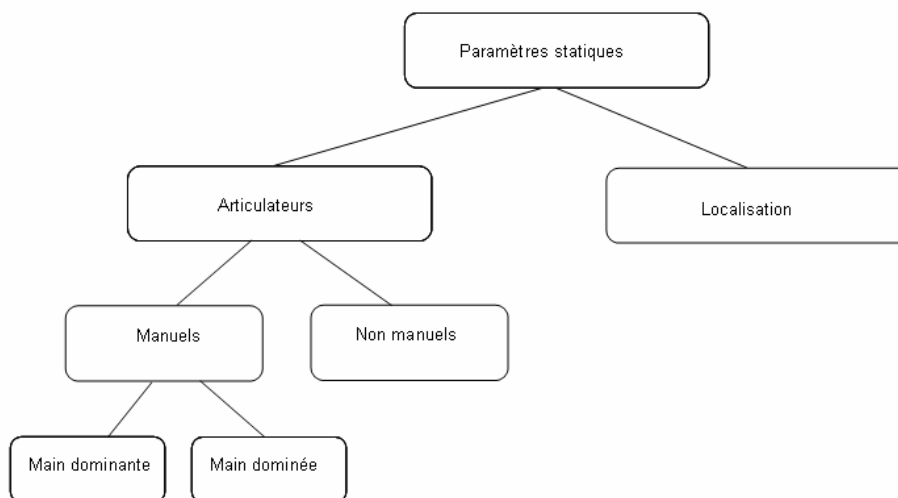
6. {} opérateur de répétition avec zéro ou plus d'une occurrence.

[] : élément facultatif.

les travaux de (Liddell et Johnson, 1989) pour la structure des signes et s'inspire des travaux de (Brentari, 1998) pour leur aspect variable. Le niveau de description articulatoire du geste va nous aider à établir une description complète et précise des phases gestuelles mises en jeu en langue des signes, quelle que soit la visée adoptée par le locuteur (iconique ou non). Par ailleurs, cette étape est nécessaire pour le contrôle du mouvement d'un avatar signeur.

L'idée de base du modèle de spécification proposé repose sur la séparation des caractéristiques statiques et dynamiques des mouvements lors de la production d'un geste.

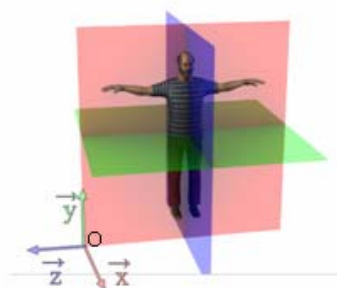
Les caractéristiques statiques regroupent d'une part la localisation du geste, et d'autre part les paramètres propres aux articulateurs manuels et non manuels, comme l'illustre la figure 9. Les caractéristiques dynamiques précisent les articulateurs impliqués dans les mouvements et décrivent ces mouvements sous la forme de trajectoires cinématiques. La séparation entre les paramètres statiques et dynamiques permet de mettre en relief l'importance du mouvement dans les gestes expressifs, et de gérer de manière pertinente les variations susceptibles d'apparaître au cours du discours.



**Figure 9.** Organisation des paramètres statiques au cours de la production d'un geste. Ces paramètres sont positionnés pendant la durée du signe et agissent au niveau de l'emplacement du signe et de la sélection d'articulateurs, manuels ou non manuels

### 5.2.1. Localisation

La localisation permet de spécifier le lieu du mouvement lors de l'exécution du mouvement. Elle est définie soit par un point dans l'espace de signation, soit par un vecteur normal à un plan, soit par un couple point/vecteur normal. Ainsi, le mouvement sera inscrit dans le plan spécifié, ou bien perpendiculaire à ce plan. Il est nécessaire d'ancrer ce plan dans l'espace de signation : pour ce faire, nous spécifions un point supplémentaire d'attache dans l'espace entourant le signeur. Ces plans sont parallèles le plus souvent aux trois plans (sagittal, frontal et transversal) généralement définis pour orienter l'humanoïde (figure 10).



**Figure 10.** Les trois plans caractérisant l'humanoïde ;  $(Oxy)$  : plan sagitta ;  $(Oyz)$  : plan frontal ;  $(Oxz)$  : plan transversal

Pour les zones de contact, nous définissons quatre segments qui sont ceux proposés par (Brentari, 1998) : la tête, le corps, les bras et les mains. Chaque segment est lui-même décomposé en plusieurs zones, ce qui définit quarante zones de contact pour le corps du signeur.

### 5.2.2. Paramètres manuels

La main est certainement le plus complexe des articulateurs. De ce fait, elle occupe une structure complexe dans notre modèle de spécification. Les articulateurs manuels regroupent la main dominante et la main dominée, respectivement nommées H1 et H2. Pour la majorité des signes, il existe une dépendance forte entre la main dominée et la main dominante. Cette dépendance peut être modélisée dans le cadre de notre description à l'aide d'un jeu réduit de contraintes. Lorsqu'un signe est bimanuel, il fait apparaître H2. Selon le mouvement, H2 a deux alternatives : copier le mouvement de H1 (avec un déphasage éventuel) ou ne pas bouger.

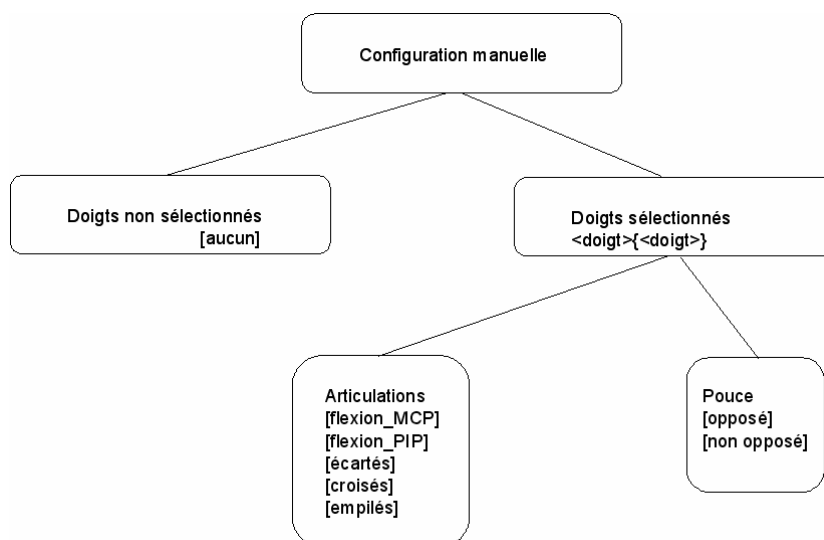
Le nombre et la nature des configurations manuelles varient suivant les auteurs. (Stokoe, 1972) en a répertorié 19, (Klima et Bellugi, 1979) 20. (Liddell et Johnson, 1989) en font état de 150. En LSF, (Braffort, 1996) propose 139 configurations dont 55 sont les principales configurations statiques. Un large éventail de configurations

manuelles peut être généré par un processus de sélection (de doigts et pouces) associé à des modifications (Lebourque et Gibet, 1999).

Une configuration manuelle est composée de doigts sélectionnés et non sélectionnés. Les doigts non sélectionnés sont ceux qui ne subissent pas de modification au cours du mouvement. Ils sont fléchis ou tendus (fléchis par défaut). Les doigts sélectionnés rassemblent deux types d'informations : une liste de doigts, et la spécification de leurs valeurs de flexion suivant leurs articulations.

Trois articulations sont mises en jeu pour chaque doigt : l'articulation métacarpo-phalangiennes, régie par le paramètre <MCP>; les articulations interphalangiennes proximales et interphalangiennes distales, spécifiées à l'aide du paramètre <PIP>, sachant qu'une relation linéaire lie les valeurs angulaires de flexions interphalangiennes distales et proximales :  $\phi_{DIP} = 2/3 \phi_{PIP}$ . Sans spécification, les valeurs aux articulations sont tendues. Il est nécessaire de dissocier le pouce des autres doigts sélectionnés au vu des propriétés articulaires du pouce par rapport aux autres doigts. Un paramètre de contact du pouce par rapport aux autres doigts est également spécifié.

La figure 11 présente la spécification de deux configurations types. Une description détaillée des articulations de la main est présentée dans (Héloir, à paraître).



**Figure 11.** Spécification des configurations manuelles

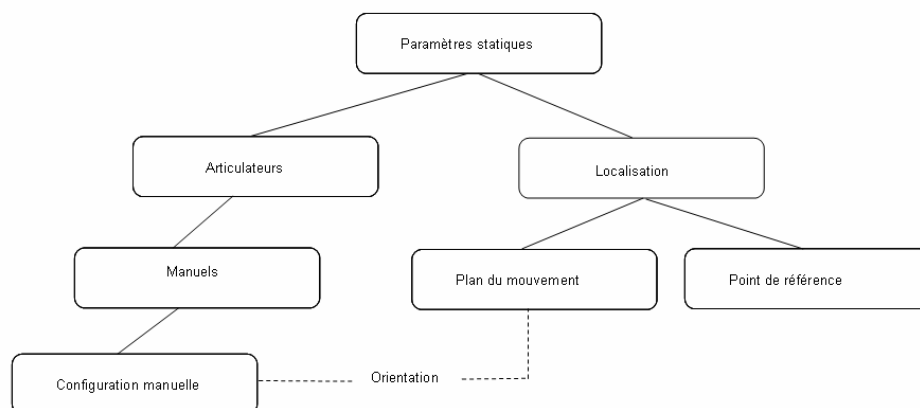
### 5.2.3. Orientation de la main

L'orientation porte sur un aspect légèrement différent des articulateurs classiques en langue des signes. Si Stokoe n'a pas choisi de reconnaître l'orientation comme l'un des paramètres constitutifs des signes, il a été reconnu par (Battison, 1978). Conformément

au modèle de (Brentari, 1998), l'orientation est en mesure de produire une différenciation lexicale entre deux signes, mais n'a guère besoin d'être exprimée par un paramètre dédié. En d'autres termes, l'orientation doit cesser d'être considérée comme un paramètre morphémique, mais plutôt comme la conséquence d'un agencement de contraintes articulaires. Il est alors possible d'exprimer une orientation manuelle comme la conséquence d'une relation entre des emplacements sur la main et des contraintes de position localisées sur le corps ou dans des plans de l'espace de signation. La figure 12 illustre la spécification des orientations.

À la différence des précédents modèles de (Losson, 2000 ; Lebourque, 1998), l'orientation des mains n'est donc pas considérée comme un paramètre constitutif à part entière, mais comme la conséquence d'un jeu de contraintes. Cette gestion implicite de l'orientation permet d'obtenir des poses plus naturelles respectant les contraintes physiologiques du signeur.

Beaucoup de signes présentent, par ailleurs, des coordinations fortes entre la main dominante et la main dominée dans un signe bimanuel. Par exemple, l'attribut d'alternance permet de spécifier le déphasage temporel du mouvement du bras de la main dominée par rapport à celui de la main dominante. De même, l'attribut de symétrie rend compte de l'aspect symétrique du mouvement des bras lors de l'exécution d'un signe.



**Figure 12.** *Spécification des orientations*

#### 5.2.4. Paramètres non manuels

Les paramètres non manuels regroupent les mimiques faciales, la direction du regard, les mouvements du cou, les haussements d'épaules et l'orientation du buste. Pour l'instant, nous nous intéressons aux mouvements du corps (épaules, tête, buste) et à la direction du regard.



#### 5.2.5. Paramètres dynamiques

Ces paramètres sont définis à plusieurs niveaux. Le premier niveau précise les articulateurs impliqués dans le mouvement. Par exemple, un mouvement peut nécessiter l'exploitation de toutes les articulations du bras, à savoir l'épaule, le coude et le poignet. Il peut au contraire être exécuté en n'utilisant que certaines articulations. Au second niveau de spécification, il s'agit de décrire les trajectoires du mouvement.

Celles-ci peuvent être inscrites dans un plan idéalisé de l'espace de signation, ce dernier étant spécifié par les paramètres statiques précédemment définis : il s'agit alors d'un geste de tracé. Parmi ces gestes, nous distinguons différents types de trajectoires : rectiligne, curviligne, circulaire, et quelques mouvements complexes.

Ces trajectoires peuvent également être orientées selon la normale par rapport à un plan idéalisé dans l'espace de signation : il s'agit alors de gestes de pointage, rectilignes ou curvilignes.

### 5.3. Modélisation multicanale

Un aspect fondamental de la langue des signes est le parallélisme des gestèmes, qui se traduit par le fait que lors de la production des gestes, plusieurs primitives de mouvements sont exécutées de manière simultanée. Il peut s'agir des mouvements des deux bras, d'une modification de la configuration manuelle simultanée à un mouvement du bras, d'une mimique faciale réalisée en même temps que des mouvements manuels, etc. Cette cooccurrence des gestèmes dans les langues des signes a été identifiée comme caractère multilinéaire des langues des signes par (Cuxac, 2000).

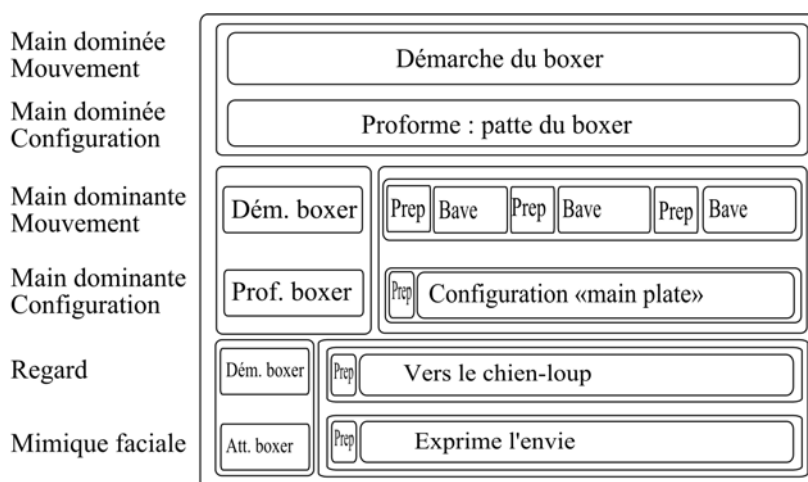
Un système de génération de la langue des signes doit donc prendre en compte cet aspect simultané et multilinéaire. Afin de limiter le nombre des combinaisons possibles des paramètres, il est nécessaire de séparer ces paramètres suivant plusieurs canaux qui peuvent être considérés comme indépendants les uns des autres. Notre modélisation suit par conséquent celle proposée par (Vogler, 2004) pour la reconnaissance et (Huenerfauth, 2006) pour la génération.

Nous considérons ainsi que les gestes s'organisent selon un ensemble de canaux participant à leur formation, chaque canal étant associé à une modalité spécifique et à un groupe d'articulateurs. L'agencement multilinéaire des différents canaux peut être illustré à l'aide d'un exemple proposé par Cuxac représentant une séquence exécutée en situation de grande iconicité (voir figure 13).

Dans cet exemple, la rencontre de deux chiens est narrée par le locuteur. L'un des chiens ressemble à un boxer, l'autre à un chien-loup. À chaque protagoniste est associé un ensemble de descripteurs le caractérisant. Chacun des chiens a également sa propre démarche. Ainsi, pour illustrer la démarche du boxer, les coudes sont écartés du corps et les poings sont serrés.

En situation de transfert personnel, le locuteur est capable de décrire un chien ressemblant à un boxer, marchant, regardant (un autre chien) avec envie et en bavant beaucoup. Cette scène a recours à différents traits expressifs simultanés qui s'expriment sur des articulateurs différents et se combinent pour former la scène finale.

Cet agencement révèle différents schémas de coordination entre les canaux qui peuvent changer au cours de l'exécution, en fonction du contexte du discours, et de l'état mental du signeur.



**Figure 13.** Partition du mouvement décrivant l'exemple tiré de (Cuxac, 2000) : « Tout en marchant, le boxer bave d'envie en regardant le chien-loup [...] », retranscrit par : « Tout en marchant (main gauche)/le boxer (main gauche + position du corps)/regarde le chien-loup (direction du regard)/avec envie (mimique faciale)/et bave (main droite)/beaucoup (mouvement de la main droite) »

## 6. Génération automatique de gestes de la LS

Nous proposons dans cette section un système de génération automatique de gestes de la LSF qui incorpore des aspects structurels propres aux LS, et permet de produire des séquences de gestes incluant une variabilité spatiale et temporelle. Cette variabilité est à rapprocher de la prosodie de l'oral, due en partie aux variations de débit, d'intensité, d'émotion, etc. On cherche ainsi à atteindre une certaine forme d'expressivité dans les mouvements générés, à la fois au niveau des signaux synthétisés, et au niveau du respect de la segmentation des gestes en unités signifiantes. Ce système permet actuellement de rejouer des séquences enregistrées, en modulant la prosodie des mouvements produits, à la fois d'un point de vue spatial et d'un point de vue temporel.

Il permet également de gérer la transition entre signes, en adaptant les phases transitoires (préparation ou rétraction) qui séparent les unités signifiantes des gestes.

### **6.1. Annotation des gestes**

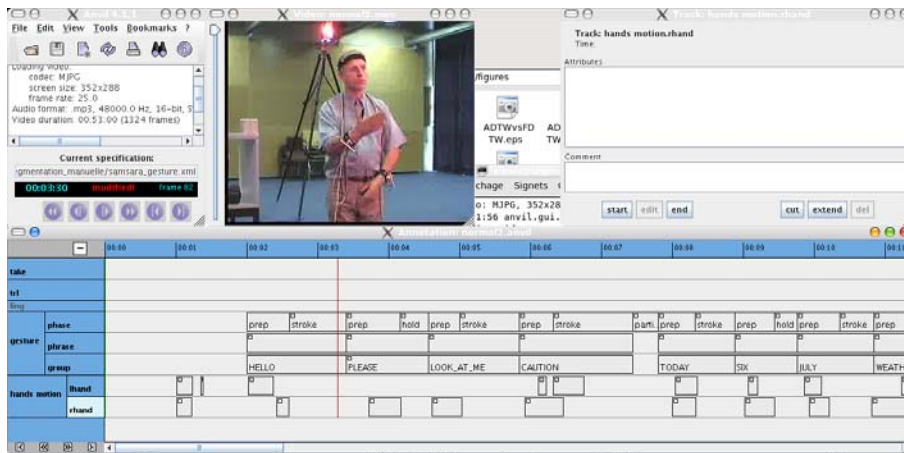
Plusieurs corpus ont été mis en œuvre à partir du choix de gestes spécifiques de la LSF. Un sujet sourd a participé à ces enregistrements, réalisés avec un système VICON MX (produit d'Oxford Metrics) composé de douze caméras infrarouges cadencées à 120 Hz, et deux gants de données Cyberglove pour la capture des configurations manuelles. Le premier corpus est constitué de vingt-deux phrases relatives à l'énoncé de bulletins météorologiques, avec un nombre limité de signes et des séquences d'enregistrement limitées à 60 secondes. Le second corpus est constitué de dix phrases traitant également de météo, réalisées avec différents styles. Parmi ces styles, nous avons retenu les styles neutre, énervé et désinvolte. Le troisième corpus contient huit séquences de messages relatifs à des incidents pouvant survenir dans une gare SNCF (cf. annexe 1). Enfin plusieurs séquences supplémentaires de signes isolés ont été réalisées, en particulier une séquence correspond à la signature d'un ensemble de villes françaises. Le contenu et la transcription de ces corpus sont donnés en annexe 1.

Afin d'étudier la succession de phases gestuelles constituant une phrase gestuelle complète, nous nous sommes appuyés sur la méthode de segmentation proposée par (Kita *et al.*, 1998). Cette méthode a initialement été appliquée pour l'annotation manuelle de gestes enregistrés sous forme de séquences vidéo. Elle s'est révélée insuffisante pour segmenter de manière non ambiguë les différentes phases du mouvement. Aussi avons nous eu recours à une segmentation manuelle assistée par une segmentation automatique. Selon Kita, les phases de mouvement sont caractérisées soit par un changement brutal de l'orientation de la main, soit par une discontinuité dans le profil de vitesse du mouvement décrit par le poignet. Comme l'identification des changements brutaux du poignet est difficile à repérer par la lecture manuelle d'enregistrements vidéo, nous avons enrichi le processus de segmentation décrit par Kita par l'introduction de deux canaux supplémentaires extraits de la représentation tridimensionnelle du mouvement. Ces canaux supplémentaires sont dédiés à l'évaluation du changement de configuration des mains au cours du temps. Les zones de changements ainsi détectées sont ensuite importées sous la forme de fichiers XML dans le logiciel d'annotation interactif ANVIL (Kipp, 2004).

Le reste du processus de segmentation a été réalisé à la main à partir des enregistrements vidéo des séances de captures. La segmentation manuelle a été réalisée le long des trois séquences de gestes capturés et a conduit à l'identification des phases présentées à la section 4. L'illustration 14 illustre l'interface d'annotation, selon les canaux que nous avons définis.

La segmentation des gestes a été réalisée suivant chacun des canaux explicités dans la section 5. Au niveau des configurations manuelles, nous avons utilisé un processus

automatique, basé sur une analyse en composantes principales, qui a été validé manuellement. Au niveau des mouvements des bras, nous avons identifié le début et la fin des gloses en détectant des pauses, et en tenant compte de la synchronicité avec les phases stables des configurations manuelles (Héloir et al., 2005).



**Figure 14.** Session d'annotation manuelle à l'aide de l'outil ANVIL. Les canaux lhand (main gauche) et rhand (main droite) sont extraits de manière automatique et assistent la segmentation

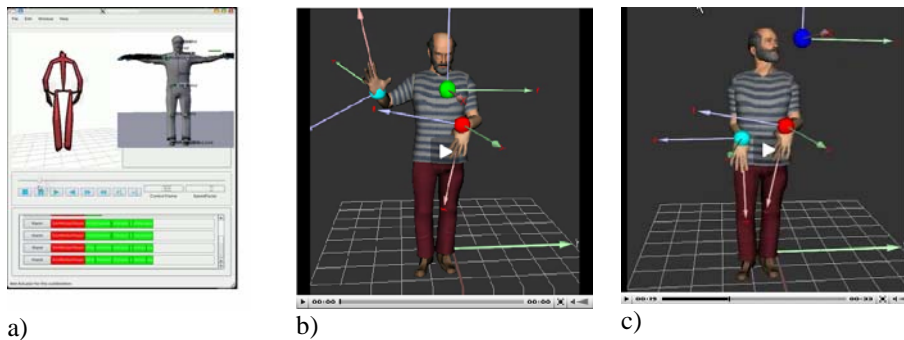
## 6.2. Contrôle du mouvement par cinématique inverse multicible

Le système de génération s'appuie sur une bibliothèque logicielle (SMR) développée dans l'équipe SAMSARA du VALORIA. Il permet, à partir d'une spécification segmentale multicanale (figure 15 a), d'engendrer les mouvements d'un signeur virtuel représenté à la figure 15 (b) et (c).

Selon le groupe d'articulateurs choisi, une méthode de contrôle d'animation est préférable à une autre. Par exemple, les mouvements « ambiants » traduisant la posture et l'attitude générale du signeur peuvent être synthétisés à l'aide de mouvements précapturés sur le corps du signeur. Notre méthode de génération de gestes comprend trois modes de génération : par rejeu pur, par interpolation ou par cinématique inverse. La méthode de cinématique inverse permet, à partir de la connaissance de la position de l'extrémité du bras, de remonter automatiquement aux angles des articulations de ce bras. C'est un problème dit redondant, car en général il existe une infinité de solutions, et il faut incorporer un certain nombre de contraintes au modèle pour trouver une « bonne » solution. Selon la spécification du geste, selon la phase de mouvement considérée et selon le groupe d'articulateurs mis en jeu, une méthode est préférable à une autre. Par exemple, les phases de préparation peuvent être avantageusement générées par interpolation.

Dans ce cas, il suffit de spécifier une posture de départ, une posture de fin et une durée pendant laquelle opérer la transition. S'ils sont disponibles, certains signes standard peuvent être directement rejoués à l'aide de leurs trajectoires articulaires. Ces trajectoires peuvent être déformées temporellement ou spatialement. Ces transformations s'appliquent, soit sur les séquences gestuelles globales, selon la méthode présentée dans (Héloir *et al.*, 2006a), soit sur les éléments signifiants des signes, selon la méthode présentée dans (Héloir et Gibet, 2007). Cette variabilité spatiale et temporelle liée au style est apprise sur les séquences de bulletins météorologiques, et reproduite ensuite sur les séquences d'incidents dans les gares. L'objectif est de produire de nouvelles séquences de gestes de la LSF, crédibles du point de vue de la naturalité des mouvements et de leur expressivité, tout en préservant la cohérence sémantique de l'ensemble.

Les séquences de gestes produites en visée iconique requièrent en revanche des méthodes de génération par cinématique inverse. En effet, l'iconicité des gestes de la LSF se traduit essentiellement par un ensemble de relations spatiales entre entités participant au mouvement produit (construction d'un espace 3D de signation, positionnement des référents dans cet espace). La méthode de cinématique inverse présentée dans (Gibet et Marteau, 1994) a été utilisée pour gérer la compression ou la dilatation spatiale des signes. Cette méthode permet d'opérer des transferts de taille ou de forme pour des signes iconiques ([NUAGE], par exemple). D'un point de vue sémantique, nous avons l'intention d'effectuer des substitutions sur les signes représentant les villes dans le corpus relatant des incidents dans les gares.



**Figure 15.** Génération de mouvement : a) Spécification interactive au niveau de l'agencement des phases du mouvement. b) Contrôle par cinématique inverse multicible : chaque main est contrôlée par une méthode d'optimisation qui minimise la distance de la main à la cible. À cette tâche d'atteinte de cible on associe une tâche visant à atteindre la cible avec une orientation spécifique de la main (trois autres cibles). c) La position de la tête ainsi que la direction du regard sont également gérées par une méthode similaire de cinématique inverse

### **6.3. Évaluation du système de génération**

Le système de génération que nous proposons a pour objectif d'aider la communication de personnes sourdes capables de s'exprimer en langue des signes française. L'avatar 3D animé doit pouvoir se substituer à des vidéos de personnes réelles, et ainsi permettre une plus grande liberté dans la visualisation et la production de gestes de la LSF. Dans le cadre de ces travaux, nous proposons une méthodologie d'évaluation en deux étapes : une évaluation objective puis une évaluation subjective. Lors de l'évaluation objective, nous quantifions l'erreur commise par le système en synthétisant les phrases des corpus après avoir opéré des transformations temporelles ou spatiales sur les données des corpus initiaux. Lors de la phase d'évaluation subjective, nous quantifions dans un premier temps la qualité segmentale de notre système de génération, puis nous déterminons la qualité de la compréhension.

Le protocole expérimental d'évaluation subjective pour la génération sera réalisé à travers une interface Web, grâce à la présentation de vidéos et à la mise en œuvre d'un questionnaire visuel réalisé en LSF. Nous nous intéressons principalement aux critères d'intelligibilité et de crédibilité du système. Le test d'intelligibilité est le plus important au vu de nos objectifs. Il s'agit dans un premier temps d'évaluer la qualité segmentale de la synthèse, et dans un second temps d'évaluer la capacité à comprendre des phrases en LSF. Le critère de crédibilité vise à qualifier la synthèse, d'un point de vue de l'attractivité des animations produites et de la qualité des styles engendrés.

Les tests d'intelligibilité portent sur l'identification de signes (ou gloses) isolés ou de phrases, ceci pour chacune des thématiques (bulletins météorologiques et séquences d'incidents dans une gare). Les sujets auront un bon niveau de connaissance de la LSF. À chaque test, la thématique est précisée. Les séquences animées apparaissent en ordre aléatoire, afin de minimiser les effets d'apprentissage. Après visualisation d'une séquence, le sujet doit cocher une réponse parmi plusieurs solutions proposées. En cas d'hésitation, ou de non-identification, le sujet a la possibilité d'indiquer les raisons de la non-reconnaissance : signe non connu, certaines caractéristiques non perceptibles (mouvement, forme de la main, expression faciale), dynamique non naturelle, problème de coordination entre modalités, angle de vue peu adapté, problème de rendu 3D, etc.

Les tests de crédibilité sont appliqués dans un premier temps aux variations de styles. Le questionnaire proposé sera réalisé en LSF. Pour chaque séquence de mouvement il sera possible de qualifier le style ou l'émotion correspondant à une séquence gestuelle (par exemple « en colère » pourra être associé à « énervé » ou « dynamique »). Dans un second temps, les expériences précédentes pourront être réitérées en proposant différents personnages (masculin, féminin), ainsi que différentes morphologies et habillages 3D d'humanoïdes virtuels.

Les premiers tests concernant la compréhension des messages produits en LSF et la reconnaissance de style sont en cours. Ils sont effectués sur les corpus présentés en annexe 1.

## 7. Conclusion

Cet article décrit une chaîne de spécification et de génération du mouvement de signeurs virtuels, qui s'appuie sur une analyse des mécanismes linguistiques impliqués dans les gestes de la langue des signes française, ainsi que sur l'étude segmentale de mouvements capturés. L'approche développée pose les fondements théoriques nécessaires pour qu'un certain nombre de mécanismes de modulation inhérents aux langues des signes soient pris en compte pour la génération et la transformation des gestes. Ces mécanismes résultent de caractéristiques propres au référencement spatial, aux variations géométriques et dynamiques inhérents aux gestes des LS. Les structures iconiques largement présentes dans les LS utilisent de façon combinée ces différents processus de modulation. Ainsi, une représentation semi-formelle des signes permet de définir à un plus haut niveau d'abstraction des opérateurs susceptibles de modifier les gestes de manière spatiale et temporelle dans des séquences discursives.

La spécification proposée est appliquée à la génération interactive de gestes de la langue des signes française. La synthèse permet, à travers une méthode sensorimotrice d'inversion cinématique, de prendre en compte les transitions entre unités gestuelles signifiantes, facilitant ainsi l'adaptation des mouvements à des contextes de discours variables. Il est possible également de déformer temporellement les mouvements synthétisés de manière à leur donner une certaine expressivité et améliorer leur plausibilité.

Ce travail a mis en évidence l'importance de travailler conjointement sur des représentations linguistiques et sémiologiques des langues des signes, et des représentations gestuelles construites à partir de l'agencement d'unités morphémiques de base. Nous espérons dans un futur proche mettre en œuvre une approche formelle permettant la traduction automatique d'une description linguistique de gestes de la LSF vers un ensemble de contraintes géométriques, cinématiques et temporelles, directement interprétables par le modèle d'animation. Une évaluation plus poussée devrait également permettre de valider les hypothèses mises en avant dans cet article.

## 8. Bibliographie

- Aznar G., Dalle P., « Computer Support for SignWriting Written Form of Sign Language ». In: *Proc. of 4<sup>th</sup> International Conference on Language Resources and Evaluation, LREC 2004*, pp. 109–110, 2004.
- Battison R., *Lexical borrowings in American Sign Language*. Silver Spring, 1978.
- Braffort, A. Reconnaissance de gestes en Langue des Signes Française, thèse de doctorat, université de Paris Sud, 1996.
- Braffort A., Lejeune F., « Spatialized semantic relations in French Sign Language : Towards a Computational Modelling ». In: *Gesture in Human-Computer Interaction and Simulation, GW 2005, Revised Selected Papers, Lecture Notes in Computer Science*, Vol. 3881, S. Gibet et al eds., Springer Berlin, pp. 37-48, 2006.

- Brentari D., *A prosodic model of sign language phonology*, Bradford Book, 1998.
- Cuxac C., La Langue des Signes Française (LSF) – Les voies de l’iconicité . *Faits de Langues*, vol. 15-16, ed. Ophrys, Paris, 2000.
- Filhol M., Braffort A., « A Sequential Approach to Lexical Sign Description », *Workshop on the Representation and Processing of Sign Languages, LREC 2006*, Genova, Italy.
- Gibet S., Lebourque T., Marteau P.F., « High level Specification and Animation of Communicative Gestures ». *Journal of Visual Languages and Computing*, Vol. 12(6), pp. 657-687, 2001.
- Gibet S., Marteau P.F., « A Self-organised Model for the Control, Planning and Learning of Nonlinear Multivariable Systems Using a Sensory Feedback », *Journal of Applied Intelligence*, Springer / Kluwer Academic Publisher, Vol. 4, Boston, pp.337-349, 1994.
- Héloir A., Courty N., Gibet, S., Multon F., « Temporal Alignment of Communicative Gesture Sequences ». *Computer Animation and Virtual Worlds*, Vol. 17, pp. 347-357, 2006a.
- Héloir A., Gibet S., Multon F., Courty N., « Captured Motion Data Processing for Real Time Synthesis of Sign Language », *In: Gesture in Human-Computer Interaction and Simulation, GW 2005, Revised Selected Papers*, Lecture Notes in Computer Science, Vol. 3881, S. Gibet et al eds., Springer Berlin, pp. 168-171, 2006b.
- Héloir A., Kervajan L., « Hiérarchisation des variations pour la génération de gestes expressifs de communication », *In: TALN 2007*, Toulouse, juin 2007.
- Héloir A., Gibet S., « A Qualitative and Quantitative Characterisation of Style in Sign Language Gestures », *Gesture Workshop 2007*, May 2007.
- Héloir A., Système de communication par agent virtuel – Aide à la communication des personnes sourdes, PhD thesis, Université de Bretagne Sud, à paraître.
- Huenerfauth M.J., Generating American Sign Language Classifier Predicates for English-to-ASL Machine Translation. PhD thesis, University of Pennsylvania, 2006.
- Kendon. A., *Human gesture.*, T. Ingold and Gibson K., editors, Tools, Language and Intelligence, 1993.
- Kennaway R., « Experience with and Requirements for a Gesture Description Language for Synthetic Animation », *In: Gesture-based Communication in Human-Computer Interaction, GW 2003, Selected Revised Papers*, Lecture Notes in Computer Science, Springer Berlin, Vol. 2915, pp. 300-311, 2004.
- Kipp, M., Gesture Generation by Imitation - From Human Behaviour to Computer Character Animation, PhD thesis, 2004.
- Kita, S., van Gijn, I., van der Hulst, H., « Movement Phase in Signs and Co-speech Gestures, and their Transcriptions by Human Coders ». *In: Proceedings of the International Gesture Workshop on Gesture and Sign Language in Human-Computer Interaction, GW 1997*, Lecture Notes In Computer Science, Vol. 1371, I. Wachsmuth et al eds., Springer-Verlag London, UK, pp. 23-35, 1998.
- Klima E. S. & Bellugi U. The signs of language. Harvard University Press, 1979.



- Lebourque T., Synthèse de mouvements naturels - Application à la Langue des Signes Française, PhD thesis, Université de Paris Sud, novembre 1998.
- Lebourque T., Gibet S., « A Complete System for the Specification and the Generation of Sign Language Gestures ». In: *Gesture-based Communication in Human-Computer Interaction, GW'99*, Lecture Notes in Computer Science, Vol. 1739, A. Braffort et al eds., Springer Berlin / Heidelberg, pp. 237-238, 1999.
- Lenseigne B., Dalle P., « Using Signing Space as a Representation for Sign Language Processing ». *Gesture in Human-Computer Interaction and Simulation, GW 2005*, Revised selected papers, Lecture Notes in Computer Science, Vol. 3881, S. Gibet et al eds., Springer Berlin, pp. 256-260, 2006.
- Liddell S.K., Johnson R.E., American Sign Language: the phonological base. *Sign Language Studies*, 64, pp. 195-277, 1989.
- Losson O., Modélisation du geste communicatif et réalisation d'un signeur virtuel de phrases en langue des signes française. PhD thesis, Université de Lille, 2000.
- McNeill D., *Hand and Mind – What Gestures Reveal about Thought*. The University of Chicago Press, Chicago, IL, 1992.
- Moody B. *La langue des signes. Dictionnaire bilingue élémentaire*. Vol. 1 et 2. Ellipses, 1993.
- Prillwitz S. L., Leven R., Zienert H., Zienert R., Hanke T., and Henning J. HamNoSys. Version 2.0. International Studies on Sign Language and Communication of the Deaf, 1989.
- Sallandre M.A., Tentative de catégorisation dans le cadre d'une grammaire de l'iconicité, PhD thesis, Université de Paris VIII, 2003.
- Stokoe W.C., *Semiotics and Human Sign Language*. Mouton, The Hague, 1972.
- Sutton V., « The SignWriting Literacy Project. Impact of Deafness on Cognition ». In Proc. of AERA Conference, 1998.
- Vogler C., « American Sign Language Recognition: Reducing the Complexity of the Task with Phoneme-Based Modeling and Parallel Hidden Markov Models », PhD Thesis, May 2003.
- Vogler C., Metaxas D., « Handshapes and Movements: Multiple-Channel American Sign Language Recognition », In: *Gesture-based Communication in Human-Computer Interaction, GW 2003, Selected Revised Papers*, Lecture Notes in Computer Science, Springer Berlin, Vol. 2915, pp. 247-258, 2004.

## **Annexe 1 – Corpus utilisés dans le cadre des travaux présentés**

### **Corpus météo**

#### ***Texte initial***

Bonjour. Attention, s'il vous plaît, regardez-moi.  
Aujourd'hui six juillet, voici le temps prévu.  
Le matin, des nuages traverseront la Bretagne.  
L'après midi, il va pleuvoir.  
Demain, le soleil brillera.  
Il fera chaud et sec.  
Ce sera le moment d'aller nager.  
Vendredi, il y aura des orages violents.  
Le vent soufflera très fort.  
La nuit, il y aura des orages violents en mer.  
Le matin, il y aura du brouillard en mer.  
À bientôt

#### **Texte transcrit**

Bonjour/s'il vous plaît/regardez-moi/attention.  
Aujourd'hui/6 juillet/le temps/prévu.  
La Bretagne/le matin/des nuages traversent.  
L'après midi/il pleut.  
Demain/le soleil brillera.  
Chaud/sec/il fera.  
Nager/aller/ce sera le moment.  
Vendredi/orages violents/il y aura.  
Le vent soufflera fort.  
La nuit/en mer/il y aura des orages violents.  
Le matin/en mer/il y aura du brouillard.  
À bientôt.

### **Corpus sncf**

Le train – 1234 – est en cours de préparation,  
la voie de départ sera affichée dans dix minutes.  
Le train – 5678 – subit un retard de – 15 mn – en raison d'intempéries.  
Le train – 1122 – subit un retard de – 10 mn – en raison d'incidents techniques.  
Le train – 3344 – subit un retard de – 20 mn – en raison d'accident de personne.  
À la suite d'un incident technique le train – 1205 – prévu voie – 4 – partira de la voie – 2 –.  
Le train à destination de Paris ne desservira pas la gare de Laval.  
En raison de la présence d'un colis suspect, le train subit un retard de – 30 mn –.

## Annexe 2 – Langage de spécification des gestes

Éléments spécifiés dans (Gibet et al., 2001)

*location* ::= **point**(*direction*, [*distance*]) |  
**point**(*near*, *body-location* [, *direction*]) |  
**point**(*contact*, *body-location* [, *direction*])

*direction* ::= **left** | **right** | **forward** | **up** | **down** |  
**left-forward** | **left-up** | **left-down** |  
**left-forward-up** | **left-forward-down** |  
**right-forward** | **right-up** | **right-down** |  
**right-forward-up** | **right-forward-down** |  
**forward-up** | **forward-down**

*distance* ::= **proximal** | **medial** | **distal** | **extended**

*body-location* ::= **head** | **ear** | **eye** | **nose** | **mouth** | **chin** | **forehead** | **temple** | **top-head** |  
**cheek** | **shoulder** | **chest** | **arm** | **elbow** | **forearm** | **wrist** | **hand**

*movement-primitive* ::= *pointing-mvt* | *straight-line-mvt* | *curve-mvt* | *ellipse-mvt* | *wave-mvt* | *zigzag-mvt* |  
*spiral-mvt*

*pointing-mvt* ::= **pointing** (*target-location*)

*straight-line-mvt* ::= **straight-line** (*start-location*, *end-location*)

*etc.*

Éléments permettant de spécifier les transferts de taille et de forme

*plan* ::= **absolute\_plan**(**X** | **Y** | **Z**)  
**relative\_plan**(*direction*, *body\_segment*)

*body\_segment* ::= **torse** | **chest** | **neck** | **left\_clavicle** | **right\_clavicle** | **left\_humerus** | **right\_humerus** | |  
**left\_radius** | **right\_radius**

*form* ::= *hand\_shape*(*basic\_form*, *num\_hand*, *size*) | {forme statique de la main}  
*hand\_shape*(*basic\_vol*, *num\_hand*, *size*) |  
*drawing\_shape*(*basic\_form*, *num\_hand*, *size*) {forme dessinée par la main}

*basic\_form*::= **square** | **triangle** | **cercle** | **elliptic** |  
*basic\_vol*::= **cone**(*radius*, *height*) | **sphere**(*radius*)  
*size* ::= **small** | **medium** | **big**

Les éléments spécifiés en gras sont des éléments du langage de description.

Les éléments spécifiés en italique sont définis par un ensemble de valeurs précisées ci-dessus.

Les éléments entre crochets sont optionnels.