



HAL
open science

Les Gestes Non Manuels en Langue des Signes Française ; Annotation, analyse et formalisation : application aux mouvements des sourcils et aux clignements des yeux

Emilie Chételat-Pelé

► **To cite this version:**

Emilie Chételat-Pelé. Les Gestes Non Manuels en Langue des Signes Française ; Annotation, analyse et formalisation : application aux mouvements des sourcils et aux clignements des yeux. domain_other. Université de Provence - Aix-Marseille I, 2010. Français. NNT : . tel-00547777

HAL Id: tel-00547777

<https://theses.hal.science/tel-00547777>

Submitted on 17 Dec 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

DOCTORAT AIX-MARSEILLE UNIVERSITE

Délivré par

Université de Provence

THESE

Pour obtenir le grade de

DOCTEUR D'AIX-MARSEILLE UNIVERSITÉ

Formation doctorale :

Cognition, Langage, Éducation

Parcours Sciences du langage mention technologie du langage

Présentée et soutenue publiquement

Par

CHÉTELAT Émilie

Le Jeudi 25 mars 2010

TITRE :

LES GESTES NON MANUELS EN LANGUE DES SIGNES FRANÇAISE,

ANNOTATION, ANALYSE ET FORMALISATION :

APPLICATION AUX MOUVEMENTS DES SOURCILS ET AUX CLIGNEMENTS DES YEUX

Sous la direction de :

Annelies Braffort (LIMSI-CNRS, Orsay) et Jean Véronis (Université de Provence)

JURY :

Annelies Braffort (Directrice)

Jean Véronis (Directeur)

Jean-Marc Colletta (Rapporteur)

Christian Cuxac (Rapporteur)

Patrice Dalle (Examineur)

Pour Alcide

LISTE DES SIGLES UTILISES	11
INTRODUCTION	1
CHAPITRE 1 : LA LANGUE DES SIGNES FRANÇAISE ET LES GESTES NON MANUELS	5
1. La Langue des Signes Française	6
1.1. <i>Approche historique</i>	6
1.1.1. Le conflit et ses origines	7
1.1.2. Situation actuelle et avancées apportées par la loi de 2005.....	12
1.2. <i>Approche linguistique</i>	14
1.2.1. Le signe et les paramètres	14
1.2.2. Le paramètre : une notion ouverte	15
1.2.2.1. Les paramètres sont trop rigides	16
1.2.2.2. Remise en cause du statut linguistique et du contenu des paramètres	18
1.2.3. Notion de structure de grande iconicité.....	19
1.3. <i>Conclusion</i>	20
2. Les Gestes Non Manuels (GNM)	22
2.1. <i>Fonctions</i>	22
2.1.1. Niveau lexical	23
2.1.1.1. Signe standard	23
2.1.1.2. Distinction de deux signes.....	24
2.1.1.3. Distinction entre l'Objet et son Action/Utilisation	25
2.1.1.4. Les modifieurs	26
2.1.2. Au niveau de l'énoncé.....	26
2.1.2.1. La modalité	27
2.1.2.2. Délimitation de transfert	29
2.1.2.3. Segmentation d'énoncés	29
2.1.3. Synthèse	30
2.2. <i>Paramètres ou Gestes non Manuels ?</i>	31
3. Conclusion	31
CHAPITRE 2 : CONTEXTE ET PROBLEMATIQUE DE LA RECHERCHE	35
1. Contexte de l'étude : Le Traitement Automatique des Langues	36
1.1. <i>Qu'est-ce que le TAL ?</i>	36
1.2. <i>Application à la LSF</i>	37
1.2.1. La génération automatique de LS.....	38

1.2.2.	Les travaux du LIMSI	40
1.3.	<i>La génération des GNM</i>	42
2.	Annotation des GNM	43
2.1.	<i>Comment décrire les GNM ?</i>	44
2.1.1.	Systèmes mis au point pour les LS.....	44
2.1.1.1.	Le système Bébian	44
2.1.1.2.	Le système D'Sign	45
2.1.1.3.	Le système HamNoSys.....	46
2.1.1.4.	Le système SignWriting.....	48
2.1.1.5.	Le système de Johnson et Liddell.....	52
2.1.2.	Système mis au point pour la description des émotions : FACS.....	52
2.1.3.	Réflexion critique autour des systèmes de notation.....	53
2.1.3.1.	Définition des attentes de l'annotation	54
2.1.3.2.	Confrontation entre les attentes et les systèmes	55
2.1.3.3.	Évaluation des systèmes.....	59
2.1.3.4.	Regard porté sur les annotations actuelles : remise en cause de la description de la posture finale	61
2.2.	<i>Quel support d'annotation ?</i>	62
2.2.1.	Les partitions	63
2.2.1.1.	La partition de Cuxac.....	63
2.2.1.2.	La partition de Bouvet	64
2.2.1.3.	La partition mixte de Fusellier-Souza et Boutora	65
2.2.1.4.	Bilan	66
2.2.2.	Les partitions multimédias	67
2.2.2.1.	Le logiciel Sign Stream.....	67
2.2.2.2.	Le logiciel Elan.....	68
2.2.2.3.	Le logiciel Ilex	69
2.2.2.4.	Le logiciel Anvil.....	69
3.	Conclusion	71
CHAPITRE 3 VERS UNE NOUVELLE APPROCHE DESCRIPTIVE		73
1.	Délimitation des contraintes et objectifs	74
1.1.	<i>Inscription dans un cadre théorique et pratique</i>	74
1.2.	<i>Contraintes dans le choix du corpus</i>	75
1.3.	<i>Définition des objectifs</i>	76
2.	Choix des supports de l'étude	77
2.1.	<i>Corpus retenu</i>	77
2.1.1.	L'origine du corpus LS-COLIN.....	77
2.1.2.	La réalisation du corpus LS-COLIN.....	78
2.2.	<i>Logiciel d'annotation retenu</i>	80

3.	Présentation de la méthodologie.....	83
3.1.	<i>Description qualitative</i>	<i>83</i>
3.1.1.	Grands principes.....	83
3.1.1.1.	Description du mouvement et non de la posture.....	83
3.1.1.2.	Décomposition des éléments à annoter	84
3.1.1.3.	Décomposition des mouvements	85
3.1.2.	Choix des attributs.....	87
3.1.2.1.	Distinction de trois amplitudes de réalisation	87
3.1.2.2.	Distinction des phases de réalisation	88
3.1.2.3.	Traitement des joues.....	90
3.1.2.4.	A chaque attribut sa couleur	91
3.1.3.	Bilan	93
3.2.	<i>Description quantitative et validation.....</i>	<i>94</i>
3.2.1.	Annotation sur les images de la vidéo.....	94
3.2.2.	Calcul sur les coordonnées des points	97
3.2.3.	Évaluation intermédiaire.....	100
3.2.4.	Analyse structurelle des GNM.....	101
3.2.5.	Bilan	102
4.	Conclusion	103
CHAPITRE 4 : RESULTATS DES ANALYSES.....		105
1.	L'annotation.....	106
1.1.	<i>L'annotation en quelques chiffres.....</i>	<i>106</i>
1.2.	<i>Synthèse des éléments traités</i>	<i>107</i>
1.3.	<i>Les données utilisées pour les analyses.....</i>	<i>109</i>
1.4.	<i>Conclusion.....</i>	<i>111</i>
2.	Présentation des résultats	112
2.1.	<i>Observation des mouvements des sourcils</i>	<i>112</i>
2.1.1.	Structure des mouvements : propriétés générales.....	114
2.1.2.	Typologie des mouvements des sourcils	117
2.1.2.1.	La segmentation d'énoncés (23,44% des mouvements étudiés).....	118
2.1.2.2.	Les concessifs (22,48% des mouvements étudiés).....	118
2.1.2.3.	La mise en évidence du rhème (14,83% des mouvements étudiés)	120
2.1.2.4.	Les transferts personnels (14,35% des mouvements étudiés).....	121
2.1.2.5.	La problématisation (13,39% des mouvements étudiés)	122
2.1.2.6.	Les fausses questions (11,48% des mouvements étudiés).....	124
2.1.3.	Bilan	126
2.2.	<i>Observation des clignements des yeux</i>	<i>127</i>
2.2.1.	Définition et propriétés	127

2.2.2.	Observation de la structure des clignements du corpus	128
2.2.3.	Proposition de typologie des clignements	130
2.2.3.1.	La segmentation (31,5% des clignements).....	130
2.2.3.2.	La mise en évidence (17,5% des clignements)	133
2.2.3.3.	Le transfert (22% des clignements) :	136
2.2.3.4.	Le clignement aspectuel de répétition (9,5% des clignements).....	138
2.2.3.5.	Le clignement personnel (9% des clignements)	138
2.2.3.6.	Les fausses questions (5,5% des clignements)	139
2.2.4.	Bilan	140
2.3.	<i>Apport de nos résultats aux connaissances actuelles</i>	141
2.4.	<i>Vers une formalisation de la typologie</i>	142
2.4.1.	Modèle Zebedee : notions et discussion	143
2.4.2.	Conventions de description	144
2.4.3.	Les structures type :	148
3.	Conclusion	155
CONCLUSION ET PERSPECTIVES		157
1.	Conclusion	157
2.	Perspectives de poursuite de l'étude	158
2.1.	<i>Annotation</i>	158
2.1.1.	Extension des corpus annotés	158
2.1.2.	Automatisation d'une partie de l'annotation	160
2.1.3.	Précision pour l'annotation d'autres GNM.....	160
2.2.	<i>Analyse</i>	162
2.3.	<i>Validation à l'aide de tests perceptifs</i>	165
2.3.1.	Propriétés du visage de synthèse.....	165
2.3.2.	Objectifs des tests de perception.....	166
2.4.	<i>Résumé des perspectives</i>	167
REFERENCES		169
1.	Bibliographie	169
2.	Webographie :	178
2.1.	<i>Textes :</i>	178
2.2.	<i>Outils et ressources :</i>	179
ILLUSTRATIONS		180
1.	Les Images	180
2.	Les figures	181

REMERCIEMENTS.....	185
ANNEXES.....	189
1. Formalisation des catégories de clignements des yeux.....	189
1.1. <i>Le clignement Pré-prise de parole.....</i>	<i>189</i>
1.2. <i>Le clignement de segmentation de groupes syntaxiques.....</i>	<i>189</i>
1.3. <i>Les clignements de pauses.....</i>	<i>191</i>
1.4. <i>Le clignement de mise en évidence du rhème :.....</i>	<i>192</i>
1.5. <i>Le clignement lié à l'objet pointé.....</i>	<i>193</i>
1.6. <i>Le clignement lié aux fausses questions.....</i>	<i>194</i>
2. Les mouvements des sourcils.....	195
2.1. <i>Le concessif de milieu d'énoncé.....</i>	<i>195</i>
2.2. <i>Le concessif de début d'énoncé.....</i>	<i>196</i>
2.3. <i>Le concessif sur tout l'énoncé.....</i>	<i>197</i>
2.4. <i>La mise en évidence du rhème.....</i>	<i>198</i>
2.5. <i>L'objet pointé.....</i>	<i>199</i>

Liste des sigles utilisés

GM : Geste Manuel

GNM : Geste Non Manuel

LV : Langue Vocale (version écrite, voir note de bas de page, page 6)

LS : Langue Signée

LSF : Langue des Signes Française

TAL : Traitement Automatique des Langues

TALS : Traitement Automatique des Langues Signées

TP : Transfert Personnel

Introduction

Les Langues Signées (LS), utilisées par les communautés sourdes, sont des langues visuo-gestuelles, dont le message est transmis par les gestes et reçu par le canal visuel. Les langues vocales, quant à elles, sont des langues audio-phonatoires, leur message est émis via le canal phonatoire et reçu grâce au canal auditif. Les canaux utilisés par les LS sont donc différents de ceux utilisés habituellement et distinguent les LS des autres langues. Les connaissances linguistiques acquises après des années de recherches et d'études sur de nombreuses langues vocales sont difficilement transposables pour les LS qui deviennent ainsi un nouvel objet d'étude linguistique seulement à partir des années 60. Ce champ de recherche est donc plus récent, ce qui explique que les connaissances sur le fonctionnement linguistique des LS soient en constante évolution.

Les LS n'impliquent pas seulement des mouvements des mains, mais également des mouvements d'autres parties du corps (par exemple des épaules, du visage ou encore de la tête). Nous appellerons ces mouvements, les Gestes Non Manuels (GNM). La manière dont les GNM participent à la transmission du sens n'est pas complètement définie. S'ils ont longtemps été assimilés à des pantomimes, ils sont aujourd'hui encore souvent comparés à l'intonation des Langues Vocales à l'oral. Mais leurs fonctions sont plus complexes et nécessitent encore des études linguistiques.

Le présent travail se situe dans le contexte du Traitement Automatique des Langues (TAL) et plus précisément de la génération automatique de Langue des Signes Française (LSF) et des GNM en particulier.

Le TAL est une discipline qui vise à traiter, analyser ou transformer informatiquement des données linguistiques. A partir de ces données, quel que soit leur format (textuels, audio, etc.), un traitement automatique offre en sortie un document dont le format, la langue ou le contenu diffère du document d'entrée. Par exemple, un document audio peut être traité afin d'être proposé en version écrite et traduite dans une autre langue. De même, un traitement spécifique permet d'obtenir la réponse à une question posée par l'utilisateur grâce, entre autres, à l'indexation automatique de documents.

Une des branches du TAL, le Traitement Automatique des Langues Signées (TALS), se concentre sur la Langue des Signes et vise à maîtriser le traitement de données linguistiques dont l'entrée ou la sortie est en LS. Par exemple en entrée, la reconnaissance de LS tente d'extraire automatiquement les signes, et à terme le sens, d'un message signé. En sortie, la génération de LS (champ dans lequel se situe cette étude) tend à proposer, par l'intermédiaire d'un personnage de synthèse signant¹, des messages générés automatiquement en LS afin de permettre à la communauté sourde d'accéder aux informations de la vie quotidienne auxquelles ils n'ont pas toujours accès actuellement.

La génération automatique de LSF pose de nombreux problèmes dont la difficulté de générer les GNM. La génération automatique d'un énoncé, dans n'importe quelle langue, nécessite en amont une modélisation, c'est-à-dire une description formelle, de cette langue. Cette modélisation implique de disposer de suffisamment de connaissances linguistiques sur la langue en question. Or, les GNM ne peuvent être modélisés en raison du peu de connaissances formelles actuellement disponibles aussi bien sur des aspects linguistiques que sur leurs structures internes.

Ce travail vise à accroître les connaissances dont on dispose sur les GNM afin, à terme, de les générer automatiquement et en contexte.

Le premier chapitre présente notre objet d'étude. A travers l'histoire des sourds et de la LSF, nous expliquons l'origine des enjeux sociolinguistiques actuels et présentons la situation législative, sociale et linguistique. Nous nous intéressons ensuite aux GNM et plus particulièrement à leurs fonctions dans la langue. Nous proposons enfin de justifier le choix de notre terminologie.

¹ Signant signifie s'exprimant en LSF.

Le second chapitre permet de situer le contexte de ce travail et présente la discipline du TAL. Nous nous intéressons à son application aux LS, autrement dit le TALS, et plus précisément à la génération automatique de LSF en général et des GNM en particulier. Nous montrons que les connaissances dont on dispose à leur sujet ne permettent pas actuellement leur génération automatique, ce qui nous conduira à proposer une nouvelle étude centrée sur la structure interne des GNM. Une telle étude étant basée sur l'observation d'un corpus, nous listons les différents systèmes d'annotation de corpus qui permettent de décrire les GNM et concluons qu'aucun de ces systèmes ne dispose de la finesse suffisante à l'étude envisagée. La première étape est donc de créer une nouvelle méthodologie d'annotation permettant de décrire avec précision l'ensemble des GNM.

Le troisième chapitre présente notre proposition de méthodologie d'annotation. Il s'agit de préciser et de justifier nos choix relatifs aux outils utilisés pour l'étude (le corpus et le logiciel d'annotation), puis d'expliquer les grands principes sur lesquels nous avons centré notre méthodologie. Enfin, nous présentons notre système d'annotation et illustrons d'exemples.

Le quatrième chapitre apporte des précisions quant à l'annotation et aux données qui en découlent, puis présente les premiers résultats de nos analyses, qui concernent uniquement les paupières et les sourcils. L'objectif étant de générer les GNM, nous proposons une modélisation des mouvements des sourcils et des paupières.

Enfin, nous concluons en proposant des pistes de poursuite de notre travail qui concernent aussi bien l'annotation d'un plus vaste corpus qu'une extension de l'annotation à l'ensemble des GNMs (joues, épaules, etc.).

Chapitre 1 : La Langue des Signes Française et les Gestes Non Manuels

Ce premier chapitre présente la LSF et les GNM. Nous proposons dans une première partie une présentation globale de la LSF sous un angle historique puis linguistique. Dans une seconde partie, nous nous intéressons aux GNM et plus particulièrement à leur rôle et à leur influence à différents niveaux de la langue et pas seulement dans la constitution d'un signe. Ceci nous conduit à remettre en cause la nomination généralement utilisée pour dénommer ces éléments. Nous proposons alors une réflexion autour des différents termes utilisés justifiant la dénomination " Gestes Non Manuels ".

1. La Langue des Signes Française

La langue des signes Française (LSF) est une langue peu connue et reconnue encore de nos jours. Que ce soit sur un plan socio-juridique ou linguistique, chaque avancée en terme de reconnaissance ou d'acquisition de savoir est lente et se heurte notamment au conflit qui oppose les défenseurs des Langues Vocales² (LV), qui pensent que seules ces dernières sont des langues, aux défenseurs des Langues Signées (LS). Il faudra par exemple attendre 2005 pour une reconnaissance légale de la LSF³.

La première section aborde la LSF d'un point de vue historique en remontant à l'origine de ce conflit afin d'en cerner les conséquences aussi bien au niveau social que scientifique. Dans une seconde section, la LSF est abordée d'un point de vue linguistique afin de positionner notre travail par rapport à certains concepts.

1.1. Approche historique

Nous retraçons l'histoire des sourds afin de préciser l'origine du conflit opposant les défenseurs de l'oralisation aux défenseurs du bilinguisme (français – LSF) dans le contexte de l'éducation. Nous présentons ensuite les conséquences de ce conflit et la situation actuelle aussi bien au niveau social que scientifique.

² Les langues peuvent se présenter sous forme orale ou écrite. Les LS ne disposent pas à l'heure actuelle de forme écrite reconnue et utilisée. Ainsi, le terme LS fait référence directement à sa forme orale, et l'expression "Langue Vocale" fait directement référence à la forme écrite d'une langue vocale. Le seul cas ambigu est l'expression "langue orale", pour laquelle nous précisons s'il est question de langue vocale ou signée, à l'aide d'expressions telles que "forme orale de la langue vocale".

³ loi n°2005-102 du 11 février 2005, Art. L. 312-9-1. - La langue des signes française est reconnue comme une langue à part entière. Disponible à l'adresse : www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000809647&dateTexte=

1.1.1. Le conflit et ses origines⁴

Jusqu'au XV^{ème} siècle : La parole est la raison

De l'Antiquité au XV^{ème} siècle, la langue vocale est considérée dans beaucoup de civilisations comme le lien entre l'Homme et Dieu, elle est ce qui distingue l'Homme de l'animal et est donc jugée sacrée.

Chez certains philosophes, ce n'est pas le cerveau qui est associé à l'intelligence mais l'ouïe car c'est elle qui permet d'apprendre (l'apprentissage passant par l'oral) puis de réfléchir et enfin de communiquer grâce au langage. C'est pourquoi les sourds, n'ayant ni l'ouïe ni le langage, sont considérés comme des êtres dépourvus d'intelligence et ne pouvant en acquérir. Les premières traces de cette pensée se retrouvent dans les écrits d'Aristote (-400, -300).

“ C'est l'ouïe qui rend les plus grands services à la pensée, puisque c'est le langage qui est cause que l'homme s'instruit, et que le langage est perçu par l'ouïe (...). Voilà bien pourquoi, parmi les hommes qui de naissance manquent de l'un de ces sens, les aveugles-nés sont plus intelligents que les sourds-muets. ”

Aristote, “ Opuscules, plan de traité de la sensation et des choses sensibles ”, Chapitre 1⁵,
traduction de J. Barthélémy Saint Hilaire.

Les sourds profonds de naissance n'ont pas accès à la LV naturellement, considérant le lien entre l'ouïe et l'apprentissage de la parole (Mugnot et Neve, 2002). Au VI^{ème} siècle, le code Justinien classe les sourds en fonction de leur aptitude à parler. Les sourds profonds de naissance sont placés au niveau le plus bas. Ils sont en conséquence privés de tout droit civique et doivent être placés sous un curateur au même titre que les fous :

“ On doit donner des curateurs aux fous, aux sourds et aux muets et à ceux qui sont atteints de maladie chronique parce qu'ils ne peuvent pas s'occuper de leurs affaires. ”

Traduction du Code Justinien, Livre 1⁶

4 Lorsque la source des informations n'est pas citée dans cette section, elles sont issues de (Lane, 1979) et (Moddy, 1983).

⁵ Traduction disponible sur : <http://remacle.org/bloodwolf/philosophes/Aristote/tableopuscules.htm>

⁶ Accessible à l'adresse : <http://www.cis.gouv.fr/spip.php?article1530>

A cette époque, les LS sont rarement assimilées à des langues, car elles ne sont pas vocales. Cette idée est défendue par le courant oraliste qui sera en conflit permanent avec le courant des défenseurs des LS.

A noter quelques rares exceptions durant cette période notamment Saint Jérôme vers la fin du IV^{ème} siècle qui est le premier à évoquer la possibilité que la gestualité utilisée par les sourds peut leur permettre d'apprendre l'Évangile presque au même titre qu'une langue le permet.

Le XVI^{ème} siècle : Une première évolution

La possibilité d'éduquer un sourd va apparaître à partir du XVI^{ème} siècle. Il s'agit simplement d'apprendre aux plus privilégiés d'entre eux, issus de riches familles, les bonnes règles de vie et tenter de les faire accéder à la parole. Cet intérêt pour l'éducation des sourds est motivé en partie par le fait que les "sourds muets" n'avaient pas le droit de posséder des biens donc pas d'héritage possible. Certaines riches familles espagnoles en mal d'héritiers Entendants ont décidé de démutiser leur seul héritier sourd afin de ne pas perdre leur nom et leurs biens en attendant la nouvelle génération. Le premier ouvrage sur la pédagogie des sourds est écrit dans cet élan par Juan Pablo Bonnet en 1620. L'apprentissage de la parole ne passait pas par une quelconque langue gestuelle mais a permis d'envisager qu'il était possible d'éduquer les sourds.

Le siècle des lumières : Émergence de la langue des signes

En 1710, Étienne de Fay, sourd de naissance, est le premier à enseigner une langue gestuelle à d'autres sourds. C'est cependant principalement en 1760 qu'émerge l'idée que les gestes peuvent exprimer la pensée grâce à l'Abbé de l'Épée, qui propose une éducation en Langue des Signes et ouvre à Paris la première école gratuite du monde dont l'enseignement se fait en LSF. Les sourds obtiennent alors, à défaut d'une reconnaissance pleine et entière, la possibilité d'acquérir un savoir et une éducation. Des représentations publiques, où l'Abbé de l'Épée interroge ses élèves sourds sur toutes sortes de matières (telles que la religion et les mathématiques), sont organisées dans toute la France. Ainsi, il souhaite démontrer que les sourds sont aussi intelligents que les Entendants et peuvent acquérir les mêmes connaissances. A la mort de l'Abbé de l'Épée, ses élèves poursuivent

son travail en France mais aussi dans le reste du monde (Laurent Clerc aux États-Unis par exemple⁷).

Les oralistes, défenseurs de la LV, sont affaiblis par les efforts et la renommée durement acquise de l'Abbé de l'Épée mais sont toujours convaincus de l'importance de l'oralité.

Le XIX^{ème} siècle : La répression

Au XIX^{ème} siècle, le courant oraliste se fait de plus en plus entendre. Il considère les langues des signes non pas comme des langues mais comme des “ singeries ” et revendique que les sourds, puisqu'ils n'ont pas de raisons d'être muets, peuvent apprendre à parler grâce aux médecins et aux professeurs.

De nombreux guérisseurs sont convaincus de pouvoir “ réparer la surdité ” et mènent des expériences comprenant toutes sortes de mutilations de l'oreille. C'est d'ailleurs en tentant d'améliorer un appareil servant à stimuler les restes auditifs que Bell, dont la femme était devenue sourde, inventa le téléphone. Il soutiendra ensuite les oralistes, allant même jusqu'à proposer un compte rendu destiné à l'Académie nationale des sciences dans lequel il se prononcera en faveur de mesures eugéniques envers les sourds :

“ Ceux qui pensent, comme moi, que la reproduction d'une race d'individus déficients serait une énorme calamité pour la société, examineront de plus près les causes des mariages entre sourds, avec l'intention d'y trouver un remède (...) la méthode la plus prometteuse pour enrayer le mal semble relever de l'adoption de mesures préventives. ”

(Bell, 1883) traduit par (Lane, 1979).

Les connaissances en médecine sur la surdité sont limitées, comme le mentionne le docteur Esser, faisant état de la conviction du docteur Itard qui pensait que le tympan ne vibrait pas et qu'il suffisait de le percer pour retrouver l'ouïe (Esser, 1831).

⁷ L'American Sign Language (ASL) et la LSF sont proches de par leurs origines communes.

Dans les écoles spécialisées, qui ne sont plus gérées par le ministère de l'éducation mais par celui de la santé, l'idée d'oraliser les sourds prend de l'ampleur. De nombreux directeurs adoptent le système allemand de Heinicke qui a fondé la première école oraliste pour sourds en 1755. En France, tandis que certains professeurs sourds tentent de s'opposer au projet d'oralisation généralisée, l'influence allemande s'étend et de nombreuses écoles oralistes ouvrent leurs portes. C'est toutefois durant cette période que Bébien propose un système de notation des LS (Bébien, 1825) et qu'a lieu le premier mariage de deux sourds en 1844, faits exceptionnels pour l'époque.

En 1880 le congrès de Milan (impliquant une mise en scène présentant de faux sourds profonds qui parlent) interdit les langues des signes dans toutes les écoles d'Europe justifiant cette décision par le fait que les LS éloignent l'enfant de la parole, que les personnes utilisant les gestes auraient plus de chance de contracter la tuberculose ou considérant également que la parole serait plus esthétique :

“ Le sourd-muet est en quelque sorte humanisé, qu'on nous permette cette expression, par la parole. En effet, on a remarqué qu'elle ennoblit tout son être en donnant à sa figure une expression plus calme, plus sereine et en tempérant les mouvements désordonnés de ses bras et de ses mains, si disgracieux parfois et si peu faits pour attirer la sympathie même des personnes qui vivent avec lui. ”

(Ministère de l'instruction publique, 1881, p. 306)

On assiste alors à une vague de répression (ouvrages traitant des LS brûlés, mains attachées dans le dos pour empêcher les sourds de signer, écriteau “ vive la parole ” à porter au cou lors de punition) accentuée par la seconde guerre mondiale (déportation, stérilisation et extermination des sourds) (Lane, 1979).

Malgré les démonstrations de l'Abbé de l'Épée, les sourds continuent bien souvent d'être considérés comme des personnes inférieures et peu intelligentes :

“ Tout le monde sait que les sourds-muets sont des êtres inférieurs à tous égards : seuls les professionnels de la philanthropie ont déclaré que c'étaient des hommes comme les autres. ”

(Regnard, 1902, p.3)

Charles Richet, prix Nobel de médecine déclare d'ailleurs en 1919:

“ A force d'être pitoyables, nous devenons barbares. C'est barbarie que de forcer à vivre un sourd-muet, un idiot, un rachitique... Ce qui fait l'homme, c'est l'intelligence. Une masse de chair humaine, sans intelligence humaine, ce n'est rien. Il y a de la mauvaise matière vivante qui n'est digne d'aucun respect ni d'aucune compassion. Les supprimer résolument, ce serait leur rendre service car ils ne pourront jamais que trainer une misérable existence. ”

(Richet, 1919, p.164)

Les LS, interdites, n'ont plus été étudiées ni enseignées en Europe pendant près d'un siècle mais ont continué secrètement d'être employées et donc d'évoluer.

Vers une reconnaissance

Le milieu du XX^{ème} siècle marque lentement un tournant pour les sourds français et la LSF, en revenant petit à petit sur quelques interdictions et en reconnaissant progressivement quelques droits aux sourds. En effet, en 1959 par exemple, la loi Buron autorise les sourds à passer leur permis de conduire. En 1977, le Ministère de la Santé abolit l'interdiction de signer dans les lieux publics qui sera complétée, en 1991, par la loi Fabius autorisant le choix d'une éducation bilingue Français - Langue des Signes. Il faudra attendre 1996 pour que le premier accueil en LSF soit mis en place à l'hôpital, 2000 pour que la présence obligatoire d'un interprète lors de la convocation d'une personne sourde en justice soit ajoutée au code pénal et 2005 pour une reconnaissance officielle de la LSF.

Les LS en tant que langues naturelles

Tandis que la LSF n'est reconnue officiellement que depuis quatre ans, les linguistes ont démontré le statut de langue naturelle des LS depuis plus de trente ans. En effet, Stokoe (1960 ; 1972) aux États-Unis (donc moins concerné par l'interdiction du congrès de Milan), a démontré la présence de la double articulation dans les LS en 1960 ce qui, d'après Martinet (1960), en légitime le statut de langue naturelle.

Cette démonstration a marqué un véritable tournant dans la linguistique, même si elle est aujourd'hui contestée, point développé dans la section 1.2.2, offrant à la recherche un nouveau type de langues mystérieuses dont le fonctionnement visuo-gestuel différerait des connaissances acquises auprès des langues audio-phonatoires.

Rapidement l'étude des LS a partagé la communauté scientifique sur l'approche à adopter. En effet les chercheurs se sont divisés en deux écoles : ceux analysant les LS à l'aide des outils et des connaissances des LV (à des fins de comparaison) et ceux abordant les LS comme un tout nouveau type de langue qui nécessite son propre vocabulaire et ses propres outils.

En d'autres termes, tandis que les premiers tentaient d'appliquer aux LS les concepts en usage pour les LV, en rejetant tous les aspects ne trouvant pas de correspondance (telle que l'iconicité, voir section 1.2.3), les seconds inventaient de nouveaux concepts propres aux LS. Ces approches et leurs origines sont abordées notamment par Cécilia Hutter, doctorante en linguistique à l'Université de Rouen, lors de la conférence d'octobre 2007 à l'URAPEDA de Bretagne⁸ :

“ Cette première démonstration (i.e. la double articulation) est représentative d'une démarche scientifique qui s'appuie sur les caractéristiques des langues orales pour étudier les langues des signes ; d'autres recherches tentent aujourd'hui d'étudier ces langues des signes à partir de leurs spécificités. Ces deux approches façonnent encore la recherche sur les langues des signes aujourd'hui et apportent des éclairages différents sur ces langues, même si ces deux démarches reconduisent toujours la distinction entre langues orales et langues des signes, que ce soit en les comparant ou en n'étudiant que les langues des signes dans leurs particularités. Cette approche des langues des signes par rapport aux caractéristiques des langues orales est liée à la linguistique “traditionnelle” et aux représentations communes de la langue qui ont longtemps associé langue et oralité. (...). Il en résulte que les langues des signes doivent sans cesse justifier leur statut de langue.”

Le premier courant a été longtemps majoritaire ce qui explique entre autres le peu de connaissances dont on dispose concernant le non manuel, alors considéré comme du mime donc extérieur à la langue.

1.1.2. Situation actuelle et avancées apportées par la loi de 2005

“...Pour que le siècle qui s'ouvre devant nous fasse du droit des sourds le devoir des Entendants”.

Présentation du rapport de et par Dominique Gillot, le 30 juin 1998

Tandis que les lois évoluent et prônent l'égalité des chances, que sous couvert d'être politiquement correct il est demandé de ne plus parler de “sourds” mais de “malentendants” ou “déficients auditifs”, la qualité de vie des sourds évolue peu.

⁸ Accessible à l'adresse : http://www.unapeda.asso.fr/article.php3?id_article=776

Selon l'UNISDA, (Union Nationale pour l'Insertion Sociale des Déficiants Auditifs) seul 10% des sourds ont accès aux études supérieures (compte rendu du congrès, Octobre 2005⁹) et plus de 80% sont illettrés (rapport Gillot, 1998¹⁰).

La population, à moins d'être confrontée au monde de la surdité, ne s'interroge pas sur les bienfondés des idées répandues. Ainsi, les convictions du passé, qui n'ont plus de raison d'être au regard des connaissances actuelles, se perpétuent encore (Dalle-Nazebi et Lachance, 2008). La loi du 11 février 2005¹¹, qui exige que tous " les établissements existants recevant du public doivent être tels que toute personne handicapée puisse y accéder, y circuler et y recevoir les informations ", est accueillie comme un réel progrès.

En effet, afin de respecter la nouvelle législation, les entreprises ont dû se sensibiliser à la surdité et des messages en LSF commencent à être visibles dans des lieux publics (notamment à la gare de l'Est à Paris¹²). Leur impact sur le grand public permet un recul des anciennes croyances (comme l'association déficience auditive/déficience mentale) de par les questionnements qu'ils suscitent.

Cette loi permet également à la LSF, après de nombreuses années de combat de la communauté sourde, d'acquérir le statut de langue. Cette avancée, si elle reste encore symbolique dans les faits, permet à la LSF d'obtenir une reconnaissance que la communauté sourde attendait depuis longtemps.

Enfin, l'impact de cette loi ne se cantonne pas à l'information du grand public ou à la reconnaissance de la LSF et devrait apporter en particulier davantage de moyens à la recherche. En effet, les travaux scientifiques sur la LSF, marginaux et surtout peu financés jusqu'alors, devraient se multiplier grâce aux besoins des entreprises de s'adapter à cette loi en trouvant un moyen de proposer des accueils en LSF. Les solutions humaines étant coûteuses et difficiles à mettre en place en raison du manque d'interprètes en France, les demandes concernent essentiellement la génération de LSF via un personnage de synthèse,

⁹ Accessible à l'adresse : <http://etsf.fnsf.free.fr/informationannonce2.htm>

¹⁰ Ce chiffre est toutefois à relativiser en raison du manque de clarté quand à l'origine des données. De plus, de nombreux autres chiffres sont avancés sans plus d'explication comme 90% selon l'organisation Opération de sauvegarde des sourds (OSS) : <http://www.oss2007.net/>

¹¹ Le texte intégral est disponible à l'adresse suivante :

<http://www.legifrance.com/affichTexte.do?dateTexte=&cidTexte=JORFTEXT000000809647&fastPos=2&fastReqId=848339230&oldAction=rechExpTexteJorf>

¹² Pour plus d'information voir chapitre 2 section 1.2.2

plus souple qu'un message filmé en LSF. De tels outils nécessitent de disposer d'un minimum de connaissances linguistiques sur le fonctionnement de la LSF afin de générer des messages corrects et compréhensibles par la population sourde.

La section suivante présente deux grands concepts linguistiques, dont la remise en cause est en lien direct avec le conflit précédemment exposé, afin de positionner notre travail par rapport à ces concepts.

1.2.Approche linguistique

Cette section vise à comprendre les conséquences du conflit évoqué précédemment sur la prise en compte du non manuel dans les LS et à situer notre travail par rapport aux différents débats linguistiques. Ainsi, la première sous-section présente le concept de paramètre, qui est discuté dans la seconde sous-section. La troisième sous-section présente la notion d'iconicité des LS. Enfin, la dernière sous-section propose une synthèse des conséquences du conflit en général tout d'abord puis pour le non manuel en particulier et nous nous positionnerons par rapport à l'iconicité et aux paramètres.

1.2.1. Le signe et les paramètres

Les LS se composent de signes¹³ qui se distinguent du "signe linguistique" de Saussure (Saussure, 1972). Chacun de ces signes est classiquement caractérisé par cinq paramètres (Moody, 1986). **La configuration** désigne la forme que la main doit adopter pour la production du signe : par exemple "main plate" ou "poing fermé". **L'emplacement** désigne l'endroit où se situe la main au début de la production du signe. **L'orientation** représente à la fois l'orientation de la paume de la main par rapport au signeur¹⁴ (vers le haut, le bas, l'intérieur - le corps du signeur - ou l'extérieur) et l'orientation de la main, autrement dit la direction montrée par les doigts. **Le mouvement** décrit le déplacement de la main, dont sa vitesse et sa direction. Enfin **les expressions du visage** regroupent les mouvements des sourcils, du front, des joues, des yeux et de la bouche du signeur durant la production du signe. Ces paramètres sont illustrés ci-dessous par la Figure 1 avec le signe [MUR].

¹³ Nous parlons ici de signe standard à savoir : « signe lexical pouvant exprimer un procès ou une entité » (Médin, Sallandre, 2006)

¹⁴ Un « signeur » caractérise toute personne communiquant en langue des signes.

On distingue la **main dominante** (la main droite pour les droitiers, la gauche pour les gauchers) de la **main dominée**. Lors de la production d'un signe, si une seule main est requise, on utilisera la main dominante (sauf si celle-ci est indisponible).

Si les deux mains sont nécessaires, deux cas se présentent : soit les deux mains vont se déplacer en symétrie, soit les deux mains vont se comporter différemment. La main dominée désignera alors souvent un référent fixe, un repère, tandis que la main dominante sera en mouvement par rapport à ce repère.

	<p>Configuration : Main plate variante angle 90°</p> <p>Emplacement : Torse.</p> <p>Orientation : Paume intérieure, doigts en direction de l'autre main</p> <p>Mouvement : Vertical de bas en haut.</p> <p>Expression du visage : Neutre.</p> <p>Les deux mains sont utilisées en symétrie.</p>
--	---

Figure 1 Description paramétrique du signe [MUR]

1.2.2. Le paramètre : une notion ouverte

Stokoe (1960), dans le but de démontrer la double articulation de l'ASL et donc le statut de langue des LS, est le premier à poser la notion de paramètre, qu'il nommera alors " aspect ". Selon lui, chaque paramètre est un élément minimal non porteur de sens, autrement dit un phonème, dont la combinaison permet de former un morphème (élément minimal porteur de sens). Aux termes de phonème et morphème il préférera les termes de chirème (du grec chir = main) ou phonème gestuel et kinème ou morphème gestuel.

C'est à Klima et Bellugi (1979) que l'on doit le terme " paramètre ", encore utilisé aujourd'hui, qui vient remplacer le terme " aspect ". Cette notion de paramètre n'a alors cessée d'être remise en question en ce qui concerne ses avantages, sa définition et son statut linguistique.

1.2.2.1. Les paramètres sont trop rigides

Filhol (2008) met en avant la rigidité des paramètres dans trois cas de figure : les valeurs exceptionnelles, les paramètres non pris en compte et la sur-spécification.

Pas de prise en compte des valeurs exceptionnelles

Filhol rappelle qu'à chaque paramètre correspond une liste finie de valeurs possibles. Or, cette liste ne tient pas compte des valeurs exceptionnelles, comme il le relève pour [LIT] (Image 1) par exemple, où la configuration, presque comme un 'H' mais avec l'annulaire et le majeur à la perpendiculaire, n'est pas recensée car est exceptionnelle. Certains signes ne peuvent donc être décrits par les paramètres, leurs valeurs n'appartenant pas à la liste des possibles. De plus, une liste figée n'est pas compatible avec la création de néologismes.



Image 1 [LIT]¹⁵

Les paramètres sont incomplets

L'auteur souligne que les paramètres actuels font référence aux mains ou au visage mais ne font pas référence au coude ou à l'avant bras. Or certains signes nécessitent une posture particulière d'un de ces éléments et ne peuvent donc être décrits comme pour [BINIOU, ECOSSE] (Image 2). Il en est de même des signes impliquant l'avant-bras tel que [BUREAU], [CAMPAGNE], [ESCARGOT]) (Image 3).

¹⁵ Sauf mention contraire, les images illustrant un signe proviennent toutes de Moody (1986, 1990).



Image 2 [ECOSSE]



Image 3 [ESCARGOT]

Les paramètres ne prennent pas en compte le contexte

L'auteur met également en garde contre la sur-spécification d'un signe qui empêche une signation naturelle en fonction du contexte notamment par exemple pour le signe [ARRETER] (Image 4) où la configuration 'angle droit' se rapprochera plus de la configuration 'B' si les mains sont plus près du corps. La rigidité des paramètres ne permet pas une adaptation au contexte mais ne permet qu'une description de signes isolés. De plus, comme l'explique l'auteur, la plupart des signes ont une réalisation variable au niveau de l'emplacement et de la taille tel que [ARBRE] (Image 5) dont " toute orientation accessible de la main dominée est valable pourvu que les doigts soient dans le prolongement du bras " (Filhol, 2008).

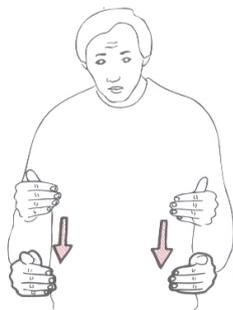


Image 4 [ARRETER]



Image 5 [ARBRE]

1.2.2.2. Remise en cause du statut linguistique et du contenu des paramètres

Les paramètres n'ont pas de statut linguistique unanime

Comme le souligne (Boutora, 2008), le statut linguistique des paramètres ne trouve pas de consensus au sein de la communauté scientifique. En effet, Stokoe leur a tout d'abord attribué le statut de phonème, dans le but de démontrer la double articulation. Or, un phonème doit être non porteur de sens, ce qui a très vite été remis en cause avec l'iconicité (Voir section 1.2.3).

Dès lors, lorsque la question n'a pas été éludée, les paramètres ont été assimilés aux phonèmes en considérant qu'un phonème peut être motivé et porteur de sens (Bouvet, 1992), soit comme des Unités Linguistiques Intermédiaires (ULI) qui sont non porteuses de sens mais potentiellement ouvertes à la signification lorsqu'elles sont combinées au sein d'un signe (Millet, 1997, 1998). Pour plus de précisions concernant le statut des paramètres et ses origines, se reporter à (Boutora, 2008).

Le nombre et le contenu des paramètres ne sont pas consensuels

Dans la partie 1.2.1, nous avons présenté cinq paramètres à savoir la configuration, l'orientation, l'emplacement, le mouvement et l'expression du visage. Or, certains de ces paramètres sont remis en cause comme par exemple le mouvement qui sera tantôt retenu comme un paramètre (Stokoe, 1960) tantôt rejeté car étant plutôt considéré comme la conséquence du passage d'un état à l'autre donc peu pertinent (Sandler, 1986). La prise en compte du non manuel est également relative et son rattachement à un paramètre reste non consensuel. Le rôle du non manuel a en effet d'abord été considéré comme trop minime pour devenir un paramètre notamment pour le modèle de Stokoe. Lorsqu'il a été intégré aux paramètres, le non manuel a parfois été associé aux mouvements de la main ou aux expressions du visage et aucun consensus n'a été trouvé (voir partie 2.2).

1.2.3. Notion de structure de grande iconicité

Alors qu'aux États-Unis les recherches sur l'ASL (American Sign Language) ont débuté dans les années 60, elles n'ont commencé en France que dans les années 80 grâce à Cuxac qui pose les bases des futures recherches françaises en proposant une vision d'ensemble du fonctionnement de la LSF. Ses premiers travaux, qui seront par la suite enrichis, approfondis et étayés, placent l'iconicité au cœur des LS.

Selon lui, un signeur a le choix entre deux modes d'expression : *dire en montrant* et *dire sans montrer* (Cuxac, 2000, pp. 27 – 29). Le premier cas utilise ce qu'il appelle des "structures de grande iconicité" c'est-à-dire qu'au lieu d'user du vocabulaire standard, l'accent va être mis sur la description visuelle de la situation. Le second cas (*dire sans montrer*) utilise le vocabulaire standard, sans intention de montrer.

En LSF, le signeur utilise les deux modes d'expression en un va et vient continu et peut même les combiner dans des proportions propres à chaque signeur. Les structures de grandes iconicités facilitent la compréhension de par leur visée illustrative, et permettent à l'interlocuteur de visualiser rapidement l'objet dont il est question. Cuxac appelle ces structures des "transferts". Il en dénombre trois types : les transferts de forme et de taille, les transferts situationnels et les transferts personnels.

Lors d'un transfert de forme/taille, la main du signeur prend la forme/taille de l'objet qu'il décrit. Par exemple, s'il parle d'une femme fine, il pourra utiliser le petit doigt et l'expression du visage (joues creuses par exemple) pour suggérer la forme globale et l'allure de la femme.

Lors d'un transfert situationnel le signeur va décrire les mouvements, l'action de la situation par rapport à un repère stable. Par exemple, dans le cas d'un chat qui monte sur une table, le signeur va d'abord spatialiser (à sa gauche par exemple) la table, puis sa main dominée va devenir un repère fixe représentant la table (souvent grâce à une proforme - forme que la main va prendre pour rappeler une caractéristique physique de l'objet auquel il fait référence – ici, la configuration de la main représentera la table), puis avec son autre main le signeur va décrire le mouvement du chat.

Enfin, lors d'un transfert personnel, le signeur prend le rôle d'un de ses protagonistes. Ainsi, son regard, son visage et son allure deviennent ceux du personnage transféré.



Image 6 Transfert personnel imagé (B. Moody & al. (IVT), 1983, p.79, 80).

Il n'est pas rare de se trouver en présence de plusieurs types de transferts en simultané (double transfert). Pour plus d'information sur ces structures se reporter à (Sallandre, 2003).

L'iconicité a été remise en cause pendant longtemps par la communauté scientifique internationale et souvent considérée comme ne faisant pas partie de la langue. Elle commence aujourd'hui à être prise en compte en dehors de l'école Cuxacienne notamment par Liddell (2003).

1.3. Conclusion

Bilan du conflit

Comme nous l'avons vu, l'opposition historique entre les oralistes et les défenseurs des LS a eu des conséquences aux niveaux social, juridique et scientifique. Les conséquences sociales concernent aussi bien la condition encore précaire des sourds aujourd'hui que les idées reçues du grand public toujours très présentes (Dalle-Nazebi et Lachance, 2008). Les deux sont en réalité intimement liées considérant par exemple la non qualification des sourds, qui est associée encore trop fréquemment à leurs capacités intellectuelles, d'où peu d'offres d'emploi intéressantes proposées aux sourds par employeurs, alors que cette non qualification découle pour une grande part du manque d'adaptation du système éducatif.

L'information est donc indispensable au même titre que la législation qui tente aujourd'hui de réparer une " injustice ancienne " (Jack Lang, 2002)¹⁶, notamment avec la loi de 2005 sur l'égalité des chances.

Enfin, ce conflit a façonné les prémisses de la recherche linguistique sur les LS dont le premier objectif a été de démontrer le statut de Langue Naturelle en essayant de calquer les connaissances des LV sur les LS. C'est ensuite dans la façon d'étudier la langue que s'est déplacé le conflit opposant ceux voulant comparer les LS aux LV et ceux ne voulant surtout pas une assimilation LS/LV. Ce débat est toujours ouvert et pose problème à chaque fois qu'un élément doit être nommé : va-t-on utiliser un mot existant pour les LV ou faut-il en inventer un nouveau ? Est-ce que les LS sont des langues à part ou faut-il les comparer aux autres langues ? (Boutora, 2008) souligne d'ailleurs que ces questions ont déjà été soulevées lors de l'étude des langues non indo-européennes à savoir : devaient-elles être étudiées avec les notions et outils venant des langues indo-européennes ?

Concernant notre étude

Notre étude porte sur le non manuel en LSF. Or, le non manuel a souvent été au cœur du conflit et a régulièrement été nommé pour démontrer que les LS n'étaient pas des langues mais des pantomimes. Ceci explique que les premiers linguistes, tel que Stokoe, l'exclurent de leurs recherches, déclarant que le non manuel ne faisait pas partie des LS, afin d'en démontrer le statut de langue en évitant les polémiques. Ainsi, peu d'études ont porté sur le non manuel même s'il a souvent été nommé, comme exposé dans la partie suivante, ce qui explique le peu de connaissances dont on dispose à ce sujet.

Notre étude vise à connaître le fonctionnement du non manuel, ses structures et ses fonctions à des fins de génération automatique. Si nous ne prétendons pas participer au débat linguistique, nous considérons l'iconicité comme acquise en LSF et utiliserons les observations de (Cuxac, 2000) comme point de départ de notre analyse.

¹⁶ Discours de Jack Lang, alors ministre de l'Education Nationale, du 13 février 2002. Disponible à l'adresse suivante : ftp://trf.education.gouv.fr/pub/edutel/actu/2002/02_13_discours_lang_languesignes.pdf

Enfin, concernant les paramètres, le seul consensus possible est de considérer qu'un paramètre est un constituant d'un signe sans chercher à lui donner un statut linguistique et c'est ainsi que nous l'employons. On peut alors se poser la question de savoir si le non manuel est un paramètre ou pas ? Est-ce qu'il est un constituant d'un signe ? Ces questions sont soulevées dans la partie suivante.

2. Les Gestes Non Manuels (GNM)

Par non manuel, nous entendons le mouvement de la tête, des sourcils, des paupières, des joues, du nez, de la bouche, des épaules et du torse ainsi que la direction du regard. Cette partie présente le non manuel à partir des connaissances actuelles, en particulier leurs fonctions linguistiques, pour proposer une réflexion sur le nom à leur donner en rapport avec leurs rôles à tous les niveaux de la langue.

2.1. Fonctions

Le non manuel et en particulier l'expression du visage a longtemps été mise de côté par les linguistes en raison du parallèle avec les pantomimes, comme nous l'avons souligné ci-dessus, puis du parallèle avec les émotions. Corina, Bellugi et Reilly ont démontré que chez les personnes sourdes, les expressions d'émotions étaient contrôlées par l'hémisphère droit du cerveau, tandis que les expressions linguistiques étaient contrôlées par l'hémisphère gauche. Dès lors, ces expressions linguistiques entrent dans la fonction langagière et ne sont donc pas assimilables à l'émotion (Corina, Bellugi et Reilly, 1999). Par la suite, les éléments non manuels ont été évoqués à de nombreuses reprises pour leurs rôles linguistiques (Baker-Shenk, 1985) mais sans pour autant avoir fait l'objet d'études spécifiques. Ainsi, cette section présente les connaissances relatives au non manuel, regroupées en fonction du niveau linguistique sur lequel il intervient. Afin de ne pas entrer dans le débat exposé précédemment, nous distinguerons uniquement deux niveaux : le lexical et l'énoncé.

2.1.1. Niveau lexical

Le non manuel intervient au niveau lexical dans quatre cas de figure :

- Dans la constitution d'un signe standard (5^{ème} paramètre) ;
- Dans la distinction de deux signes ;
- Dans la distinction entre l'objet et son action et/ou utilisation ;
- Dans la réalisation de modificateurs (adjectifs qualificatifs, adverbes, etc.).

2.1.1.1. Signe standard

Comme abordé dans la section 1.2.3, le 5^{ème} paramètre représente l'expression du visage. Or cette dernière fait partie des paramètres justement parce que si le signe est produit seul, sans expression du visage, il est considéré comme tronqué (Jouison, 1995). Tel est le cas par exemple de [LAQUE] (Image 7).



Image 7 [LAQUE]

2.1.1.2. Distinction de deux signes

La LSF, comme toutes les langues, possède de nombreux homonymes appelés homosignes tels que [VERRE] et [CONSERVE] (Image 8) ou [CHOCOLAT] et [VIDE] (Image 9).

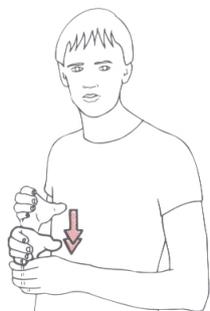


Image 8 [VERRE]



[CONSERVE]



Image 9 [CHOCOLAT]



[VIDE]

D'autres signes en revanche sont proches des homosignes, leurs paramètres manuels étant identiques, et se distinguent grâce au non manuel comme nous le montrent les paires minimales ci-dessous (Moody, 1986, 1990).

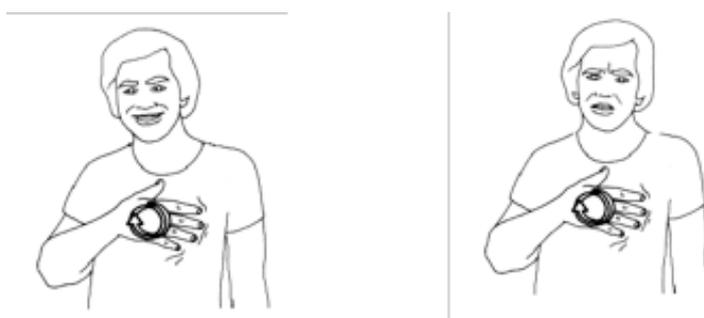


Image 10 [CONTENT] – [MAL AU CŒUR]

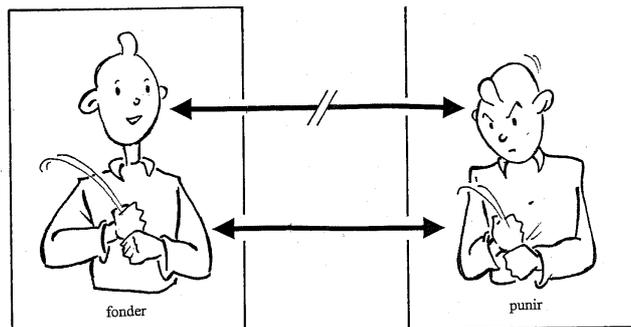


Image 11 [FONDER] – [PUNIR]. Images extraites de (Verlaine)

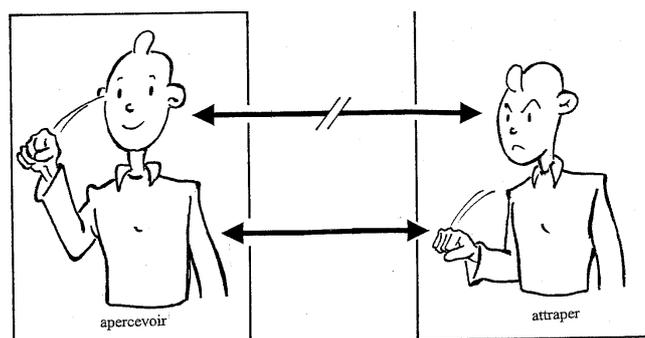


Image 12 [APERCEVOIR] – [ATTRAPER]. Images extraites de (Verlaine)

2.1.1.3. Distinction entre l'Objet et son Action/Utilisation

Enfin, il est fréquent que la distinction entre l'objet comme [PIPE] et son action: [FUMER LA PIPE] (Image 13) se fasse grâce au non manuel (ici la bouche qui fait " P.P.P ", (Moody T2)). L'Image 14 nous montre le mouvement manuel pour [NUIT] et [LA NUIT TOMBE] qui se distinguent notamment par un plissement des yeux en même temps que l'abaissement des bras (Vergé, 2001).

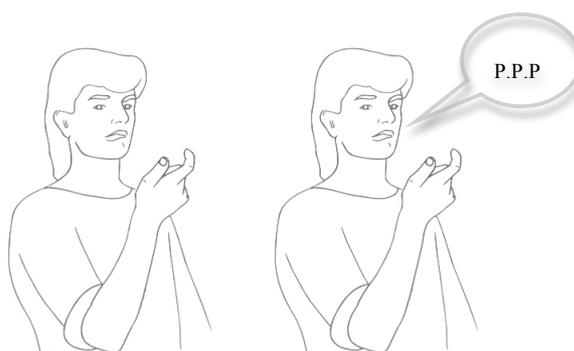


Image 13 [PIPE] – [FUMER LA PIPE]



Image 14 [NUIT] – [LA NUIT TOMBE]

2.1.1.4. Les modifieurs

Le non manuel permet la réalisation des modifieurs et ce aussi bien en LSF (voir ci-dessous) que dans la langue des signes américaine (Baker, 1979 ; Neidle et al, 2000) et dans bien d'autres encore.

Selon (Cuxac, 2000), les qualificatifs sont réalisés grâce à la mimique faciale et peuvent être objectifs ou subjectifs selon l'orientation du regard. Par exemple, dans le cas d'un discours rapporté, si le locuteur exprime " gros, volume important" il peut s'agir soit de l'avis du personnage rapporté, soit de l'avis du locuteur. Il décrit ainsi douze expressions relevées dans son corpus vidéo et leur interprétation : " grande quantité, grand nombre de " ; " gros, volume important " ; " flasque, gélatineux " ; " spongieux " ; " vite " ; " minuscule " ; " petit " ; " normal " ; " mince " ; " fort, costaud (grand) " ; " pointu, piquant ". Certaines de ces expressions peuvent se combiner entre elles.

De plus, associés à certains signes, les qualificatifs prennent des valeurs de quantifieurs (" beaucoup " par exemple). Le degré d'ouverture des yeux peut également remplir cette fonction adverbiale (Vergé, 2001) en ce qui concerne les valeurs spatiales et les valeurs temporelles (plus les yeux sont plissés, plus l'objet/l'époque concerné(e) est loin(taine)).

2.1.2. Au niveau de l'énoncé

Le non manuel intervient à de nombreuses reprises sur l'énoncé, apportant par exemple à lui seul la modalité. Il va ici jouer un rôle comparable à la ponctuation de l'écrit ou à la tonalité de l'oral. Ainsi, il est indispensable à la bonne compréhension du message final.

2.1.2.1. La modalité

D'après les travaux de Cuxac

La modalité s'exprime grâce à une expression du visage souvent associée à un regard spécifique ainsi qu'à des mouvements de tête et/ou de torse (Cuxac, 2001). Cuxac propose une description de la réalisation de chaque mode :

“ Il (i.e. le conditionnel) se réalise formellement par une inclinaison et un léger mouvement de recul du visage vers l'arrière; le regard, désinvesti, est dirigé vers le haut, les sourcils sont relevés ”

(Cuxac, 2000, p 226).

Les mimiques faciales à valeur modale peuvent se diviser en deux catégories : celles exclusivement à valeur modale dans le contexte standard et celles qui peuvent parfois faire partie des catégories précédemment évoquées (modificateurs) dans un contexte de grande iconicité.

Cuxac distingue cinq mimiques modales à savoir la “ normalité ” ; le “ conditionnel ” ; “ l'hypothèse mentale ” ; “ duplicité ” et “ aboulie ”.

“ Duplicité ” indique une intention de “ duperie, tromperie ” sans que les autres personnes n'en soient informées. Dans le cas où cette dernière condition n'est pas respectée (moquerie ouverte) le signeur n'aura pas recours à la duplicité. De plus, cette mimique est toujours accompagnée d'un verbe précis (faire semblant, mentir, faire exprès de, tromper, etc.).

“ Aboulie ” fait référence à ce qui est hors du contrôle du protagoniste soit par manque de volonté ou de possibilité d'action, soit par ignorance. De plus sa réalisation varie d'un locuteur à l'autre.

Concernant les mimiques non exclusivement modales, qui peuvent être tantôt modales, tantôt des “ compléments de manière ”, Cuxac en dénombre treize : “ l'impératif ” ; le “ volitif ” ; “ l'incitatif ” ; le “ réprobatif ” ; “ l'ironique ” ; le “ dubitatif ” ; “ l'assertif ” ; le “ capacitif ” ; la “ problématisation ” ; le “ concessif ” ; “ l'interrogatif ” ; le “ négatif ” ; “ l'interro-négatif ” (Cuxac, 2000, pp. 230 – 232).

Rôle particulier des sourcils

Selon Vergé, les sourcils sont particulièrement importants pour réaliser la modalité et la posture adoptée par les sourcils pour un mode précis sera maintenue durant toute la production de l'énoncé (Vergé 2001).

Dans les autres LS

Kolod souligne que la modalité en ASL se détecte grâce à une combinaison de mouvements non manuels :

“ Movement of the signer’s eyes, face, and head (...). When these movements are combined they indicate whether a sentence is a question, an assertion, a command, or a conditional. ”.

(Kolod, 2004, p 3)

Interrogative, conditionnel et topicalisation dans les autres LS

Les sourcils relevés sont présents à chaque interrogative fermée et à chaque conditionnelle. Cela est le cas aussi bien pour la Langue des Signes Américaine (ASL) (Liddell, 1980 ; Baker-Shenk, 1985) que pour la Langue des Signes Suédoise (SSL), (Bergman, 1984), Danoise (DSL), (Vogt-Svensen, 1990) et Néerlandaise (NSL) (Coerts, 1992). Il semble donc que l'élément “sourcils levés” soit un élément présent dans plusieurs marqueurs complexes (composés de plusieurs éléments non manuels) de la modalité. Leur présence ne suffit pas à elle seule à indiquer le mode mais est un élément indispensable du marqueur complexe de chaque mode. Coulter note par exemple qu'aucune phrase contenant des sourcils relevés n'est assertive et que le syntagme qui porte le levé de sourcil fonctionne toujours comme un topique (Coulter, 1978).

En Langue des Signes Québécoise (LSQ), Dubuisson montre que ce sont les mouvements de la tête qui permettent de marquer le type d'interrogation : une question ouverte sera marquée par une inclinaison de la tête vers l'arrière tandis qu'une question fermée sera marquée par une inclinaison avant (Dubuisson, 1996 ; Dubuisson & al, 1999). Enfin, Dubuisson, Lelievre, Parisot et Rancourt mettent en avant le rôle syntaxique des sourcils en particulier au niveau des topicalisations et des conditionnelles où le passage des sourcils relevés aux sourcils froncés peut rendre l'énoncé agrammatical (Dubuisson, Lelievre, Parisot et Rancourt, 1999).

2.1.2.2. Délimitation de transfert

Le transfert personnel est délimité par le regard : avant le transfert le regard est posé sur l'interlocuteur, au début du transfert le regard quitte rapidement celui de l'interlocuteur et ce durant toute la durée du transfert puis en fin de transfert le regard se pose à nouveau sur l'interlocuteur (Cuxac, 2000) comme le montre la Figure 2.

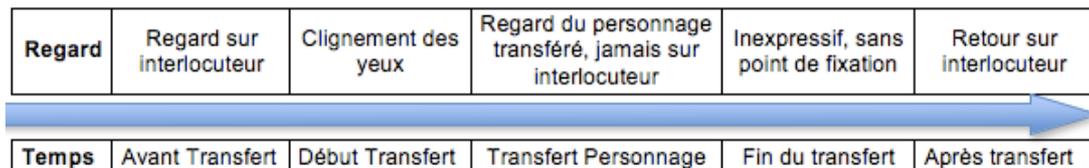


Figure 2 rôle du regard lors d'un Transfert personnel

De plus, le regard, associé à la mimique faciale, aux balancements corporels, voire à l'orientation du corps, de la tête et des épaules, permet de distinguer les différents types de transferts (Sallandre, 2005).

2.1.2.3. Segmentation d'énoncés

Le non manuel joue également un rôle important dans les segmentations d'énoncés et en particulier pour délimiter les groupes syntaxiques. Ainsi, Cuxac observe une relation étroite entre les balancements/rotations du corps et les frontières syntaxiques, tandis que Vergé souligne que les clignements des yeux délimitent des " groupes de sens " comme le font les pauses à l'oral ou la ponctuation à l'écrit.

2.1.3. Synthèse

En conclusion, le non manuel intervient à différents niveaux de langue. Nous proposons ci-dessous un récapitulatif en fonction de chaque élément non manuel et du niveau linguistique dans lequel il intervient.

	Transfert de Taille/Forme	Transfert Situationnel	Transfert Personnel	Général
<i>Regard</i>	- Installe la forme; - Accompagne le déploiement - Annonce fin du transfert	- Installe locatif; - Installe l'actant; - Anticipe point d'arrivée	Regard du personnage. transféré, ne croise jamais celui de l'interlocuteur	- Régit l'interaction; - Vers le haut pour hypothèse mentale; - Construction d'une ref spatiale;
<i>Mimique Faciale</i>	Rôle qualifiant et quantifiant	Nature aspectuelle de l'action	Etat d'esprit du personnage. transféré.	- Valeur aspectuelle (parfois). - Valeur modale (souvent). - Valeur déterminants grammaticaux du nom et fausse question
<i>Clignement des yeux</i>			Début et fin de transfert	- Démarcation de frontières syntagmatiques; - Début et fin d'énoncé; - Annonce une hypothèse mentale.
<i>Corps</i>			- Change en début de TP; - Relâche en fin de TP; - Retour à position non marquée	Balancement = marqueur de frontière syntaxique; si appuyée = valeur assertive marquée

Tableau 1 Intervention du non manuel selon les niveaux linguistiques

Enfin, le tableau 2 ci-dessous propose une synthèse du rôle du non manuel dans les transferts selon Cuxac (2000).

	Niveau lexical	Niveau énoncé		
<i>Regard</i>		Réalisation de la modalité	Accord verbal	Délimitation d'un transfert personnel, distinction du type de transfert
<i>Mimique Faciale</i>	Sourcils, bouche, yeux: - Réalisation d'un signe, - Distinction de paires minimales - Expression de la modalité	Sourcils, bouche, yeux, joues: réalisation de modifieurs	Sourcils: - Topicalisation, - Conditionnel	Distinction du type de transfert
<i>Clignement des yeux</i>			Délimitation des groupes syntaxiques	Passage d'un personnage à l'autre dans les transferts personnels
<i>Corps</i>	Inclinaison du buste: - Réalisation d'un signe - Distinction de paires minimales	Mouvement de la tête et du buste: réalisation de la modalité	Délimitation de groupe syntactique	Distinction du type de transfert

Tableau 2 Récapitulatif des rôles du non manuel en fonction de l'élément qui intervient.

2.2. Paramètres ou Gestes non Manuels ?

Si l'on considère les paramètres, hors de tout débat linguistique, comme les éléments qui constituent un signe, on peut se demander si le non manuel est un paramètre.

Comme nous l'avons vu ci-dessus le non manuel intervient bien dans la constitution d'un signe mais il intervient également aux niveaux morphologique et syntaxique. Ainsi, il n'est pas uniquement une composante d'un signe et la dénomination " paramètre " est en cela réducteur. Il en est de même pour la notion de trait qui fait référence aux éléments qui constituent un paramètre ou à un paramètre (Boutora 2008).

Pour procéder au choix d'un terme, une observation de sa définition marque un préalable nécessaire. Il est important que cette dernière réponde aux différents composants que l'on souhaite appréhender : le regard, le mouvement des sourcils, des yeux, de la bouche, des joues, des épaules, du torse, de la tête et des bras. Selon le dictionnaire de l'Académie Française neuvième édition, le geste se définit de la façon suivante : " attitude, mouvement du corps, mimique : mouvement du corps, notamment du bras, de la main, de la tête ".

Selon le TLF1¹⁷ le terme " geste " vient du latin " gestus " (attitude, mouvement du corps, geste) et désigne " une activité corporelle particulière d'une personne, un mouvement extérieur du corps ou de l'une de ces parties ".

Le terme " geste " évoque donc bien l'idée de mouvement de parties du corps. En ajoutant " non manuels " il est donc possible d'exclure les mains et d'obtenir ainsi uniquement les mouvements des parties du corps évoquées précédemment. De plus, les LS sont des langues visuo-gestuelles, le terme " geste " semble donc approprié. Nous considérons que ces langues sont composées de Gestes Manuels (GM) et de Gestes Non Manuels (GNM).

3. Conclusion

L'objectif de ce chapitre était de présenter globalement notre objet d'étude à savoir la LSF et en particulier les GNM. Nous avons tout d'abord proposé une approche historique afin de

¹⁷ Trésor de la Langue Française Informatisé, disponible sur : <http://atilf.atilf.fr/tlf.htm>

comprendre l'origine du conflit LS – LV et d'en percevoir les conséquences au fil de l'histoire.

Nous avons pu constater que, si les progrès scientifiques ont permis d'invalider les croyances d'alors (tel que le lien entre l'ouïe et l'intelligence par exemple), celles-ci sont pourtant restées d'actualité pour la population et ont eu des conséquences désastreuses aussi bien au niveau de la législation que de la condition des sourds. Actuellement, l'empreinte de l'histoire est toujours présente et à bien des niveaux. Nous avons par exemple observé son influence sur la recherche linguistique, où la préoccupation première a été de prouver que les LS étaient bien des langues ce qui a engendré deux courants de recherches opposés et peut expliquer par exemple le peu de connaissances dont on dispose actuellement sur les GNM.

L'observation de la situation actuelle, en particulier au niveau législatif, nous a conduit à introduire la loi de 2005, réel tournant dans l'histoire puisqu'elle permet enfin d'envisager une amélioration de la condition des sourds. En effet, cette loi permet d'une part la reconnaissance officielle de la LSF comme langue à part entière mais devrait également permettre, grâce à l'obligation d'accessibilité, l'explosion des recherches linguistiques et informatiques portant sur les LS. Dans les deux disciplines, les chercheurs tentent de disposer de connaissances suffisantes sur le fonctionnement de la langue afin de la modéliser.

Notre étude portant tout particulièrement sur les GNM, nous avons alors dressé un état des lieux des connaissances relatives à ces derniers. Ceci nous a permis de constater que leur fonction portait aussi bien sur le lexique que sur l'énoncé et nous a conduit à mener une réflexion autour du nom à leur donner. C'est ainsi que nous avons adopté le terme Geste Non Manuel (GNM).

L'objectif de ce travail est de générer automatiquement des GNM. Or, la génération nécessite de disposer d'informations précises sur le fonctionnement de ce que l'on souhaite générer. Le chapitre suivant propose de replacer le contexte de l'étude en présentant tout d'abord la discipline dans laquelle se situe cette étude, puis l'application de cette discipline à la LSF en général et aux GNM en particulier. Nous verrons ainsi quelles informations sont nécessaires à la génération des GNM et pourront ainsi définir les informations que notre étude devra permettre d'obtenir

Chapitre 2 : Contexte et problématique de la recherche

Ce chapitre a pour objectif de replacer notre étude dans son contexte - à savoir le TAL (Traitement Automatique des Langues) - et d'en présenter la problématique.

Une présentation de cette discipline et de ses applications nous permettra de nous intéresser plus précisément à la génération automatique de LSF et des GNM en particulier. Comme nous le montrons, les connaissances structurelles dont nous disposons concernant les GNM ne permettent pas leur génération automatique ce qui nous conduit à envisager une nouvelle étude dont la première étape est l'annotation d'un corpus.

Nous nous intéressons ensuite aux outils disponibles pour mener à bien cette annotation. En premier lieu, il s'agit de définir la façon dont nous allons décrire ce que nous observons sur le corpus vidéo. Nous présentons donc les systèmes de notations existants et proposons une évaluation critique de chacun de ces systèmes au regard de nos besoins. En second lieu, il s'agit de choisir le support d'annotation le plus adéquat à cette étude. Nous présentons chacun de ces supports avec leurs avantages et inconvénients afin de déterminer lequel répond le mieux à nos attentes.

1. Contexte de l'étude : Le Traitement Automatique des Langues

Cette partie vise à présenter le contexte de notre étude à savoir le Traitement Automatique des Langues (TAL) en général et le Traitement Automatique des Langues Signées (TALS) en particulier.

La première section propose une définition du TAL. La seconde section aborde le TALS et plus précisément la génération automatique de LSF. Nous nous intéressons plus particulièrement à la génération automatique des GNM qui, comme nous le montrons dans une troisième section, pose encore de nombreux problèmes en raison du peu de connaissances structurelles dont nous disposons à leur sujet. Ainsi, il est nécessaire de disposer de données plus précises ce qui implique de mener une réflexion autour de la méthodologie qu'il conviendrait de mettre en œuvre pour répondre à nos objectifs. Nous concluons donc qu'il paraît pertinent de procéder à une nouvelle annotation centrée sur les GNM.

1.1. Qu'est-ce que le TAL ?

Le TAL est une discipline à mi-chemin entre la linguistique et l'informatique visant à créer des programmes informatiques qui, à partir d'une entrée (textuelle, vidéo, sonore, etc.) propose un traitement qui offrira une sortie différente. Le traitement peut agir sur des textes ou en générer :

“ Il peut agir sur des données linguistiques (c'est à dire les textes) pour les corriger, les condenser ou les traduire. Bien souvent, cette transformation comprend une étape intermédiaire qui vise à extraire des textes leur représentation : elle est appelée analyse des langues naturelles. Le terme représentation (...) désigne toute traduction du texte dans un système autre que la langue naturelle et qui rend explicite des informations implicites dans le texte (...). Dans ce premier type de traitement, l'entrée est donc un texte et la sortie un nouveau texte ou une représentation de texte.

Le TAL peut aussi faire l'opération inverse : il prend alors en entrée la représentation de texte, pour produire un texte en langue naturelle (...). Cette opération est appelée génération des langues naturelles. ”

(Bouillon, Vandooren, 1998, pp. 5, 6)

Par exemple, la reconnaissance automatique de la parole part d'une source orale, l'analyse (trouve les équivalences phonèmes/morphèmes, teste les segmentations, etc.), puis en propose une version écrite. La plupart de ces outils utilisent les statistiques afin d'augmenter leurs taux de réussite notamment pour la traduction automatique (Léon, 2008).

La section suivante aborde l'application du TAL à la LSF, à savoir le Traitement Automatique des Langues Signées (TALS) et présente plus particulièrement la génération automatique de LSF.

1.2. Application à la LSF

Les techniques du TAL peuvent être utilisées pour la création d'outils permettant de faciliter la communication entre les sourds et les Entendants en proposant un pont entre les LS et les LV. Dans l'idéal, ce pont serait un traducteur automatique permettant de traduire instantanément de la LSF au français (ou autre LV) et réciproquement. Cependant, les études linguistiques étant encore récentes (comme évoqué ci-dessus) et les connaissances sur l'organisation et le fonctionnement des LS encore limitées, la traduction automatique est plutôt un objectif à long terme¹⁸.

Les recherches actuelles visent à apporter une aide concrète au quotidien des sourds, par exemple en leur facilitant l'accès aux informations de la vie quotidienne. Pour se faire, il est nécessaire d'accroître les connaissances linguistiques disponibles sur les LS en ayant recours, si possible, à des outils d'analyse ou de traitement automatique.

Nous distinguons trois champs de recherches informatiques : l'analyse automatique de corpus vidéo (Lefebvre-Albaret, Dalle, 2009), la reconnaissance automatique de mouvement (Cooper, Bowden, 2009) et la génération automatique de LSF (explicité ci-dessous). Notre étude visant à approfondir les connaissances dont nous disposons sur les GNM à des fins de génération automatique, seul ce dernier champ de recherche sera abordé plus en détail.

¹⁸ Rappelons que la traduction automatique entre deux LV pose encore de nombreux problèmes tel que les ambiguïtés lexicales (Léon, 2009).

La première sous-section revient sur la notion de génération automatique de LSF et illustre les travaux en la matière à travers un exemple de projet. La seconde sous-section présente plus particulièrement les travaux de génération automatique de l'équipe M&TALS du LIMSI, contexte dans lequel se situe nos travaux.

1.2.1. La génération automatique de LS

La génération automatique de LS implique la création d'une forme linguistique de la LS à partir d'une représentation formelle de son contenu. Cette forme est visualisée par l'animation d'un signeur virtuel¹⁹ (ou avatar signant). Il existe quatre méthodes de création d'animation 3D (ou d'animation d'un signeur virtuel):

- La capture de mouvements : chacun des signes est pré-enregistré par une personne équipée de capteurs. Les données de ces capteurs vont être transmises à la représentation informatique du squelette du signeur virtuel, lui permettant ainsi de rejouer exactement ce qui vient d'être signé.
- La rotoscopie : la base de données est réalisée par un infographiste par recopie de séquences vidéos. Les signes sont ensuite concaténés avec une gestion de la coarticulation.
- La génération automatique d'animation : chaque signe est modélisé, décrit, afin d'être interprété puis joué automatiquement par le module d'animation
- Une nouvelle méthode utilise une base de données de mouvements capturés plus ou moins modifiés. Ces mouvements sont concaténés avec une gestion de la coarticulation.

Nous illustrons ci-dessous deux types de génération automatique d'animation 3D à travers le projet européen VISICAST²⁰ (2000 – 2003) et son prolongement le projet européen eSign (2002 – 2004). Dans une première version, les animations étaient basées sur la capture de mouvements, puis dans une deuxième version, elles étaient générées syntaxiquement.

¹⁹ Par signeur virtuel, également appelé avatar signant, nous entendons un personnage virtuel en 3D qui s'exprime en langue des signes.

²⁰ Virtual Signing: Capture, Animation, Storage and Transmission. Informations disponibles sur : <http://www.visicast.co.uk>

Le projet VISICAST est un projet du programme européen IST²¹ (Bangham et al., 2000) qui a débuté en janvier 2000. Ce projet, partenariat entre le Royaume-Uni, l'Allemagne, la France et les Pays-Bas, avait pour objectif de rendre les informations et services plus accessibles aux sourds, en proposant une traduction signée via un signeur virtuel. Deux applications sont nées de ce projet : une application Web proposant une version signée d'un site web (Verlinden et al., 2001) et une application " face à face ", TESSA²² (Cox et al., 2002) (Image 15), dont l'objectif était de faciliter la communication entre les sourds et l'administration. Par exemple, TESSA a été proposée lors d'une démonstration pour être utilisée dans les postes anglaises. Ces applications utilisent des gestes isolés contenus dans une base de données constituée grâce à la capture de mouvements.



Image 15 TESSA

Le projet s'est ensuite orienté vers une animation générée automatiquement (Kennaway, 2001) à partir d'un langage de haut niveau créé durant le projet nommé SIGML (Elliot et al., 2000). SIGML est un dérivé au format XML du système d'HamNoSys qui permet de décrire un signe selon ses paramètres (voir la seconde partie de ce chapitre).

Le projet s'est prolongé à travers le projet eSign²³ (Kennaway, 2003) de septembre 2002 à septembre 2004, dont l'objectif était de faciliter l'accès aux informations du gouvernement aux sourds grâce à un signeur virtuel.

De ce projet est né le signeur virtuel GUIDO qui propose actuellement une traduction signée de certains sites notamment en Allemagne, en Hollande et en Angleterre. L'Image 16 est une copie d'écran d'une page web anglaise (deaf connexions) qui propose une version signée d'une partie du site grâce à GUIDO.

²¹ Information Society Technologies

²² Text and Sign Support Assistant, informations disponibles sur :
<http://www.visicast.co.uk/news/Tessa.html>

²³ Essential Sign Language Information on Government Networks. Informations disponibles sur :
<http://www.visicast.sys.uea.ac.uk/eSIGN/index.html>

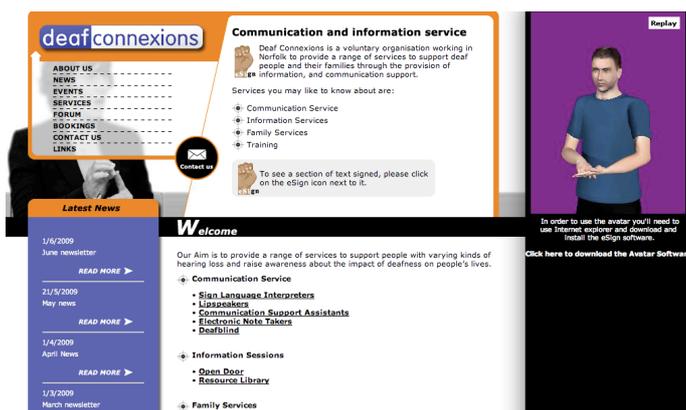


Image 16 GUIDO dans deaf connexion²⁴

1.2.2. Les travaux du LIMSI

L'équipe M&TALS (Modélisation et Traitement Automatique de la Langue des Signes) créée en mai 2004 et nouvellement rattaché au groupe ILES (Information, Langue Écrite et Signée) du LIMSI travaille sur la modélisation de la LSF à partir de corpus vidéos à des fins de génération automatique. Les travaux en cours portent sur différents niveaux de traitement :

- **L'annotation de corpus et modélisation :** Observation d'un corpus, annotation (pouvant passer par la mise en place de la méthodologie d'annotation), afin d'obtenir une modélisation des phénomènes concernés. Dans l'équipe il s'agit actuellement de la modélisation de la coarticulation (Segouat, 2009) et des GNM (Chételat et al., 2008).
- **La modélisation informatique :** Modélisation des signes en contexte (Filhol, 2008), de l'espace de signation et des énoncés spacialisés (Braffort et Dalle, 2008).
- **L'implémentation de prototype pour l'évaluation :** Création d'un module d'animation du signeur virtuel nommé Octopus qui permet actuellement de rejouer et de concaténer des animations préenregistrées réalisées par l'infographiste. De plus, il s'agit de créer un module de génération automatique d'animation 3D

²⁴ Copie d'écran de la page internet : <http://www.deafconnexions.org.uk/dev/>

(Delorme, 2009). A terme, l'objectif est de proposer une plateforme de génération automatique de LS, nommée GénéALS (Bolot & al., 2006).

- **La valorisation et le transfert :** Par l'intermédiaire de Websourd²⁵, le module d'animation Octopus est utilisé dans le cadre d'un projet de signeur virtuel pour la SNCF. Il permet, intégré à un logiciel de Websourd, de produire en LSF les messages audios diffusés dans les gares.

Ces travaux sont illustrés par la Figure 3.

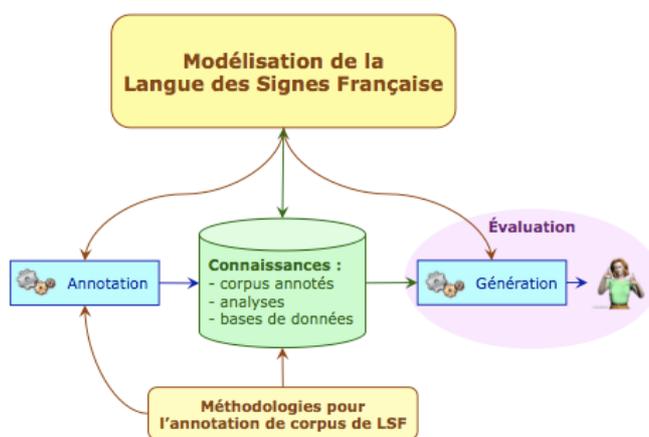


Figure 3 Organisation des travaux de M&TALS²⁶

Actuellement, l'équipe participe au projet européen Dicta-Sign²⁷ qui vise à rendre la communication en ligne, l'interaction sur internet, plus accessible aux sourds. L'objectif est de créer des sites en LS permettant aux sourds de participer, d'interagir en ajoutant, commentant, modifiant et partageant des informations (par exemple les forums mais également les blogs et les encyclopédies ouvertes) afin de faciliter leurs contributions.

Actuellement, lorsque cela est possible, la contribution des sourds se fait par vidéo grâce à une webcam. Or, ce type de procédé ne garantit pas l'anonymat de la personne. N'importe qui est donc susceptible de reconnaître l'auteur du message, ce qui n'encourage pas la participation des sourds. De plus, modifier ou ajouter une précision à une vidéo postée

²⁵ www.websourd.org

²⁶ Schéma emprunté à (Braffort, 2008)

²⁷ www.dictasign.eu

nécessite de refaire une nouvelle vidéo, contrairement à un post écrit qu'il suffit d'éditer pour le corriger.

L'objectif du projet est donc de permettre de publier sa contribution en utilisant un signeur virtuel. Pour ce faire, l'utilisateur devra signer son message devant une WebCam. Ce message sera alors convertit en une représentation de la LS puis joué par un signeur virtuel. Ainsi, l'anonymat sera préservé et les messages seront modifiables aisément. Ce projet vise également à proposer un dictionnaire à entrées gestuelles par webcam et un système de traduction de LS vers une autre LS (Efthimiou et al., 2009).

1.3. La génération des GNM

Comme l'illustrent les projets ci-dessus, le TALS pourrait être d'une aide précieuse pour diminuer les inégalités qui séparent encore les sourds des Entendants, en permettant un accès aux informations courantes grâce à la génération automatique.

Celle-ci implique de disposer d'un modèle des signes et de la langue. Les recherches actuelles visent à tenir compte des spécificités des langues gestuelles à savoir la gestion de l'espace et de l'iconicité. Cependant, " la prise en compte du regard, de la mimique faciale et des mouvements du corps reste encore à développer " (Pelachaud et al., 2004). A l'heure actuelle, les GNM ne sont produits par les signeurs virtuels que de façon non automatique, à savoir une succession d'interpolations entre des images clés définies manuellement et de façon empirique.

Générer des GNM nécessite de connaître suffisamment leur fonctionnement ainsi que la structure interne et la dynamique de leurs mouvements, afin de les modéliser. Par exemple, pour une hausse de sourcil, il est nécessaire de savoir quand elle a lieu, quelle partie du sourcil se lève, de combien et sur quelle durée, combien de temps cette hausse sera tenue, etc. Or, si de nombreuses études (exposées dans le chapitre 1) ont permis de démontrer l'importance des GNM et quelques unes de leurs fonctions, aucune n'a permis d'apporter les connaissances suffisantes pour leur génération automatique.

De telles connaissances proviennent de l'observation et de l'annotation minutieuse de ces phénomènes sur un corpus vidéo. Comme aucune étude n'a porté sur les mécanismes mis en œuvre dans la production de GNM, aucune annotation ne les décrit suffisamment. Il semble donc pertinent d'envisager une nouvelle étude centrée sur les GNM dont la première étape est l'annotation de ces phénomènes. La question se pose alors de savoir comment décrire ce que l'on voit ? Comment annoter le corpus vidéo afin d'obtenir les informations souhaitées pour une modélisation des GNM ? Nous proposons ci-dessous un inventaire des différents systèmes de description existants afin de déterminer lequel est le plus adéquat pour conduire notre étude.

2. Annotation des GNM

Pour mener à bien une étude linguistique, “garder une trace” (Sallandre, 2003) de ce que l'on observe peut s'avérer utile, principe même d'une annotation. Pour cela, il faut disposer d'un outil permettant de décrire les éléments souhaités. Dans le cas d'une langue gestuelle, l'annotation peut se révéler d'autant plus problématique qu'il n'existe pas de version écrite de cette langue ni un standard de transcription de la gestualité (tel que l'est l'API²⁸ pour les LV).

En réponse, certains chercheurs (exposés ci-dessous) ont utilisé des mots de la LV tandis que d'autres ont inventé des systèmes graphiques. Dans la littérature, ces systèmes sont souvent regroupés en fonction de leur finalité (les systèmes de transcription (Jemni ; Elghoul ; Makhlouf, 2007) et les systèmes d'annotation (Aznar, Dalle, 2005)), de leur niveau de traitement mais également en fonction de leur linéarité (les systèmes mono linéaires, les partitions manuscrites ou multimédias (Sallandre, 2003)).

Nous nous situons du point de vue de l'annotateur et présentons ces systèmes sous l'angle du processus d'annotation et non de leur fonction linguistique. Nous considérons donc la distinction suivante : il existe tout d'abord des formes graphiques qui permettent, entre autre, “d'écrire” ou de décrire un signe ou une composante de celui-ci. D'autre part, il existe des supports d'annotation dans lesquels il est possible d'utiliser les mots de la LV ou les formes graphiques évoquées précédemment.

²⁸ API : Alphabet Phonétique International

La première section présente les différents systèmes existants qui permettent de décrire le non manuel (GNM ou émotions) et un positionnement par rapport à chacun d'eux. Une seconde section propose un tour d'horizon des supports d'aide à l'annotation mis au point pour répondre à la nécessité de synchroniser les supports avec la vidéo du corpus.

2.1. Comment décrire les GNM ?

Cette partie présente les différents systèmes graphiques, permettant de décrire le non manuel, créés spécifiquement pour les LS (première sous-section) ou pour l'expression de émotions (deuxième sous-section). A noter que nous présentons également dans la première sous-section le modèle de Johnson et Liddell qui ne permet pas encore de décrire le non manuel, mais dont on retiendra le découpage temporel placé au cœur de la description.

La troisième sous-section liste les qualités que le système devra posséder pour la bonne conduite de cette étude et nous permettra de nous positionner par rapport à chacun des systèmes présentés.

2.1.1. Systèmes mis au point pour les LS

Nous présentons brièvement ci-dessous cinq systèmes mis au point pour décrire ou écrire les LS qui tiennent compte des GNM.

2.1.1.1. Le système Bébien

(Bébien, 1825) propose le premier système de notation connu pour les LS en 1825 dans son ouvrage intitulé "Mimographie" (Image 17). Ce système, pictural, prenait en compte quatre paramètres (" mouvement " ; " confirmation de la main " (configuration) ; " partie du corps " (emplacement) ; " points physiologiques " (mimiques faciales)).

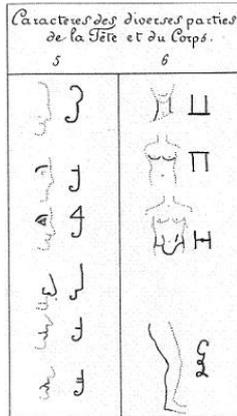


Image 17 Extrait de la seconde planche de Bébian : « Caractéristique des diverses parties de la tête et du corps »²⁹

2.1.1.2. Le système D’Sign

(Jouison, 1995) commence ses recherches sur la LSF dès 1977 et confère au corps tout entier un rôle capital. En 1985, il entreprend de réaliser un système de notation intitulé D’Sign (“ D ” pour Discours) qu’il ne cesse d’améliorer jusqu’à sa mort en 1991. Son système est basé sur l’importance du corps et du signeur en Langue des Signes. Il décrit l’orientation du regard, la localisation du signe, la configuration manuelle et le mouvement du corps (Figure 4).

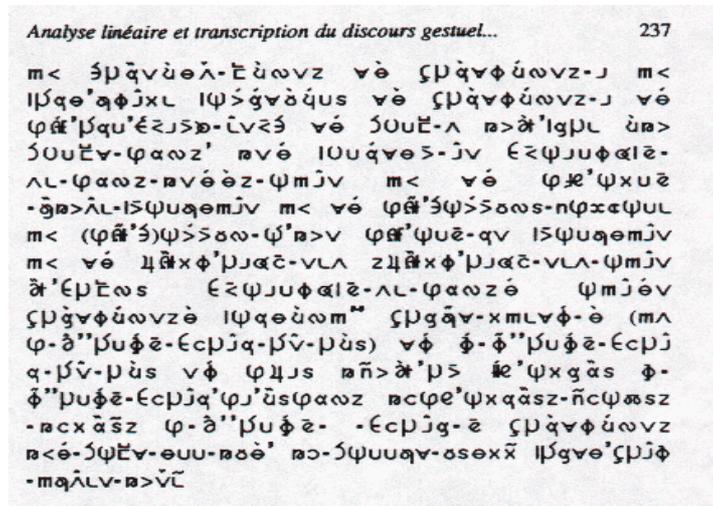


Figure 4 Résultat d’une transcription de D’Sign (Jouison, 1995, pp. 237)

La transcription commence généralement par deux symboles³⁰ représentant les “ formes-mouvements ” du visage.

²⁹ Image extraite de <http://www.cis.gouv.fr/spip.php?article1322>

³⁰ On appelle symbole tout signe conventionnel abrégé.

D'Sign permet de distinguer quatre orientations du regard et d'indiquer lorsque le signeur suit des yeux quelque chose (main, production, etc.). Il est également possible de transcrire trois types de souffle, les haussements de sourcils, ainsi qu'un ensemble d'expressions faciales du type : “ crispation du visage ”, “ courbure convexe des éléments mobiles de la face ” ou encore “ visage et orifice labial arrondis ”.

2.1.1.3. Le système HamNoSys

HamNoSys³¹ (Hambourg Notation System) (Prillwitz, Siegmund et al., 1989) est un système qui se veut “ phonétique ”. Il est utilisé pour la transcription de corpus vidéo et comprend un ensemble de plus de deux cents symboles. En réalité, il reprend les paramètres de Stokoe (l'orientation, la configuration, le mouvement et l'emplacement (Sotkoe, 1960)) et y ajoute les expressions du visage. La version initiale ne prenant pas en compte ces dernières, elles ont été ajoutées et se trouvent donc à l'écart dans le système comme nous le montre la Figure 5.



Figure 5 HamNoSys : [WHAT] dans l'énoncé “ What quote three bears ? ”

La partie gauche décrit le geste manuel : les deux ronds indiquent une symétrie inversées des mains, les symboles suivants détaillent respectivement la configuration, l'orientation de la main et de la paume, l'emplacement (au niveau des épaules), les mouvements (vers extérieur d'abord puis vers intérieur pour la main dominante, inversement pour la main dominée) et la dynamique (mouvements courts). Le cadre rouge de la Figure 5 indique les GNM. Ici, une hausse de sourcil (impliquée dans l'interrogation).

HamNoSys distingue chaque élément non manuel comme nous l'indique la Figure 6 (distinction des deux sourcils, des yeux, des oreilles, des joues, des épaules, etc.). Ces distinctions servent à l'origine à indiquer la position de la main.

Cependant, les symboles spécifiques n'étant pas encore suffisamment développés pour les GNM, les concepteurs suggèrent d'utiliser les symboles de localisation afin de faire référence à l'élément non manuel:

³¹ Site d'HamNoSys: <http://www.sign-lang.uni-hambourg.de/projects/HamNoSys.html>

“ Nonmanual components are not developed sufficiently for HamNoSys. Up to now movements of other articulating parts than hands have to be described with the known movement inventory of HamNoSys. Necessary detailed description can not be done at present. ”

Manuel HamNoSys de la version 4.0³²

	left to the left	left	between both	right	right to the right
~ eye brows	◻ ~	▪ ~	~	~ ◻	~ ◻
∞ eyes	◻ ∞	▪ ∞	∞	∞ ◻	∞ ◻
∪ ears	◻ ∪	▪ ∪		∪ ◻	∪ ◻
} cheeks	◻ }	▪ }		} ◻	} ◻

Figure 6 Extrait des symboles de localisation sur le visage d’HamNoSys³³

Concernant les mouvements, il faut également utiliser les symboles conçus pour la transcription des mouvements manuels. Cette liste comprend les mouvements de base tels que haut, bas, gauche, droite, diagonale etc. et un ensemble de mouvements plus complexes comme les mouvements circulaires. Pour transcrire un GNM, il faut donc combiner les symboles de localisation et de mouvement manuel. L’Image 18 illustre des exemples de ces combinaisons.



Image 18 Secousse de la tête - Hausse de l’épaule gauche³⁴

La mobilité du visage est particulière car elle ne dispose pas d’articulation ce qui a pour conséquence que peu de symboles de mouvements manuels sont adéquats pour les GNM.

HamNoSys est un système souvent utilisé comme point de départ des systèmes de reconnaissance et de génération automatique (par exemple SiGML cité plus haut) car son encodage informatique est relativement aisé.

³² Disponible à l’adresse suivante :

<http://www.sign-lang.uni-hamburg.de/projekte/hamnosys/hns4.0/hns4.0eng/contents.html>

³³ Disponible à l’adresse suivante : <http://www.sign-lang.uni-hamburg.de/projects/hamnosys.html>

³⁴ Disponible à l’adresse suivante:

<http://www.sign-lang.uni-hamburg.de/projekte/hamnosys/hns4.0/hns4.0eng/contents.html>

2.1.1.4. Le système SignWriting

SignWriting³⁵ (SW), (Sutton, Gleaves, 1995) est un système qui se veut “phonétique” utilisé pour écrire la LS. Il repose sur un ensemble de 425 symboles répartis en 60 groupes qui, théoriquement, peuvent représenter les configurations, orientations, emplacements, zones de contact, mouvements (des mains, de la tête du torse et des épaules), le regard et les expressions du visage de toutes les LS. Il est actuellement utilisé dans plus de 30 pays. Le système a été historiquement conçu à partir d’un autre système prévu pour la prise de notes des chorégraphes, puis adapté pour les LS.

SW utilise des symboles iconiques pour définir les parties du corps et les configurations manuelles, le noir et blanc pour l’orientation, des étoiles pour représenter les contacts et tout un ensemble de symboles de mouvement.

Les deux exemples ci-dessous (Figure 7 et Figure 8) permettent de mieux appréhender le système :

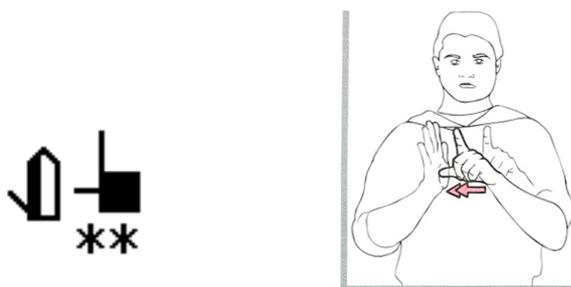


Figure 7 Écriture de [LOI] en SW³⁶.– illustration de [LOI]

La main plate est signifiée par le symbole de gauche, le noir représente le dessus de la main tandis que le blanc désigne la paume par rapport à la perception du signeur. Ici, la main de gauche est donc verticale, de profil, la paume vers la droite. Pour le symbole de droite (représentant la main droite) le carré signifie qu’il s’agit d’un point fermé, les deux traits que l’index et le pouce sont dressés et le noir indique que c’est le dessus de la main qui est face au signeur (donc paume vers l’extérieur). Les deux étoiles représentent deux contacts.

³⁵ Site de SignWriting : <http://www.signbank.org>

³⁶ A noter que ce signe, disponible sur le site officiel de SW de Belgique, ne comporte pas de flèche de mouvement de la main dominante vers la main dominée. On ne sait donc dans quel sens le mouvement doit se faire. Image disponible sur : <http://www.signwriting.org/>



Figure 8 Ecriture de [TRISTE] en SW³⁷

Ce signe, légèrement plus complexe, inclut l'expression du visage nécessaire à sa production. Le blanc de la main indique que la paume est face au signeur. Les ronds sur le visage de droite associés aux étoiles indiquent les points de contact (un au front et un au menton). La double flèche indique un mouvement vers le bas : le premier contact est en haut, puis la main descend pour le second point de contact. La forme de la bouche indique qu'il faut avoir les commissures des lèvres vers le bas.

Or, si nous regardons la vidéo associée au signe [TRISTE] sur le site de SW (et sur bien d'autres³⁸), on constate que la bouche est neutre mais surtout il apparaît un mouvement latéral de la tête et un engagement du torse probablement indispensable pour le maintien du regard et donc de l'interaction, comme l'illustre l'Image 19.

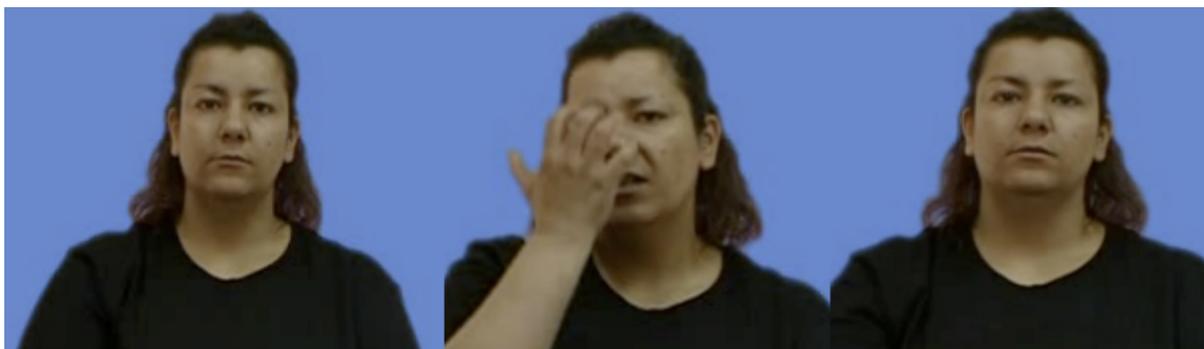


Image 19 Avant la production : tête droite; Pendant la production : tête tournée; Après la production : tête droite.³⁹

Le seul élément non manuel décrit par SW pour [TRISTE] est non pas le torse, les épaules ou les sourcils (pourtant froncés) mais la bouche. Bien sur, SW n'est pas le seul à décrire [TRISTE] de cette façon. Néanmoins, ce signe est rarement réalisé ainsi à moins d'un transfert personnel ou autre procédé iconique⁴⁰ (Image 20).

³⁷ Disponible sur : <http://www.signbank.org>

³⁸ Par exemple sur : <http://www.lsf dico-inj smetz.fr/index.php?page=video>

³⁹ Disponible sur : <http://www.pisourd.ch/index.php?theme=dicocomplet&id=585>

⁴⁰ Nous pouvons alors nous demander s'il y a une influence inconsciente des représentations prototypiques des émotions que l'on a tous, dans la transcription des LS. De même, les descriptions actuelles sont-elles crédibles dans l'ensemble de ces systèmes ?



Image 20 [TRISTE] sur le site “ LSF, dictionnaire et + ”⁴¹

Le système SW détaille relativement bien les expressions du visage, les mouvements de la tête et du visage ainsi que ceux du torse. Pour les expressions du visage, SW distingue par exemple dix groupes de symboles comprenant le front, les sourcils, les yeux et le regard, les oreilles et les joues, la respiration et le nez, la bouche, les dents, le menton et les autres parties (Sutton, 2002). Les symboles des expressions du visage (sept pour chaque sourcil, Figure 9), permettent de distinguer une élévation et un abaissement de différentes parties des sourcils : hausse de l’intérieur uniquement par exemple. De même un ensemble de symboles permet de désigner les différents emplacements pertinents de la langue (en tant qu’élément du visage) comme l’illustre la Figure 10.



Figure 9 : Symboles SW pour désigner les postures possibles des sourcils et le front⁴²

⁴¹ <http://www.lsfplus.fr/>

	BaseSymbol 572	Tongue Sticks Out Far
	BaseSymbol 573	Tongue Licks Lips
	BaseSymbol 574	Tongue Tip Between Lips
	BaseSymbol 575	Tongue Tip Touches Inside Mouth
	BaseSymbol 576	Tongue Inside Mouth Relaxed
	BaseSymbol 577	Tongue Moves Against Cheek
	BaseSymbol 578	Tongue Center Sticks Out
	BaseSymbol 579	Tongue Center Inside Mouth

Figure 10 Extrait des symboles de SW utilisés pour désigner l'emplacement de la langue⁴³

Enfin, les symboles des yeux permettent d'isoler quatre ouvertures (ouverts, plissés, moitié ouverts, grand ouverts) ainsi que trois fermetures (fermeture classique, fermeture froncée et clignement), comme nous le montre la Figure 11.

Yeux			
	Yeux ouverts		Yeux grands ouverts (étonnement)
	œil droit ouvert		Yeux à moitié ouverts
	œil gauche ouvert		œil droit à moitié ouvert
	Yeux plissés		œil gauche à moitié ouvert
	œil droit plissé		Yeux à moitié fermés
			œil gauche plissé
			Yeux fermés
			œil droit fermé
			œil gauche fermé
			Yeux fermés serrés ensembles
			œil droit à moitié fermé
			œil gauche à moitié fermé
			Clignement d'yeux
			Clin d'œil droit
			Clin d'œil gauche

Figure 11 Extrait des symboles de pour désigner les formes que les yeux peuvent adopter⁴⁴

⁴² Disponible sur: <http://www.signwriting.org/>

⁴³ Disponible sur : http://www.signbank.org/swis/data.php?subset=&sg_code=55072&bs_code=*#sgd

⁴⁴ Disponible sur: <http://www.signwriting.org/>

2.1.1.5. Le système de Johnson et Liddell⁴⁵

Contrairement aux systèmes précédents, le système de Johnson et Liddell (Liddell, 1989) repose sur la temporalité : sur l'axe du temps se trouvent un ensemble d'unités temporelles (appelées TU pour timing unit). Chacune de ces TU est soit une posture soit une transition entre deux postures. Chaque posture (et transition) est définie par un ensemble de traits binaires (par exemple opposition du pouce et des doigts) plus précis que les descriptions globales pour les formes de main par exemple. On distingue une posture⁴⁶ d'une transition de par la stabilisation de tous ses traits. Actuellement le système ne prend pas en compte les GNM mais ils devraient être intégrés selon la même approche de temporalité (Filhol, 2008).

2.1.2. Système mis au point pour la description des émotions : FACS

Après avoir établi l'aspect universel des sept émotions de base (neutralité, joie, tristesse, surprise, peur, colère et dégoût), (Ekman et Friesen, 1978) mettent au point FACS (Facial Action Coding System) pour tenter de décrire l'expression des émotions sur le visage à l'aide de critères formels. Ainsi, ils se basent sur la dynamique des muscles faciaux pour définir tous les mouvements élémentaires possibles du visage. FACS décompose ces mouvements en 46 actions unitaires provoquées par les muscles. Chaque point d'émergence d'un muscle est représenté par un numéro sur le visage. Une ligne partant de ce numéro indique "l'attachement du muscle au tissu mou du visage. Lorsqu'un muscle se contracte, le mouvement se fait dans la direction du chiffre et a pour effet de rider la peau de manière perpendiculaire à la ligne de traction des muscles" (Kaiser, 2006⁴⁷). La Figure 12 illustre le système de FACS avec l'expression du dégoût.

⁴⁵ Robert E. Johnson & Scott K. Liddell (in preparation). "Sign language Phonetics: Architecture and Description"

⁴⁶ Les postures sont appelées TU non transitoire par opposition aux TU transitoires.

⁴⁷ KAISER S. (2006). "Les expressions faciales". Support de cours de l'Université de Genève, récupéré à l'adresse suivante : <http://dokéos.unige.ch/courses/7217F/document/Historique.pdf?cidReq=7217F>

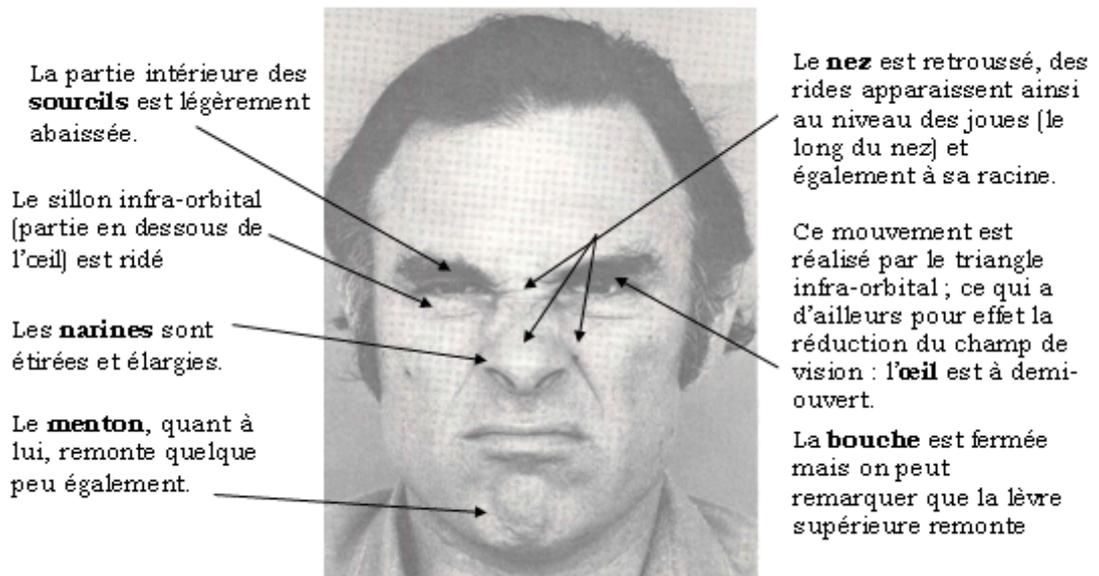


Figure 12 Expression du dégoût, analyse de FACS⁴⁸

2.1.3. Réflexion critique autour des systèmes de notation

Chaque annotation est faite dans un but et pour une étude spécifique. Par exemple, l'annotation ne sera pas la même pour une étude portant sur le lien entre la main et le regard et une étude sur la coarticulation. En effet, le choix du système de notation dépend des éléments à annoter, du degré de précision et de la finesse souhaitée, qui varieront d'une étude à l'autre. Il semble donc important de définir préalablement la granularité souhaitée (Quelle finesse ? Quel degré de précision ? etc.) afin de la confronter avec chacun des systèmes graphiques et définir ainsi lequel de ces systèmes répondra le mieux aux exigences de l'étude.

Afin de déterminer quel est le système le plus adéquat à utiliser pour notre annotation, soit le système le plus pertinent pour la description des GNM, nous proposons tout d'abord de rappeler les objectifs de notre étude afin de définir les caractéristiques souhaitées du système. Dans un second temps, nous proposons d'évaluer chacun des systèmes graphiques présentés précédemment en fonction de ces caractéristiques. Ceci nous conduira, dans un troisième temps, à proposer un regard critique sur les annotations actuelles. Enfin, nous proposerons une réflexion autour de la description d'une posture statique dont nous montrerons les limites notamment en terme de précision.

⁴⁸ Figure disponible sur : www.ecsa.ucl.ac.be/personnel/.../exprfacialefasc1.doc

2.1.3.1. Définition des attentes de l'annotation

L'objectif de cette étude est d'acquérir des connaissances précises concernant les GNM afin de les modéliser informatiquement. Concrètement, nous souhaitons savoir quand les GNM sont significatifs et comment. Il est donc nécessaire que l'annotation porte sur deux niveaux : au niveau de l'énoncé pour noter la présence d'un GNM et au niveau local pour décrire la structure des GNM.

Par ailleurs, nous considérons que les connaissances sur les GNM ne sont pas encore exhaustives et souhaitons donc que l'annotation permette de relever le maximum d'indices possibles, même si leur importance n'a pas été démontrée jusqu'ici.

Tout en étant fine, l'annotation devra respecter le principe d'économie, devra être simple et suffisamment lisible pour faciliter l'analyse ou tout du moins la partie qualitative de l'analyse. En effet, plus la description est fine et précise, plus le nombre d'informations est important et donc l'annotation difficilement réalisable et visuellement peu perceptible. Nous souhaitons donc un système simple d'utilisation et qui favorise l'analyse qualitative de par son aspect visuel.

Enfin, l'annotation devra être rapidement exploitable par un module d'animation du visage. Cela permettra de procéder à des tests de validation de la méthodologie durant l'annotation et de proposer des ajustements et des modifications si nécessaire. Cependant, proposer une description formelle nécessite de disposer de données précises afin que chaque symbole permette de caractériser un élément précis. Le critère de précision est donc un critère indispensable à la description formelle. Les critères de précision et de description formelle seront donc abordés dans une même section.

Le système utilisé pour l'annotation des GNM devra donc posséder les caractéristiques suivantes :

- **Précision et finesse** : Le système de notation doit permettre de distinguer toutes les postures statiques, même deux proches afin d'étudier le nombre de variations possibles et les marges de liberté des mouvements. Il doit en outre permettre d'annoter des éléments dont l'importance n'a pas été établie.

- **Simplicité** : L'annotation doit être simple, lisible et économique. Par “ lisible ” nous n'entendons pas que le système doit permettre une traduction française des transcriptions mais plutôt qu'il doit permettre une visualisation de ce qu'il se passe au niveau du non manuel. Nous souhaitons par exemple que les clignements des yeux apparaissent immédiatement et de façon claire, sans avoir à les chercher parmi un ensemble de symboles compliqués.

2.1.3.2. Confrontation entre les attentes et les systèmes

Nous proposons ci-dessous de comparer successivement chacune des attentes dégagées avec les systèmes de notation présentés précédemment, afin de déterminer si ces derniers permettent la description souhaitée.

Précision et finesse

Il s'agit ici tout d'abord de ne pas assimiler deux signes légèrement différents, même si cette différence n'est pas perceptible visuellement, mais de les distinguer afin de déterminer par la suite si cette différence est discriminante ou non. Par exemple, l'Image 21 nous montre deux hausses de sourcils d'amplitude différente. Nous cherchons à savoir quel système permet de les distinguer.



Image 21 Exemple de deux hausses de sourcils différentes⁴⁹.

FACS répond à cette attente. De par la décomposition des muscles faciaux, il permet de distinguer les hausses de différentes parties du sourcil. De plus, l'intensité du mouvement est notée par une lettre (de A à E selon l'intensité du mouvement) ce qui permet de distinguer deux hausses de degrés différents. La description non pas de la posture des éléments non manuels, tels que les yeux ou les sourcils, mais des muscles mis en œuvre pour obtenir cette posture est particulièrement intéressante dans ce système.

⁴⁹ Images extraites du corpus LS-Colin (voir partie 2.1 du chapitre 3).

D'Sign, HamNoSys et SignWriting ne permettent pas de distinguer deux hausses de sourcils différentes au niveau de l'amplitude (une hausse plus marquée que l'autre) car ils proposent un seul symbole pour tout type de hausse. Toutefois, SW permet de distinguer deux hausses dont l'unique différence serait la partie du sourcil levée. Par exemple, il distingue une hausse de l'extérieur du sourcil d'une hausse de l'intérieur du sourcil.

Par ailleurs, les symboles d'HamNoSys comme ceux de D'Sign ou de SW offrent des informations sur une posture statique ce qui empêche la description de certains GNM. En effet, un GNM comporte au moins deux postures : une posture de départ et une posture d'arrivée voire à nouveau la posture de départ (dans d'autres cas il s'agit de l'adoption d'une nouvelle posture, voir Figure 13).

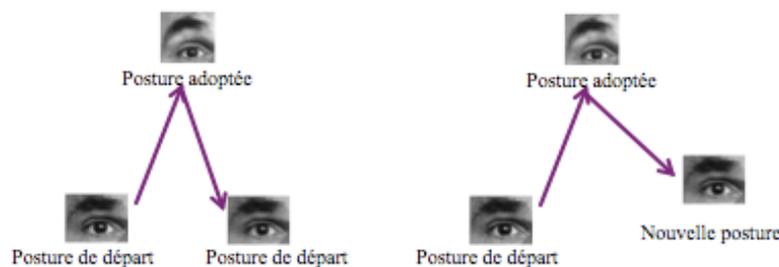


Figure 13 postures contenues dans un GNM

Or, si l'on veut décrire puis analyser ce GNM, il est nécessaire de distinguer chacune de ces postures afin d'en connaître les propriétés et ce pour deux raisons essentielles : d'une part le GNM peut être interrompu et avoir un rôle différent de s'il avait été produit entièrement.

Par exemple, la Figure 14 illustre une hausse de sourcils interrompue par une baisse. La hausse n'est donc pas achevée et pose donc la question du label qui doit être attribué à ce GNM à savoir hausse ou baisse des sourcils ? D'autre part, le sens peut être différent en fonction du temps écoulé entre la première et la seconde posture (autrement dit le temps de réalisation). Par exemple, dans le cas d'une hausse de sourcil (avec autres GNM associés) la hausse rapide peut indiquer la surprise tandis qu'une hausse lente peut exprimer l'incrédulité (sans soudaineté). Cet exemple est illustré par la Figure 15.

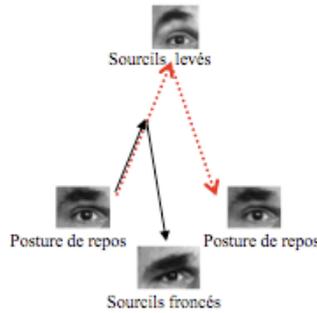


Figure 14 Schéma d'une hausse de sourcils normale (traits pointillés) et d'une hausse interrompue suivie d'un froncement (traits pleins)

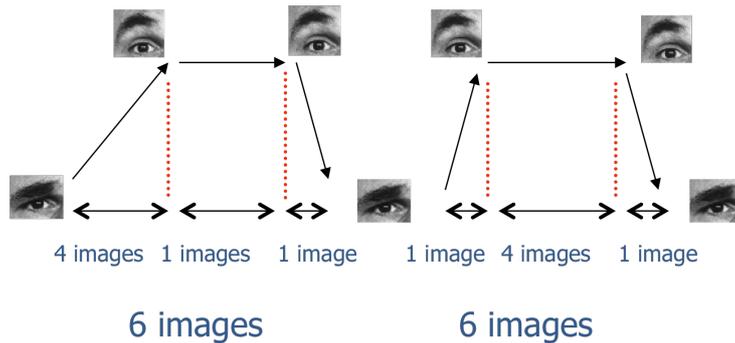


Figure 15 Deux hausses de sourcils à temps de réalisation variable qui peuvent exprimer l'incrédulité (à gauche) et la surprise (à droite)

La Figure 14 illustre bien les limites de ces systèmes. En effet, ce type de hausse est différent d'une hausse aboutie mais ne sera pourtant pas distingué. De même que pour la Figure 15 où la durée finale est la même mais le temps de réalisation est différent et probablement le sens véhiculé aussi. Cependant, il existe n'existe qu'un label qu'il est possible d'attribuer pour définir les hausses ou baisses sans tenir compte de l'interruption ou du temps de réalisation variable.

Le système de Liddell et Johnson est le seul qui permettrait (s'il traitait du non manuel) de prendre en compte ces différentes réalisations car il repose sur la notion d'axe temporel et de mouvement.

Enfin, il s'agit également de décrire toutes les postures, mêmes celles dont l'importance n'a pas été établies. Or, la plupart des systèmes actuels ne permettent pas de répondre à cette attente. En effet, ils proposent une liste fermée de postures possibles il n'est donc pas possible de décrire de nouvelles postures. FACS, en permettant de combiner l'action des différents muscles, est le seul à répondre à ce critère.

De tous les systèmes, seul FACS répond au critère de précision il est donc le seul susceptible de proposer une description formelle répondant à nos attentes. Ses descriptions reposent sur les mouvements des muscles, donc sur un critère objectif et sans ambiguïté. Il est donc le seul système à proposer une description formelle pouvant être utilisée par un module d'animation du visage.

Simplicité, lisibilité et économie

HamNoSys ne répond à aucun de ces critères ne serait-ce que de par sa transcription linéaire. En effet, les LS s'exprimant par le biais de nombreux articulateurs simultanément, inscrire à la suite des phénomènes qui se passent en réalité en simultanément rend la relecture plutôt délicate et l'ensemble du système peu économique, comme l'illustre la transcription de [MONTAGNE] (Figure 16).

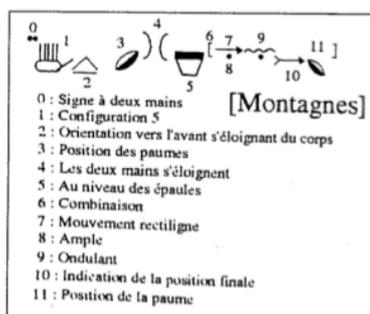


Figure 16 Détail de la transcription de [MONTAGNE] d'après (Martin Dupond, 1995).

De plus, concernant le non manuel, le fait que les symboles utilisés soient ceux utilisés pour décrire les mouvements manuels engendre des ambiguïtés et nécessite une interprétation humaine. Par exemple, lorsqu'un symbole de mouvement de la main est utilisé pour caractériser un mouvement de la tête, ce mouvement est difficilement interprétable comme l'illustre la figure ci-dessous.



Figure 17 HamNoSys : tête tournée

La Figure 17 symbolise “ tête tournée ”. Or, le symbole de mouvement semble indiquer un tour complet de la tête, ce qui est évidemment impossible. Comme il ne fournit pas d'autres indications, son interprétation est délicate.

FACS ne correspond pas non plus à nos attentes de simplicité et de lisibilité. En effet, il demande de longues heures d'apprentissage pour ensuite ne pouvoir être déchiffré que par des personnes initiées. Il ne permet donc pas la simplicité requise pour l'analyse future des données. Il est cependant relativement économique : chaque description se compose d'un chiffre et d'une lettre parmi un nombre très restreint de lettres et de chiffres.

SW, de par sa nature plutôt iconique, répond presque au critère demandé. Presque, car le nombre de ses symboles est beaucoup trop élevé ce qui rend le système complexe et peu économique.

D'Sign demande un apprentissage de la signification de l'ensemble des symboles ce qui le rend difficilement lisible et peu pratique pour l'analyse envisagée.

Ainsi, aucun des systèmes ne répond à ces critères qui devront faire l'objet d'une attention particulière lors de la réflexion sur la méthodologie à adopter pour l'annotation.

2.1.3.3. Évaluation des systèmes

Nous avons dressé le panorama de nos attentes et les avons confrontées avec chacun des systèmes. Nous proposons maintenant une rapide évaluation de chacun des systèmes afin d'en relever les avantages et les inconvénients.

D'Sign

D'Sign propose des symboles pour la description des mouvements du corps d'une richesse et d'une finesse remarquable. Cependant, ce système se veut une forme graphique de la LSF ce qui, en plus de la forme même des symboles, le rend difficilement exploitable pour notre étude.

HamNoSys

Ce système présente l'avantage d'être adapté à l'animation de signeur virtuel de par son encodage informatique relativement aisé. De plus, certains de ses symboles sont simples et peuvent tout à fait correspondre à ce que nous souhaitons. Par exemple les flèches de mouvements non complexes (droite, gauche, etc.) manquent certes d'informations sur l'amplitude du mouvement mais sont simples, lisibles et aisément interprétables.

Nous aurions donc pu envisager d'utiliser ce système comme point de départ tout en l'adaptant notamment en créant de nouveaux symboles pour plus de précision. Cependant, le fait que ce système ne tienne pas du tout compte de la dynamique des mouvements ni du contexte en fait un système inadéquat pour notre étude. Nous retenons toutefois ses points forts pour la création de la méthodologie finale.

SignWriting

Les symboles utilisés pour la langue (en tant qu'élément du visage) semblent pertinents pour notre étude. Concernant les autres parties du corps, le trop grand nombre de symboles rend ce système peu économique et certaines des postures proposées sont peu réalisables : par exemple la "joue droite tendue vers le bas" (voire même difficilement conceptualisable). Par ailleurs, SW permet une distinction des différentes parties des sourcils mais ne permet pas leurs combinaisons (par exemple l'extérieur levé et l'intérieur baissé).

En outre, le système manque de précision quant à l'amplitude des mouvements. Pour toutes ces raisons nous n'utiliserons pas ce système pour notre étude mais en retenons la distinction des parties des sourcils.

FACS

Cet outil semble le plus adéquat pour notre étude : basé sur des critères formels et objectifs, il propose une description fine de l'ensemble des éléments non manuels en décomposant chacun de ces éléments en fonction des muscles en action. De plus, il permet de distinguer les différents mouvements selon leur intensité ce qui nous semble particulièrement pertinent.

FACS est cependant long à maîtriser (environ 100 heures d'apprentissage pour appréhender le système), difficilement déchiffrable, mais surtout il manque de précision pour les études temporelles. Cet outil ne propose en effet aucune description de l'aspect temporel, d'où un temps de réalisation des mouvements qui reste méconnu. C'est pourquoi nous pourrions nous inspirer de ce modèle sans pour autant pouvoir l'utiliser.

La Langue Vocale

A ces différents systèmes s'ajoute la LV sous forme de gloses⁵⁰ ou de description qui permet, sous certaines conditions, de remplir chacun des critères individuellement mais jamais la totalité. En effet, de manière générale, l'utilisation du lexique de la LV engendre des descriptions de nature symbolique (ex : "lèvres serrées et rétractées" ou "moue dubitative" (Cuxac, 2000)). Or, si la signification globale de ces descriptions (du moins une représentation de cette signification) émerge rapidement chez l'interlocuteur, l'accès à cette signification ne pourra se faire sans instanciation pour les logiciels d'animation. Nous pouvons par exemple nous demander qu'est-ce qu'une "hausse de sourcil" ? Il manque ici des informations caractérisant cette hausse. Bien sur, il est possible de fournir ces informations grâce au lexique des LV mais cela rendrait alors la description compliquée, difficilement déchiffrable et peu économique.

En réalité, l'ensemble des systèmes s'est heurté au problème de l'équilibre entre précision et simplicité. Plus on souhaite décrire de façon précise, plus cela va demander du temps et de la complexité d'où une lisibilité et une exploitation analytique immédiate difficile. Ceci est d'autant plus vrai avec les mots des LV. En effet, afin de décrire de façon claire, sans présumé ni ambiguïté, il faudrait des termes extrêmement précis et nombreux. D'où au final, pour des raisons d'économie, l'utilisation d'abréviations puis de symboles. Quand bien même une description textuelle arriverait à remplir l'ensemble de ces critères sans équivoque, elle ne pourrait être traitée par le module d'animation de signeur virtuel sans un (long) traitement intermédiaire. Notre annotation n'utilisera donc pas le lexique des LV.

2.1.3.4. Regard porté sur les annotations actuelles : remise en cause de la description de la posture finale

Les différents systèmes que nous avons vus s'attellent à décrire la posture finale (appelée aussi posture adoptée). Cependant, ces descriptions posent de nombreux problèmes pour le traitement des GNM. Nous avons par exemple évoqué la difficulté de distinguer deux postures proches (comme deux hausses de sourcils) et de déchiffrer des postures sans interprétation humaine (nombreux présupposés).

⁵⁰ Par Glose nous entendons un mot de la LV utilisé pour nommer un signe. Par exemple [HOMME] désigne le signe « homme » en LSF.

De plus, comme nous l'avons vu, les descriptions actuelles attribuent un label à la posture finale. Or, nous avons rappelé qu'un GNM contient au moins deux postures (la posture initiale et la posture adoptée) et qu'il semble important de les distinguer pour ne pas assimiler les GNM interrompus et les GNM aboutis. Par ailleurs, ces deux postures sont reliées par un mouvement pour passer de l'une à l'autre, dont la durée peut varier tout comme la posture finale peut-être tenue plus ou moins longtemps. Vergé souligne d'ailleurs l'importance de la durée de réalisation de la fermeture des yeux lors d'un clignement (Vergé, 2001), autrement dit la durée pour passer de paupières ouvertes à paupières fermées. La structure temporelle des GNM (par exemple le temps pour passer de la première à la seconde posture) n'est pourtant pas prise en compte par les annotations, elle n'est donc pas étudiée et l'étendue de son influence est inconnue. Pour la production par un signeur virtuel, cela se traduit par une réalisation des mouvements à durée standard dont seule la tenue peut varier.

Décrire les postures finales peut également s'avérer particulièrement problématique et engendrer de nombreuses ambiguïtés et imprécisions concernant la bouche. En effet, la bouche peut adopter un très grand nombre de postures différentes en raison des différents axes sur lesquels chacune de ses parties peut évoluer : les commissures des lèvres peuvent varier sur l'axe horizontal (droite – gauche) et l'axe vertical (haut – bas), tandis que les lèvres évoluent aussi bien sur un axe frontal (avant – arrière) que sur un axe vertical (haut – bas). Dès lors, la bouche peut adopter un grand nombre de postures en jouant sur les différents axes, postures difficiles à nommer et à distinguer.

L'ensemble des problèmes évoqués ci-dessus est du aux limites que posent la description d'une posture. De plus, ce type de description ne prenant pas en compte la dynamique non seulement de la langue mais en particulier des GNM, nous considérons qu'il n'est pas adéquat à notre étude et qu'il sera nécessaire de mener une réflexion sur une description plus adaptée.

2.2. Quel support d'annotation?

En s'appuyant sur le caractère multidimensionnel de la LSF, des systèmes adaptés à l'annotation des LS ont été mis au point et en particulier les tableaux (ou partitions) qui permettent de respecter la simultanéité des éléments intervenant dans les LS sur un axe temporel.

Par la suite, d'autres besoins ont émergé, notamment celui de procéder à des annotations informatisées afin de permettre une diffusion des annotations entre linguistes (plus malléables que les supports papiers utilisés précédemment) mais également de les modifier plus rapidement. D'autre part, s'est révélé l'intérêt d'intégrer la vidéo du corpus à l'annotation afin de les synchroniser. C'est ainsi que sont apparus les premiers logiciels d'annotation multimédia.

Nous proposons tout d'abord une brève définition du système de partition illustrée de quelques exemples. Dans un second temps, nous présenterons les logiciels d'annotation multimédias existants.

2.2.1. Les partitions

Une partition est un système de transcription multilinéaire sous forme d'un tableau. Elle permet " une lecture temporelle d'un paramètre (axe horizontal) ou une lecture simultanée de tous les paramètres (axe vertical) " (Sallandre, 2003). Chaque ligne représente un des éléments que l'on souhaite annoter (en fonction de l'étude) et la description peut se faire avec n'importe lequel des systèmes graphiques évoqués précédemment ou avec des mots de la langue vocale.

Nous présentons ci-dessous trois exemples de partition : les deux premiers utilisent le lexique du français et se distinguent par le choix et l'organisation des éléments à annoter. La dernière partition illustre un exemple de description mixte : à l'aide de la LV et de SW. Soulignons que les partitions présentées ci-dessous ne sont pas issues d'une analyse centrée sur les GNM, même si ceux-ci sont présents. Il ne s'agit donc pas de savoir si l'une ou l'autre est pertinente pour notre étude, mais simplement d'illustrer les possibilités d'une partition.

2.2.1.1. La partition de Cuxac

La partition de (Cuxac, 2000) est composée de plusieurs lignes parallèles (illustrée par la Figure 18) puis, sur une deuxième page, une reprise et une numérotation des éléments de la partition permettant une traduction approchée puis littérale de l'énoncé.

MD	Etats-Unis	
D et G		
MG		
C		
V		
R		public _____
MF		(garantie) lippe quantitative (très) OB-

Figure 18 Extrait de la partition de Cuxac p.352

Sa partition distingue cinq parties. La première comprend les trois premières lignes (“ MD ” : main droite ; “ D et G ” : mains droite et gauche ; “ MG ” : main gauche) et donne des informations sur la position et les mouvements manuels. L’avantage des trois lignes est de pouvoir distinguer aisément les signes produits avec une main des signes produits avec deux mains en symétrie ou non.

La seconde partie est la ligne notée C qui décrit le mouvement du corps (“ recul, à gauche ”, “ recentre ”, etc.). La troisième partie est très proche de la seconde et explicite les mouvements de la tête (“ recul tassé ”, “ oui oui ”, etc.) sur la ligne notée V. Les deux dernières parties sont consacrées au regard noté R (“ balayage ”, “ neutre ”, “ vers public ”, “ fermé ”...) et aux mimiques faciales sur une ligne notée MF (“ surprise ”, “ fronce ”, “ grimace ”, “ embarrassé ”, péjorative ”...). La partition de Cuxac permet de mettre en relation le non manuel avec le procès de l’énoncé.

2.2.1.2. La partition de Bouvet

(Bouvet, 1996) propose une partition qui se distingue de celle de Cuxac de par les lignes décomposant les signes (Figure 19). En effet, la première ligne est consacrée aux sourcils (élevés ou non) et la deuxième au regard (porté sur l’intersigne⁵¹ ou non). L’ordre différent entre les lignes de cette partition et les lignes de celle de Cuxac s’explique par le fait qu’ici, l’ordre respecte visuellement ce qui est perçu par l’observateur (la partition représente le visage, donc les sourcils sont au-dessus du regard par exemple) contrairement à l’organisation de la partition de Cuxac qui débute par les signes manuels qui représentent, selon lui, “ un moyen de repérage plus stable ”.

⁵¹ On appelle intersigne un interlocuteur d’une discussion en LS.

La partie manuelle de la partition de Bouvet comprend quatre lignes qui distinguent la main dominée de la main dominante tandis que la ligne médiane ne diffère pas d'une partition à l'autre et comprend les deux mains ensemble. Ces lignes ne peuvent être remplies que de traits indiquant de quelle main il s'agit. Une quatrième ligne permet alors de noter le signe qui est produit, par exemple [POISSON] [VOIR] etc. Les paramètres manuels ne sont donc pas du tout détaillés contrairement aux GNM. La ligne suivante intitulée " tête " permet de noter à l'aide de flèches les mouvements de la tête et leur lien avec les gestes manuels. Enfin, la dernière partie est composée de quatre lignes pour la partie inférieure du visage : gonflement des joues, bouche, labialisation, double menton. Là encore le début et la durée de réalisation sont notés à l'aide de traits parallèlement aux gestes manuels.

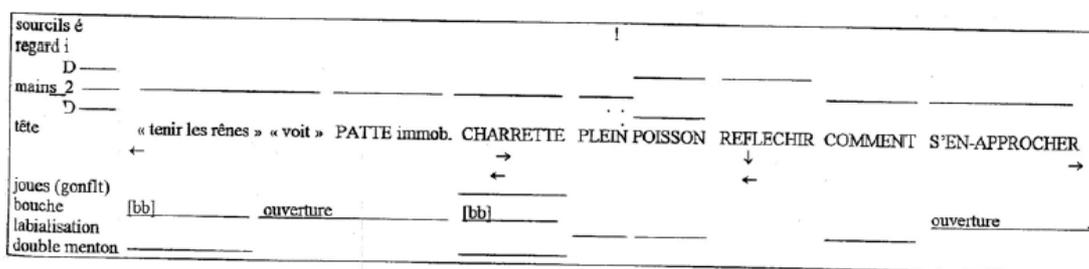


Figure 19 Extrait de la partition de Bouvet p.66

Si la partition de Bouvet précise davantage les GNM que celle de Cuxac, elle ne décrit toutefois pas ces derniers et note uniquement leur présence.

2.2.1.3. La partition mixte de Fusellier-Souza et Boutora

Dans le cadre d'une étude contrastive sur les moyens d'annotation, les auteurs confrontent en particulier les systèmes de transcription en partition et le système SW (Fusellier-Souza et Boutora, 2005). Ainsi, elles ajoutent une ligne à leur partition pour ajouter les symboles de SW correspondants. (Figure 20).

00:31	00:32	00:32	00:32	00:32	00:33
60	61	62	63	64	65
					
45	46	47	48	49	50
					
(?)					1
	1	1	1	1	
D-----	D-----	E°dv'conf'-----	D-----	D-----	D-----
moi	"accueillir"	"heberger"	"au sud de Paris"	"au sud de Paris"	"le trajet"
MP			DS	DS	

Figure 20 Extrait de la partition

Nous mentionnons cette étude principalement afin d'illustrer un exemple mixte : utilisation de symboles et de LV dans la partition. Cependant, cette étude repose sur l'opposition de SW et des partitions alors qu'ils sont, selon nous et dans le cadre de cette étude, complémentaires et que l'opposition se situe dans la partition au niveau de l'utilisation des symboles ou de la LV. En effet, dans le cadre bien précis de l'utilisation de SW pour les GNM, il est tout a fait possible d'utiliser les symboles de SW séparément pour chaque élément du visage dans une partition. Par exemple, une ligne dédiée aux sourcils qui utilisera les symboles SW appropriés, une ligne pour les yeux, etc. Cette utilisation des symboles de SW est probablement plus difficile pour d'autres éléments que les GNM en raison de l'interdépendance des paramètres (emplacement, mouvement, etc.) dans SW.

2.2.1.4. Bilan

Contrairement à une annotation linéaire, les annotations qui utilisent les partitions permettent de prendre en compte la simultanéité des éléments et respectent donc la multilinéarité des LS. De plus la lecture des éléments est plus aisée et facilite les analyses. Il est possible de choisir quels éléments doivent être annotés en parallèle. Ce type de support permet donc d'adapter l'annotation aux objectifs de l'étude.

Ces annotations en partition ont durant longtemps (et encore un peu aujourd'hui) été produites manuellement sur une feuille de papier. Or, hormis la difficulté de respecter la temporalité, il était en plus relativement délicat de fusionner deux annotations. Ainsi, des chercheurs ont mis au point des logiciels de partitions multimédias.

2.2.2. Les partitions multimédias

Les partitions multimédias sont des logiciels associant la vidéo du corpus à l'annotation, comme nous l'avons brièvement abordé. Globalement, ces logiciels possèdent tous une (ou plusieurs) fenêtre(s) pour la (les) vidéo(s) en haut, une fenêtre de commande de la vidéo à gauche, une fenêtre pour placer les frontières des segments à annoter et enfin une fenêtre d'annotation. Nous présenterons ci-dessous les spécificités de chacun des quatre logiciels qui existent actuellement.

2.2.2.1. Le logiciel Sign Stream

Sign Stream⁵² est un logiciel d'annotation réalisé par le "National Center for Sign Language and Gesture Ressource" de l'Université de Boston. Il offre la possibilité de traiter jusqu'à quatre vidéos et d'annoter des dialogues. Ses partitions possèdent une sous partie consacrée aux expressions du visage comme l'illustre la Figure 21.

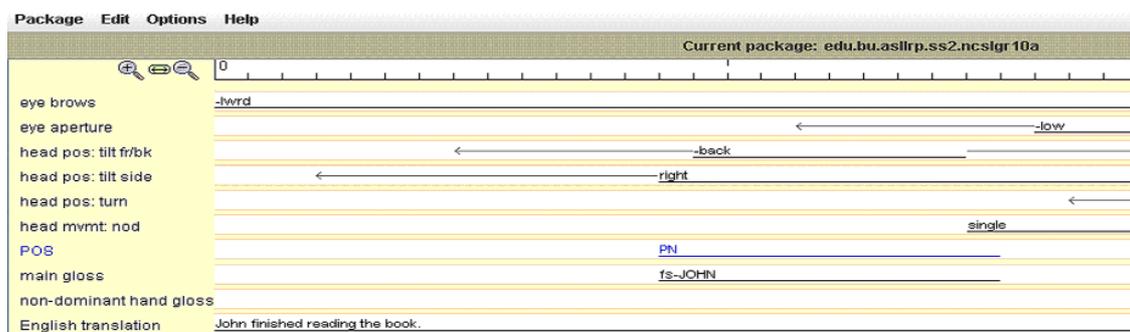


Figure 21 Sous-partie de la partition Sign Stream

⁵² Site de SignStream : <http://bu.edu/asllrp/SignStream/>

Le logiciel dispose d'un outil de recherche qui permet, de façon très intuitive (via des listes déroulantes), de retrouver un phénomène. Cependant, l'annotation ne peut utiliser que des gloses accompagnées de traits indiquant le début et la fin du phénomène. Il n'est donc pas possible d'annoter à l'aide d'images ou de symboles. Comme expliqué dans la partie précédente, l'utilisation de la LV présente de nombreux inconvénients pour cette étude, nous n'utiliserons donc pas ce logiciel.

2.2.2.2. Le logiciel Elan

Elan a été réalisé par le "Max Planck Institute for Psycholinguistics"⁵³ (Pays-Bas). Il dispose également d'un outil de recherche et permet de traiter jusqu'à quatre vidéos en même temps. Ce logiciel présente l'avantage de proposer une exportation des annotations sous différents formats et outils d'analyse. Il permet également l'affichage de courbe. Cependant, les annotations ne peuvent utiliser que des caractères Unicode et l'aspect visuel de l'annotation ne facilite pas une analyse qualitative comme l'illustre l'Image 22.

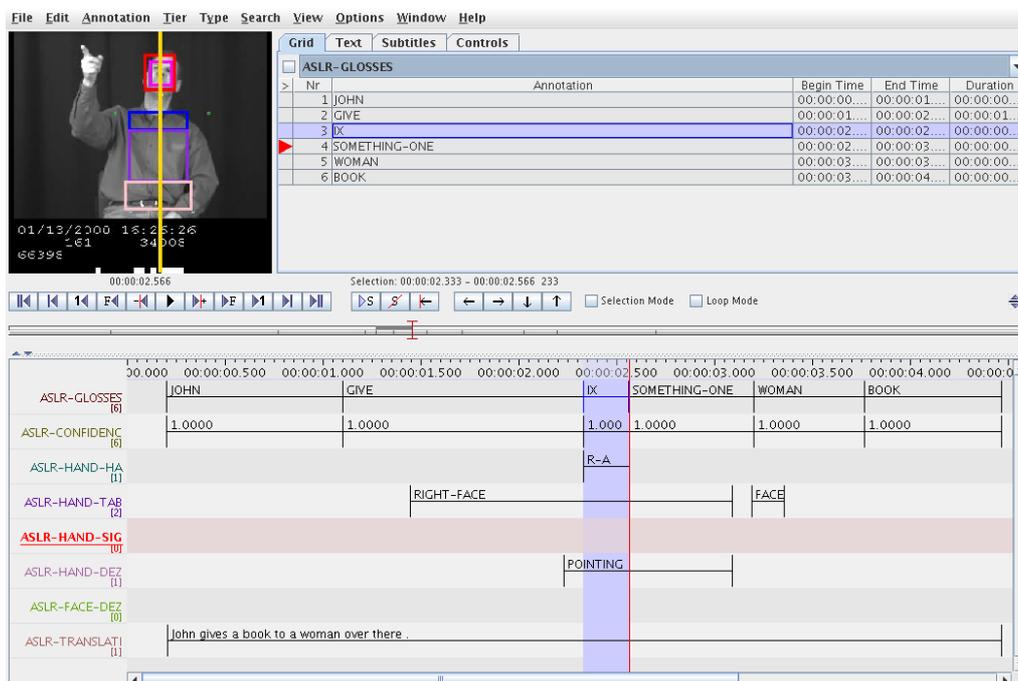


Image 22 Illustration du logiciel Elan

⁵³ Pour télécharger Elan : <http://www.mpi.nl/tools/elan.html>

2.2.2.3. Le logiciel Ilex

Ilex a été conçu par l'institut de l'université de Hambourg⁵⁴. Dans sa version actuelle, l'annotation ne permet d'utiliser que des symboles HamNoSys. Comme les deux précédents systèmes, un outil d'analyse est intégré. Celui-ci repose sur des requêtes qui permettent notamment d'isoler toutes les réalisations d'un phénomène dans l'annotation. La structure de l'annotation est réalisée par l'utilisateur. L'avantage majeur de ce logiciel est d'être associé à une base de données qui permet de stocker toutes les réalisations possibles d'un signe donné ainsi que sa description ce qui facilite l'annotation et l'analyse (Image 23).

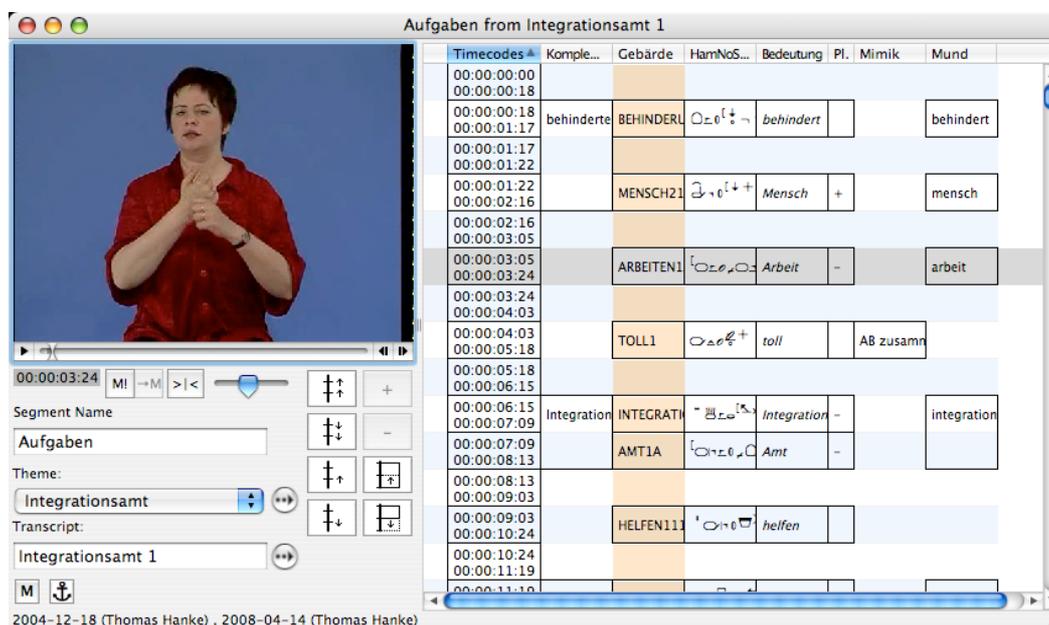


Image 23 Illustration du logiciel Ilex.

2.2.2.4. Le logiciel Anvil

Anvil⁵⁵ a été créé par Mickael Kipp (Kipp, 2004) au DFKI (University of the Saarland and German Research Center for Artificial Intelligence)⁵⁶ en Allemagne. Ce logiciel permet de choisir les éléments que l'on souhaite annoter ainsi que le format de la description (à l'aide d'images, de symboles ou encore de mots), comme l'illustre l'Image 24. Toutes ces caractéristiques sont définies dans un fichier XML.

⁵⁴ Site d'Ilex : www.sign-lang.uni-hambourg.de/ilex/

⁵⁵ Annotation of Video and Spoken Language.

⁵⁶ Plus d'informations à l'adresse suivante : <http://www.dfki.de/~kipp/anvil>

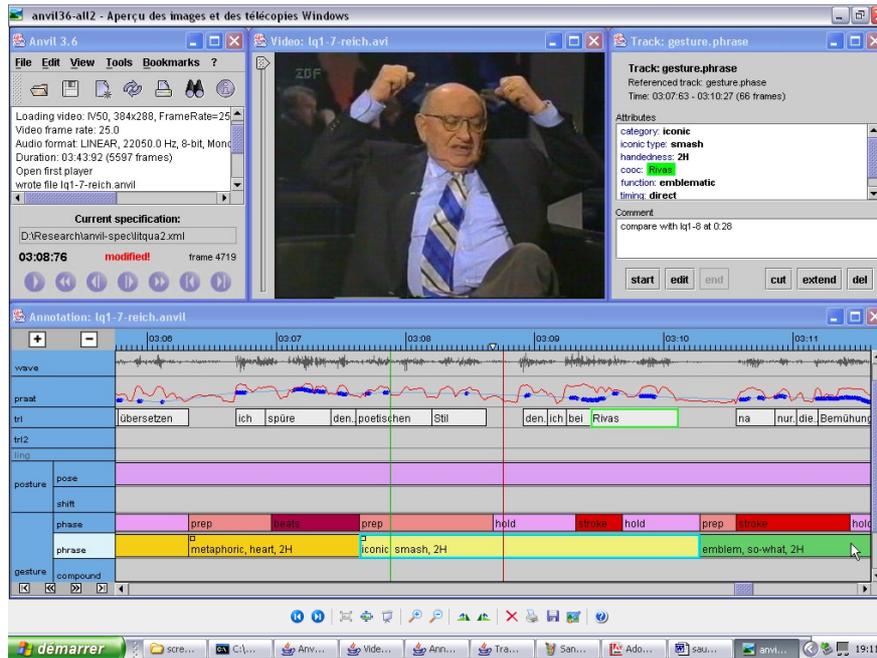


Image 24 Interface d'Anvil⁵⁷

Il est également possible d'afficher des courbes ainsi que le signal acoustique (lorsqu'il s'agit d'une LV). Enfin, les éléments de l'annotation, encore appelés "segments", peuvent prendre une couleur spécifique si l'utilisateur le souhaite notamment dans le but de mettre en avant certains phénomènes. Anvil est également le seul à permettre une annotation à l'aide de mots, d'images ou de symboles de son choix. En revanche, Anvil ne dispose d'aucun outil d'analyse incorporé et ne permet pas de traiter plusieurs corpus simultanément. De plus, il ne permet pas une exportation de l'annotation obtenue mais uniquement des données brutes (format XML ou tableur), et est en perpétuelle évolution par le concepteur (et uniquement lui, le code n'étant pas en open source). Chaque nouvelle version dispose de nouvelles fonctionnalités mais n'en propose plus certaines des versions précédentes. Il faut ainsi jongler entre différentes versions pour utiliser l'ensemble des fonctionnalités.

Chacun de ces logiciels de partitions multimédias offre différentes possibilités. Par exemple, Ilex est particulièrement intéressant pour l'annotation rapide de signes manuels tandis que Sign Stream est plus adéquat pour le traitement simultané de plusieurs vidéos. Ainsi, le choix du logiciel d'annotation dépend des attentes de l'annotation pour l'étude souhaitée.

⁵⁷ Interface tirée du site de M Kipp (<http://www.dfki.de/~kipp/anvil>)

3. Conclusion

Comme nous l'avons vu, le TAL, et plus particulièrement la génération automatique de LSF, peut permettre d'apporter une aide concrète à la communauté sourde. Les travaux actuels s'attachent à prendre en compte les spécificités des LS et en particulier l'espace et l'iconicité. Les GNM, s'ils sont pris en compte, ne peuvent pas être générés de manière automatique en raison du manque de connaissances dont nous disposons en particulier concernant leur structure interne. Ce constat nous a conduit à proposer une nouvelle étude s'inscrivant dans une recherche linguistique de corpus. Cette étude nécessite une première phase d'annotation de corpus.

Nous nous sommes alors interrogés sur la façon de décrire les GNM et avons passé en revue les différents systèmes existants à savoir aussi bien les outils de transcription et de notation que les systèmes de description des émotions. Ces systèmes se sont avérés insuffisants en terme de précision ou trop complexes pour l'étude souhaitée mais ont permis de préciser nos exigences en terme de finesse et de granularité tout en offrant une base pratique de réflexion pour l'élaboration d'un outil descriptif plus adapté à cette étude.

Nous avons ensuite présenté les logiciels d'éditeurs de partitions en relevant les avantages et inconvénients de chacun. En fonction de la structure et de la forme que nous donnerons à l'annotation, nous choisirons le logiciel qui correspondra au mieux à nos attentes.

Nous pouvons alors mener une réflexion autour de la création de la méthodologie la plus adaptée à notre étude, qui réponde aux attentes évoquées précédemment, tout en respectant l'équilibre entre précision et simplicité.

Chapitre 3 Vers une nouvelle approche descriptive

L'observation des annotations actuelles permet une mise en évidence des problèmes soulevés par le traitement des GNM. A partir de ces problèmes, la structure de l'annotation peut prendre forme en tentant d'apporter des solutions tout en répondant aux exigences fixées préalablement. Il s'agit de la phase d'élaboration de la structure qui comprend aussi bien le choix du corpus à annoter (à la base de l'évaluation), le choix des éléments à étudier ainsi que le choix du format des attributs (symboles, mots, images, etc.).

Ce chapitre propose tout d'abord, à partir des limites dégagées précédemment en matière de description des GNM, de définir les différentes contraintes et objectifs de la réalisation d'une nouvelle structure d'annotation. Dans une seconde partie, nous présenterons les solutions apportées par notre nouvelle approche descriptive ainsi que la méthodologie adoptée.

1. Délimitation des contraintes et objectifs

Nous proposons ci-dessous de replacer notre annotation dans un cadre théorique et pratique puis de définir les contraintes relatives au corpus et les objectifs que nous souhaitons atteindre.

1.1. Inscription dans un cadre théorique et pratique

Du point de vue linguistique, nous formulons l'hypothèse que les mouvements des articulateurs en LSF reposent sur la notion d'amplitude. En effet, chaque signeur possède son rythme, comme chaque locuteur des LV a son propre flux de parole. Cela signifie que chaque mouvement se définit par rapport au rythme propre au signeur. Ainsi, nous considérons que la différence entre [MARCHER] et [COURIR] est uniquement la vitesse à laquelle le second est réalisé par rapport au premier. Ce paramètre devra être pris en compte par la structure de l'annotation. Par exemple, une forte hausse des sourcils se devra d'être définie par rapport à la hausse moyenne de chaque signeur.

Du point de vue informatique, l'annotation a pour objectif une description des GNM qui sera utilisée pour animer des signeurs virtuels dans le cadre de la génération automatique de LSF. Dans cette perspective, elle doit reposer sur des critères formels et objectifs afin d'éviter toute mauvaise interprétation par les systèmes. En effet, pour que le personnage de synthèse adopte la posture demandée, une définition de cette posture en termes clairs et précis, évitant au maximum les présupposés, doit être proposée.

Par ailleurs, l'annotation doit permettre une implémentation rapide dans un système de génération automatique afin de procéder à des tests sur un visage animé. Ces tests permettront d'affiner et d'ajuster les descriptions. La structure de l'annotation doit donc être rapidement exploitable.

1.2. Contraintes dans le choix du corpus

Le choix du corpus est essentiel à l'élaboration de la structure d'annotation. En effet, une observation rigoureuse du corpus permet de définir les éléments qu'il semble essentiel de décrire et offre une vision plus claire de la façon d'y parvenir. C'est pourquoi le choix du corpus doit être l'objet d'une première longue réflexion pour répondre au mieux aux exigences de l'étude choisie.

A l'oral, dans les LV, les gestes co-verbaux varient d'une personne à l'autre en fonction de son origine socioculturelle, de sa région ou encore de son sexe. Il est donc fort probable qu'il en soit de même pour les LS. Ceci pose le problème suivant : comment savoir si un GNM fait partie de la langue ou alors s'il est propre au signeur. Pour évaluer cet aspect, le corpus doit donc être composé de différents signeurs d'origine, d'âge et de sexe différents.

De plus, afin de permettre une analyse précise des éléments du visage, la vidéo doit être d'excellente qualité et disposer d'un gros plan du visage. En effet, il peut être délicat de distinguer si les yeux sont fermés ou plissés lorsque l'image est petite. Cependant, si la majorité des éléments à étudier se situent sur le visage, il n'en reste pas moins que les coudes et les mains doivent également être visibles, soit pour l'annotation, soit pour la compréhension du message.

Enfin, le contexte d'enregistrement du corpus et le thème abordé sont deux éléments essentiels à prendre en compte dans le choix du corpus. En effet, selon l'objet de l'étude, le corpus choisi ne sera pas le même. Pour notre analyse, celui-ci doit refléter au maximum la réalité et être le moins biaisé possible. Par exemple, on peut supposer qu'un sourd ne signera pas de la même manière s'il est en présence d'Entendants qui débute en LSF ou en présence d'autres sourds, il peut notamment articuler ou pas le mot français qu'il est en train de signer. Ce phénomène n'est qu'un exemple parmi de nombreuses situations pouvant biaiser la nature de la langue filmée. La valeur de l'étude qui en découlerait risquerait d'être limitée. Ainsi, il serait préférable que l'enregistrement du corpus se déroule dans un contexte de personnes signantes (afin de minimiser l'impact des Entendants) et utilise le moins possible de procédures d'élicitation (afin de disposer de données naturelles).

Par ailleurs, si le thème abordé peut être imposé, il serait pertinent que le degré de liberté d'expression des signeurs soit vaste. En effet, plus le discours sera imposé, plus les productions des signeurs risquent d'être formatées ou du moins biaisées par la situation.

1.3. Définition des objectifs

La méthodologie de l'annotation doit permettre de résoudre les problèmes évoqués précédemment. Il s'agit tout d'abord de décrire toutes les postures complexes que certaines parties du corps (comme la bouche et la tête) peuvent adopter de la façon la plus précise possible et en prenant en compte la dynamique du mouvement et de la LSF.

Il faut ensuite distinguer les différentes réalisations possibles d'un même mouvement. Par exemple si un GNM n'est réalisé qu'en partie, l'annotation doit le distinguer du même GNM qui serait abouti. Si un GNM est plus accentué qu'un autre, l'annotation doit également être capable de les différencier.

Par ailleurs, la dynamique des GNM doit être prise en compte afin de permettre une étude spécifique de leur structure temporelle. Ainsi, la posture de départ, la transition 'posture de départ - posture adoptée' et la posture adoptée doivent être distinguées.

Enfin l'annotation doit considérer le principe des amplitudes relatives en LSF en proposant des descriptions reposant sur ce concept. En effet, chaque élément doit être décrit en fonction d'une position standard. Ainsi, une description sera du type "sourcils élevés de x% par rapport à la hausse moyenne". Cette dernière sera propre à chaque signeur.

Concernant le support, les logiciels d'annotation multimédia sont particulièrement pertinents pour notre étude et ce pour plusieurs raisons. D'une part ils permettent de lier mais surtout de synchroniser la vidéo du corpus avec l'annotation. Cette synchronisation offre la possibilité de considérer l'axe temporel et donc de connaître la durée précise de chaque phénomène observé. D'autre part, lorsqu'un segment est isolé, il est possible de le visionner plusieurs fois, image par image afin notamment d'ajuster avec précision l'emplacement des frontières du segment (le début et la fin). Il permettra donc d'obtenir la description précise que nous souhaitons.

Nous avons rappelé ci-dessus les différentes contraintes et objectifs de notre étude. A partir de ces remarques, la structure d'annotation peut être définie. Nous présentons dans la partie suivante les outils sur lesquels vont s'appuyer cette étude.

2. Choix des supports de l'étude

Nous proposons de présenter les outils que nous avons choisis pour cette étude à savoir le corpus et le logiciel d'annotation. La première section décrit l'origine et le contexte d'enregistrement du corpus LS-COLIN. La seconde section présente le logiciel Anvil et s'attache en particulier à en expliquer son fonctionnement ses fonctionnalités.

2.1. Corpus retenu

Considérant la complexité et le matériel nécessaire à la constitution d'un corpus LSF exploitable pour notre étude, nous avons observé les corpus pré-existant afin d'observer si un correspondait à nos critères. Suite aux contraintes que nous avons dégagées, nous avons choisi un extrait du corpus LS-COLIN qui comprend treize signeurs sourds racontant individuellement leur journée du 11 septembre 2001.

L'intérêt de ce corpus est d'une part un grand nombre de signeurs de différents âges, sexes et origines, et d'autre part une grande liberté d'expression : chaque signeur raconte ce qu'il souhaite concernant le 11 septembre.

2.1.1. L'origine du corpus LS-COLIN

Le corpus LS-COLIN, dont est issu le corpus de cette étude, comprend au total six heures d'enregistrement, et a été créé lors du projet LS-COLIN "Langues des Signes : Analyseurs privilégiés de la faculté de langage ; apports croisés d'études linguistiques, cognitives et informatiques (traitement et analyse d'image) autour de l'iconicité et de l'utilisation de l'espace". Ce projet a vu le jour dans le cadre de l'Action Concertée Incitative (ACI) cognitive 2000 "Langage et Cognition" (Cuxac, LS-COLIN, 2002) mise en place par le Ministère de l'Education Nationale, de la recherche et de la Technologie en 1999. Pendant 24 mois, il a regroupé les Universités de Paris 8, Paris 4, Toulouse 3 et le CNRS (LIMSI). Un des buts a été de mener des études linguistiques et informatiques des Langues des Signes en synergie. Cela a inclus l'étude de nouveaux outils utiles à l'analyse des Langues des Signes aussi bien pour les linguistes que pour les informaticiens.

2.1.2. La réalisation du corpus LS-COLIN

D'une étude approfondie des besoins de chacun et des thèmes de recherches ont émergé les caractéristiques que le corpus devait adopter. Ainsi, concernant le matériel tout d'abord, trois caméras numériques professionnelles ont été utilisées (Figure 22) : une de face en plan américain (permettant diverses analyses telles que les paramètres manuels ou encore une analyse globale des liens entre les différents paramètres, etc.), une de face en contre-plongée (pour une analyse des GNM ou exclusivement des expressions faciales), et enfin une dernière caméra au dessus du signeur, particulièrement intéressante pour étudier les mouvements du tronc intervenant entre autres dans les phénomènes de transfert de grande iconicité.

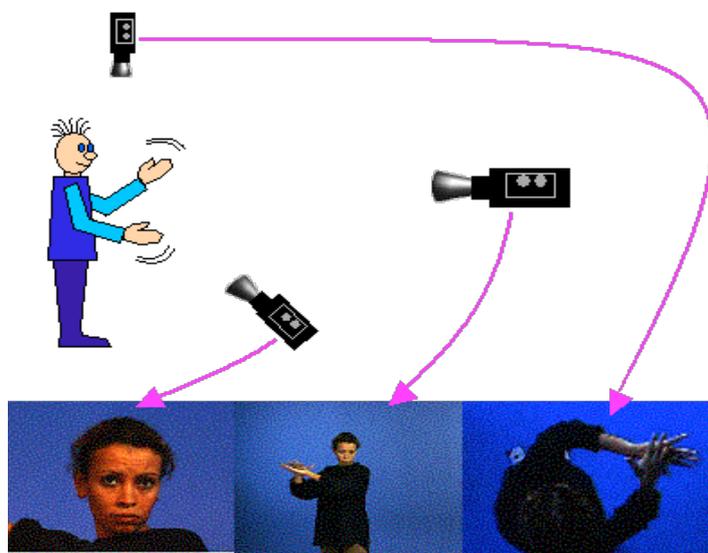


Figure 22 Plan des caméras

Afin de réduire au maximum les ombres, six projecteurs ont été utilisés ainsi que des réflecteurs de part et d'autre du signeur.

En outre, il était demandé aux signeurs d'être vêtu de noir et de porter des manches longues. Ainsi, cela permettait de trancher avec le fond bleu du plateau. Signalons que les manches longues noires trouvent leur légitimité dans le fait qu'uniquement les mains sont ainsi de couleur "peau" ce qui est très utile pour le traitement automatique d'image tel que la détection des mains.

Par ailleurs, les treize signeurs ont été recrutés pour leur variété sociolinguistique et pour leurs compétences en LSF (voir Tableau 3).

Signeur	Age	Sexe	Région	Réside
S1	23	M	Isère	Lyon
S2	23	F	Toulouse	Toulouse
S3	27	F	Parisienne	Paris
S4	29	F	Algérie	Marseille
S5	29	M	Pyrénées	Lyon
S6	29	M	Parisienne	Paris
S7	30	M	Dijon	Paris
S8	31	F	Parisienne	Paris
S9	31	M	Parisienne	Paris
S10	39	M	Maroc	Paris
S11	55	M	Algérie	Paris
S12	65	F	Allier	Paris
S13	NC ⁵⁸	M	NC	NC

Tableau 3 Méta données des signeurs

Le tournage a eu lieu à L'INJS (Institut National de Jeunes sourds) de Paris ce qui a permis une immersion dans le milieu de la surdité et une influence moindre des Entendants. Enfin, chaque signeur était reçu par un linguiste qui lui expliquait en LSF le déroulement du tournage.

Les signeurs s'adressaient directement au linguiste qui les avait reçus. Il leur était demandé de raconter deux histoires imagées sans texte, en noir et blanc puis de s'exprimer au choix sur un sujet (soit sur le passage à l'euro, soit sur les événements du 11 septembre 2001), d'expliquer une recette de cuisine, de choisir un thème du programme d'enseignement en linguistique de Paris 8 et de l'aborder à la façon d'un cours (uniquement pour les signeurs ayant suivi ce cursus) et enfin de raconter une nouvelle fois la première histoire filmée. Cette grande diversité de thèmes a permis d'étudier différents phénomènes de la LSF. Par exemple, les histoires sont très utiles pour étudier les structures de grande iconicité (terme explicité dans la section 1.2.3 du chapitre 1).

Dans le cadre de notre étude, il était nécessaire que la situation soit la moins imposée possible et ce afin de déterminer dans une situation " naturelle " quels GNM étaient utilisés et à quel moment. Les cours de linguistique ont donc été écartés car, du fait de la nature explicative du discours, il est probable que les GNM diffèrent d'une situation " classique ". De même, le corpus d'histoire a été constitué en utilisant des images où l'histoire était racontée. Le biais étant alors trop important, nous ne l'avons pas utilisé.

⁵⁸ NC signifie Non Communiqué

Les récits du 11 septembre et de l'euro semblaient donc les moins contraints et les plus pertinents. Le choix a alors été totalement arbitraire par rapport à l'intérêt personnel face à ces sujets.

2.2. Logiciel d'annotation retenu

Nous avons retenu le logiciel Anvil, essentiellement car il est le seul actuellement qui permet d'inclure des symboles de notre choix dans l'annotation. De plus, il est simple d'utilisation (très intuitif), personnalisable pour tout type d'analyses, et permet de définir une couleur pour chaque attribut, comme expliqué plus en détails ci-dessous, ce qui facilite l'analyse qualitative.

Ce logiciel est composé de la façon suivante (voir explications en dessous de la Figure 23) :

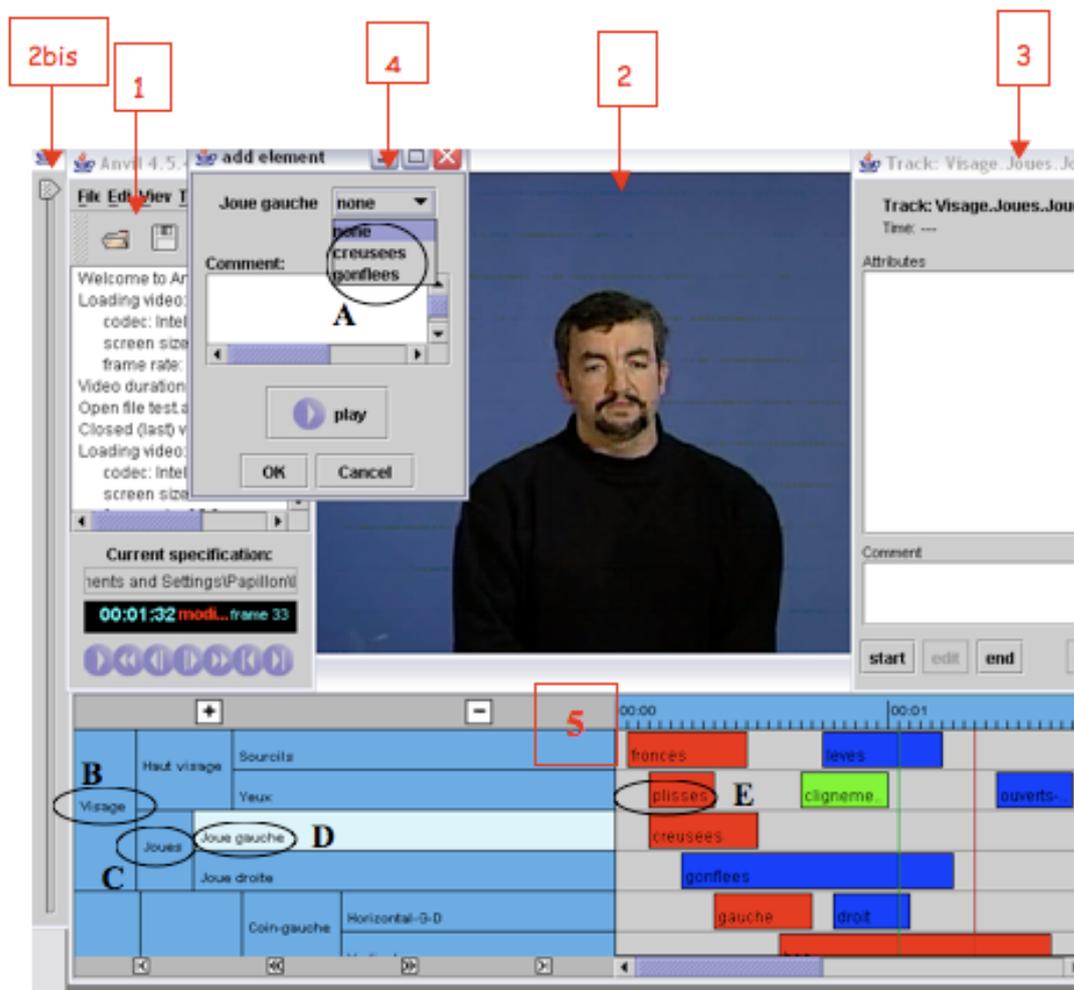


Figure 23 Interface d'Anvil

- La fenêtre (1) contient les fonctionnalités standard telles que “ ouvrir ”, “ enregistrer ”... ainsi que le clavier de commandes de la vidéo. Il est par exemple possible de regarder la vidéo image par image, par séquence ou encore à vitesse normale.
- La fenêtre centrale (2) permet la visualisation de la vidéo avec la possibilité de choisir la vitesse de défilement (ascenseur (2bis)).
- La fenêtre de droite (3) affiche les attributs (et commentaires) donnés à une séquence. Des boutons de commande, en bas de la fenêtre, permettent en outre de placer les frontières des segments de chaque piste⁵⁹.
- La fenêtre (4) apparaît lorsque l'on a placé les frontières d'un segment (se reporter à la lettre E sur la Figure 23). Elle permet de lui donner les attributs (lettre A sur la Figure 23) souhaités et d'ajouter des commentaires.
- Enfin, la fenêtre d'annotation est située tout en bas (5). Un des grands avantages d'Anvil, comme nous l'avons évoqué succinctement, est qu'il permet une grande liberté aussi bien dans le choix des éléments à annoter que dans le choix de la forme des attributs. La fenêtre d'annotation (5) est créée par Anvil à partir d'un fichier de spécification XML entièrement défini par l'utilisateur. Nous pouvons observer sur la Figure 23 la première fenêtre d'annotation réalisée pour cette étude, elle a depuis subi de nombreuses modifications.

La partie gauche de la fenêtre d'annotation énumère les parties que nous souhaitons étudier (les yeux, les sourcils, etc.). Chacun de ces composants peut se décomposer autant de fois qu'on le souhaite, par exemple, les “ mains ” peuvent se décomposer en “ main gauche ” et “ main droite ” puis ces deux composants peuvent encore se décomposer en “ configuration ”, “ emplacement ”, etc. Nous obtenons donc un groupe (lettre B sur la Figure 23) dans un groupe (C) autant de fois qu'il est nécessaire. Les groupes finaux (les plus détaillés) s'appellent des “ pistes ” ou encore des “ tracks ” (D). Ce sont sur les pistes que l'annotation sera finalement réalisée sous forme de segments (E) auxquels vont être affectés un ou plusieurs attributs (“ plissés ” dans le cas de E).

⁵⁹ Par pistes, nous entendons les éléments finaux que nous allons décrire par exemple « joue gauche » : (D) sur la Figure 23.

La partie droite de la fenêtre contient l'annotation en tant que telle, à savoir les segments et les attributs. Les éléments annotés prennent une couleur spécifique en fonction des attributs qui leur sont affectés. Cela est déterminé dans le fichier de spécification et sera ici utilisé pour valider des hypothèses et isoler des phénomènes de manière visuelle donc qualitative. Par exemple, si l'on considère qu'à chaque fois que les joues sont creuses, les sourcils sont froncés, on peut se demander dans quelles proportions ces éléments sont liés. Nous leur attribuons la même couleur et observons leur présence dans l'annotation. Ainsi, nous pouvons savoir après une rapide observation de l'annotation si les deux phénomènes sont systématiquement présents ensemble ou non et isoler les cas où un seul de deux phénomènes est présent.

Par ailleurs, outre la possibilité d'étudier leur corrélation, il est possible d'isoler des éléments extérieurs à un phénomène. Par exemple si toutes les pistes contiennent du jaune au même moment et qu'une des pistes contient du bleu, on peut supposer que l'élément intrus dénote un nouveau phénomène (si le bloc jaune reflète la modalité, l'élément bleu peut être représentatif d'un adjectif par exemple). Enfin, certains phénomènes isolés, comme par exemple le clignement des yeux, seront mis en avant grâce à une couleur spécifique (vert dans le cas présent). Tout ceci est illustré dans la Figure 24.

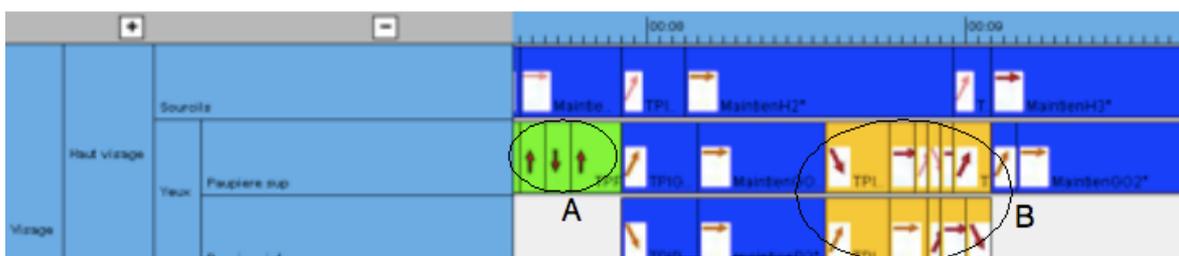


Figure 24 Extrait de l'annotation

Nous pouvons observer ici les clignements de yeux (bloc A) ainsi que la présence de jaune (bloc B) à l'intérieur d'un bloc bleu (bloc foncé). Les couleurs permettent ainsi de mettre en avant un phénomène particulier qu'il nous est alors possible d'étudier. Ici, le signeur vient de demander " quoi ? " et il répond à sa question par : " 11 septembre ". Le levé de sourcils et les yeux grands ouverts (bloc foncé) sont donc tenus jusqu'à la fin de sa réponse. Cependant, le signeur s'aperçoit qu'il a oublié de préciser qu'il parlait d'une date, il signe donc : [QUOI ONZE DATE ONZE SEPTEMBRE]. Le plissement des yeux (bloc B) dure le temps de [DATE] mais les sourcils restent levés jusqu'à la fin de l'énoncé.

3. Présentation de la méthodologie

Notre méthodologie d'annotation comporte deux types de description : une description qualitative qui, de par son caractère visuel, facilite l'analyse en contexte et une description quantitative qui permet d'obtenir des données numériques de la position des éléments annotés.

3.1. Description qualitative

Nous proposons ci-dessous une méthodologie d'annotation qui permet de décrire les GNM de façon simple et économique. Nous explicitons les grands principes sur lesquels nous nous sommes basés dans une première section, puis nous présentons la méthodologie dans une seconde section. Enfin, nous proposons un bilan de cette méthodologie ce qui nous permettra d'envisager une méthodologie supplémentaire, plutôt quantitative, nécessaire afin de valider et de définir notre description.

3.1.1. Grands principes

3.1.1.1. Description du mouvement et non de la posture

Comme nous l'avons vu, un GNM comporte deux postures (la posture initiale et la posture adoptée). Nous considérons que la description ne doit pas porter sur ces postures mais sur les mouvements permettant de passer de la première posture à la seconde. Par exemple, si l'on reprend notre figure représentant les mouvements d'un GNM (Figure 25), il s'agira de décrire les flèches qui conduisent d'une posture à l'autre. Ainsi, la description de la dynamique des GNM sera placée au cœur de la méthodologie et il sera possible de donner la description précise de chaque élément (paupière, bouche, etc.) mais également de décrire leurs *phases* et *amplitudes de réalisation* (explicités dans la section 3.1.2)

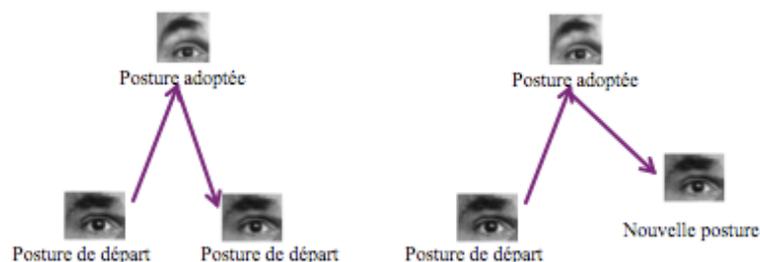


Figure 25 Les mouvements compris dans un GNM

3.1.1.2. Décomposition des éléments à annoter

Notre étude ne portant que sur les GNM, les éléments à annoter sont en nombre restreint et sont donc traités individuellement. Il s'agit d'observer les yeux, les sourcils, la bouche, les joues, la tête, les épaules et les coudes. Cependant, afin de limiter autant que possible les descriptions faisant appel à des présupposés, nous prenons le parti de décomposer au maximum chacun des éléments à annoter afin que les annotations ne concernent que des parties d'élément. Ainsi au lieu de décrire la bouche et de lui donner comme attribut " ouverte en avant ", nous décrivons séparément la lèvre supérieure et lèvre inférieure.

Concernant la bouche, une autre distinction a lieu : les commissures des lèvres sont dissociées des lèvres. Ainsi, nous obtenons un groupe (bouche) puis deux sous-groupes (lèvres et coins) puis nous distinguons lèvre supérieure de lèvre inférieure et le coin gauche du coin droit. La bouche est donc décomposée en quatre composantes dont les mouvements se décrivent de façon simple à savoir " gauche, droite, haut, bas, avant, arrière ". Toutes les formes de la bouche sont donc descriptibles grâce à des mouvements simples de ces quatre composantes. D'autre part, ce type de découpage permet d'éviter toute ambiguïté issue de la description d'un élément possédant une multitude de mouvements complexes, et donc de positions possibles, tout en minimisant le risque de présupposé.

Les paupières sont, elles, traitées individuellement : " paupière supérieure " et " paupière inférieure ". Cela s'explique par le fait que ce qui distingue un " plissement des yeux " d'un " regard vers le bas " n'est pas le degré d'ouverture des yeux mais le fait que dans le premier cas la paupière du bas est levée tandis que dans le second cas elle est en position standard (Image 25). Cependant, la description des mouvements de la paupière du bas n'a lieu que lorsque le mouvement est flagrant et discriminatoire d'un regard vers le bas.



Image 25 Yeux plissés – Regard vers le bas (Images extraites de notre corpus)

La langue quant à elle pose un tout autre problème : elle peut être visible ou non. Cependant, dans chacun de ces cas, elle peut modifier la forme d'un autre élément. Il faudra donc parfois noter la présence de la langue qui déforme un élément, sans que celle-ci ne soit visible (nous pourrions observer une protubérance sur la joue par exemple, mais pas la langue en tant que telle).

Enfin, les joues sont dissociées, d'une part en raison de la langue qui peut en déformer une et pas l'autre, et d'autre part car un certain nombre de réalisations implique l'utilisation d'une seule joue ([ENNUYEUX] par exemple).

3.1.1.3. Décomposition des mouvements

Comme nous l'avons vu pour la bouche, il est intéressant de décomposer les éléments pour obtenir des composantes qui n'évoluent que sur des axes simples. De la même manière, les mouvements complexes sont décomposés en mouvements simples. Par exemple, un mouvement en diagonal (Figure 26) est décomposé en un mouvement latéral et un mouvement vertical. Ceci permet une annotation plus aisée et offre en outre la possibilité d'annoter l'amplitude de chaque axe : à savoir pour une diagonale, il est par exemple possible de noter si le mouvement latéral est plus accentué que le mouvement vertical ou inversement. Toutes les diagonales peuvent donc être décrites ainsi que tout mouvement complexe. Le principal avantage de cette méthode est de disposer d'un nombre restreint de symboles pour annoter tout type de mouvements.

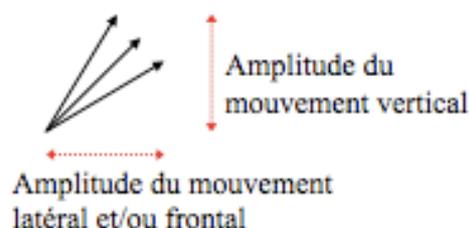


Figure 26 Décomposition d'une diagonale

Les pistes vont donc décrire les mouvements des composantes des éléments à annoter : par exemple si l'on reprend la bouche, nous retrouvons les deux sous-groupes (lèvres et commissures). Chacun de ces sous-groupes est à nouveau décomposé en deux groupes (lèvre supérieure et lèvre inférieure pour le premier groupe et commissure gauche et droite pour le second). Enfin ces derniers sont décomposés en fonction de l'axe sur lequel ils peuvent évoluer (horizontal, vertical, etc.), ces derniers groupes sont les " pistes " (ou " track "). La Figure 27 montre un extrait de ces pistes.

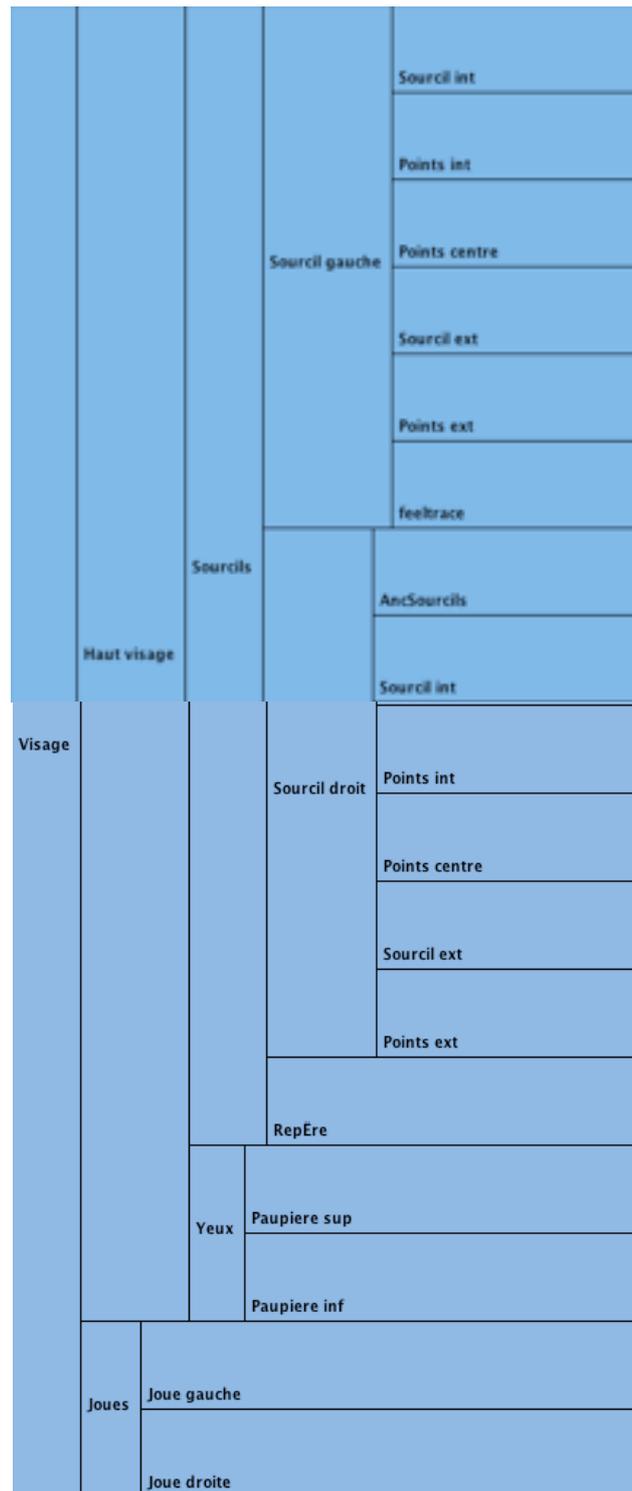


Figure 27 Structure des pistes de l'annotation

3.1.2. Choix des attributs

Dans le but d’obtenir des annotations claires et objectives, il semble plus pertinent d’utiliser des symboles à la place des mots de la langue française. En effet, cela permet d’éviter d’une part toute ambiguïté de sens (que signifie “ sourcils levés ” ?) et d’autre part tout type d’amalgame entre les deux langues (français et LSF). De plus, cela présente l’avantage de pouvoir exprimer plusieurs phénomènes grâce à un seul symbole, comme nous allons le voir plus en détails ci-dessous. Enfin, les symboles facilitent la perception visuelle des structures récurrentes de l’annotation, en favorisant ainsi l’analyse syntaxique..

Les symboles utilisés sont simples et en nombre restreint (Figure 28) pour une clarté et une appropriation plus aisée. Cela est rendu possible par la décomposition maximale de tous les mouvements comme nous l’avons détaillé précédemment. Ainsi, toutes les composantes peuvent utiliser les mêmes symboles.



Figure 28 Symboles les plus fréquents

3.1.2.1. Distinction de trois amplitudes de réalisation

Nous déclinons trois amplitudes : une standard et deux autres (atténuée et amplifiée) - distinguées dans l’annotation par l’intensité et l’épaisseur du symbole (Figure 29). Plus le mouvement est amplifié, plus le symbole est foncé et épais.



Figure 29 Amplitudes de réalisation

Au cours de l’annotation, nous allons donc relever une réalisation qui semble comporter une amplitude standard. Par “ standard ” il faut entendre une réalisation d’amplitude moyenne présente plus souvent que les autres. Elle servira ainsi de référence pour définir les autres, respectant ainsi le principe des amplitudes relatives des LS.

La direction des flèches de la Figure 29 indique qu’il s’agit de flèches de contact (voir Tableau 4). En effet, pour distinguer un contact d’un “ presque contact ” (par exemple des yeux fermés par rapport à des yeux très plissés donc à peine ouverts), nous utilisons deux types de flèches différentes : les flèches de contact (verticales) et les flèches de “ non contact ” (diagonales).

Fonctions		Symboles
Contact		↓ ↑
Elévation		
TPIH		↗
Tenue		→
TPHI		↘
Droite		
TPID		→
Tenue		—
TPDI		←
Avant (ex : langue, épaule...)		
<u>TPIAvt</u>		●
Tenue		—
<u>TPAvtl</u>		●

Fonctions		Symboles
Contact		↓ ↑
Baisse		
TPIB		↘
Tenue		→
TPBI		↗
Gauche		
TPIG		←
Tenue		—
TPGI		→
Arrière (ex : langue, épaule...)		
<u>TPIArr</u>		●
Tenue		—
<u>TPArrl</u>		●

Tableau 7 Symboles utilisés

3.1.2.2. Distinction des phases de réalisation

Par phases de réalisation, nous nommons les différents mouvements qui constituent un geste comme explicité dans la Figure 25 (section 3.1.1.1). Chaque GNM est donc décrit de la façon suivante : mouvement conduisant de la posture initiale à la posture adoptée, tenue de ce mouvement puis mouvement conduisant à la posture initiale ou à une autre posture (exemple Figure 30). Les phases de réalisation sont ainsi isolées et peuvent être sujettes à une analyse particulière (par exemple : le sens est-il le même lorsqu’une posture est adoptée très lentement que lorsqu’elle est rapide ?).

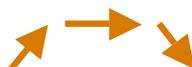


Figure 30 Phases de réalisation d’une hausse sans contact suivie d’une tenue puis d’un retour à la posture initiale

Les symboles de la figure ci-dessus sont les phases d'une hausse (de sourcil, d'épaule, de paupière, etc.) sans contact. En effet dans le cas des yeux et de la bouche, il peut y avoir contact (entre les deux paupières ou les lèvres), dans ce cas, la première et dernière flèche seront verticales. S'il s'agit d'une baisse (de paupière, sourcil, etc.) et non d'une hausse, les trois phases seront également représentées mais les flèches de mouvements seront dans le sens inverse.

Comme nous pouvons voir dans le Tableau 4, la liste des symboles utilisés est extrêmement limitée pour plus d'économie. Cela est rendu possible par le fait que chaque élément se définit par rapport au contexte. Ainsi, les tenues ne sont par exemple pas distinguées entre elles (tenue en position⁶⁰ haute et tenue en position basse sont décrites avec le même symbole) mais elles sont définies par le symbole précédent. Outre l'économie, l'intérêt de ce système est de placer les amplitudes relatives au cœur de la description de la LSF.

Par exemple, dans la Figure 31, le premier symbole de la première ligne indique une hausse atténuée de sourcil.

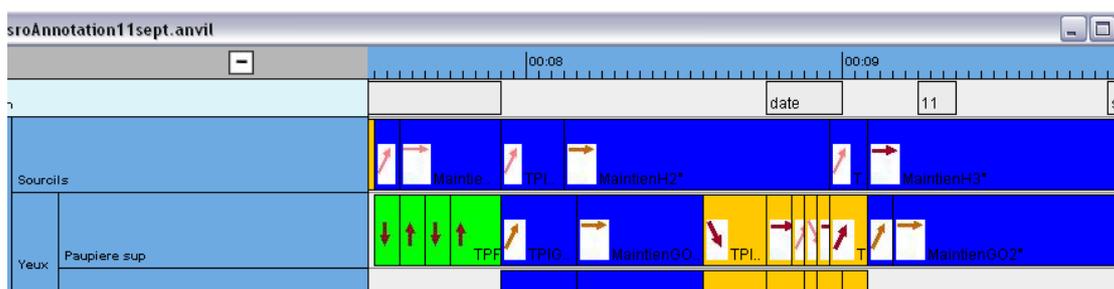


Figure 31 Extrait de notre annotation

Puis, après un court maintien de la position (deuxième symbole), le signeur lève davantage les sourcils (troisième symbole). Ceux-ci seront alors en position élévation moyenne (quatrième symbole). La flèche utilisée ensuite pour décrire la nouvelle hausse de sourcil (cinquième symbole) est une flèche de degré 1 (soit une élévation faible). Comme les sourcils étaient déjà élevés (de niveau 2), cette flèche indique une élévation faible par rapport à leur position de départ, soit une position haute (niveau 3) au final, sixième symbole.

⁶⁰ Nous utilisons « position » dans le sens du « lieu où une chose est placée ». <http://www.mediadico.com/dictionnaire/definition/Position/1>

Ce système permet ainsi de décrire tous les mouvements non manuels de la LSF en usant d'un minimum de symboles, leur définition reposant sur la notion d'amplitude relative.

3.1.2.3. Traitement des joues

Les joues quant à elles évoluent sur un nouvel axe qui comprend les mouvements intérieur/extérieur. Cet axe n'étant pas utilisé pour décrire les autres éléments, de nouveaux symboles ont été créés. Cependant, étant assez proche du mouvement avant/arrière les mêmes symboles ont donc été repris avec une orientation différente (Tableau 5). A noter que dans les cas où la langue, non visible, est à l'origine de la déformation de la joue, un astérisque sera ajouté aux symboles, illustré dans le Tableau 5 pour "joue droite gonflée".

Fonctions	Symboles
Intérieur (joues creusées)	
Extérieur (joues gonflées)	
tendue	

Tableau 5 Symboles utilisés pour les joues

Les caractéristiques de l'annotation demeurent les mêmes que pour les autres éléments. En effet, les amplitudes sont toujours représentées par différentes couleurs et les trois phases de réalisation sont également distinguées.

Nous avons également proposé un traitement particulier pour le regard. Cependant, l'annotation étant longue et laborieuse, nous ne l'avons pas annoté, et n'avons donc pas testé cette méthodologie. Nous prenons donc le parti de ne pas présenter cette méthodologie ici mais dans les perspectives (chapitre 5).

La tête et les coudes n'ont pas été traités davantage mais sont actuellement décrits par (Filhol, 2008).

3.1.2.4. A chaque attribut sa couleur

Le logiciel Anvil permet de donner une couleur à chaque attribut. Nous avons utilisé peu de couleurs, à des fins de lisibilité. Ces couleurs nous permettent de distinguer des phénomènes bien distincts à savoir par exemple les hausses des baisses. Après une première phase de tests s'est révélé l'intérêt de distinguer également les contacts entre deux éléments, des rapprochements sans contact (par exemple distinguer les yeux plissés, des yeux fermés ou idem pour la bouche) puis les clignements, des fermeture des yeux, afin de les mettre en valeur pour l'analyse. Nous avons également regroupé les phénomènes souvent présents ensemble (comme par exemple une grande ouverture des yeux et une hausse de sourcil) et leur avons attribué la même couleur. Enfin, certains éléments fonctionnent ensemble en ayant des mouvements inverses comme par exemple les yeux ou les épaules. Dans le cas d'une grande ouverture des yeux, la paupière supérieure monte tandis que la paupière inférieure baisse. Il en est de même pour les épaules : lors d'une rotation du torse vers la droite par exemple, l'épaule gauche va avancer tandis que l'épaule droite va reculer. Tous ces mouvements fonctionnent ensemble et nous leur attribuons la même couleur. Ainsi, nous utilisons sept couleurs :

- Bleu pour les hausses des sourcils et de la paupière supérieure, pour la baisse de la paupière inférieure (pour les yeux grand ouverts) et pour les joues gonflées.
- Jaune pour le froncement des sourcils, la baisse de la paupière supérieure, la hausse de la paupière inférieure (yeux plissés), et les joues creuses.
- Vert pour les clignements des yeux.
- Rouge pour les contacts des yeux et de la bouche.
- Rose dont l'utilisation est très ponctuelle : uniquement dans le cas où, après une fermeture des yeux, le signeur ouvre très peu la paupière supérieure. Dans ce cas précis on ne peut, en effet, ni utiliser le jaune qui qualifie une baisse, ni utiliser le bleu qui qualifie une paupière levée plus que la normale.
- Violet et bleu clair pour distinguer les mouvements des épaules et du torse.

Tous ces éléments interagissent entre eux et proposent une description détaillée du mouvement comme nous l'illustrent les Figure 32, Figure 33 et Figure 34 extraites de notre annotation.

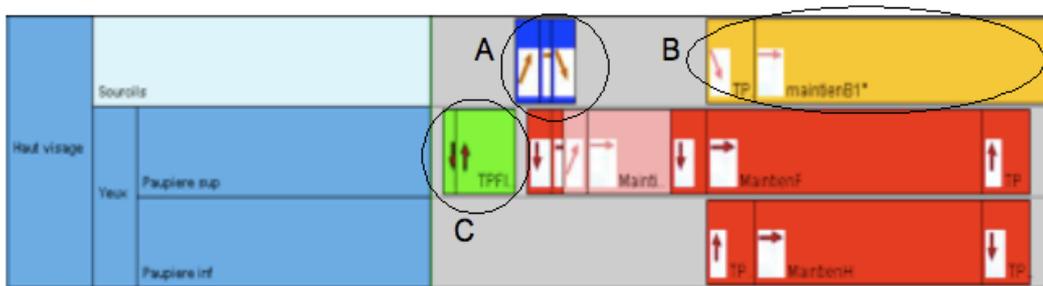


Figure 32 Extrait de notre transcription du corpus LS-COLIN

Le premier bloc de la première ligne (bloc A) décrit une hausse de sourcils “ standard ” (flèches marron : ” standard ”) avec ces trois phases de réalisation. A droite (bloc B), un léger froncement de sourcils (flèches claires : petite amplitude). De même, sur la seconde ligne (paupière sup.), le premier bloc (C) montre un clignement des yeux, le second bloc montre une fermeture des yeux (flèche de contact, voir Tableau 4, page 88) maintenue. Enfin, au lieu d’avoir un retour à la posture initiale, il va les réouvrir légèrement (bloc plus clair⁶¹). Ce qui distingue un retour à la posture initiale d’une légère réouverture est la couleur de la flèche (niveau d’amplitude). Si l’annotation commence avec des flèches “ standard ” (marron) et que le retour à la posture initiale se fait à l’aide d’une flèche claire, alors il ne s’agit pas d’un retour à une posture initiale mais une nouvelle posture comprise entre la posture adoptée et la posture initiale.

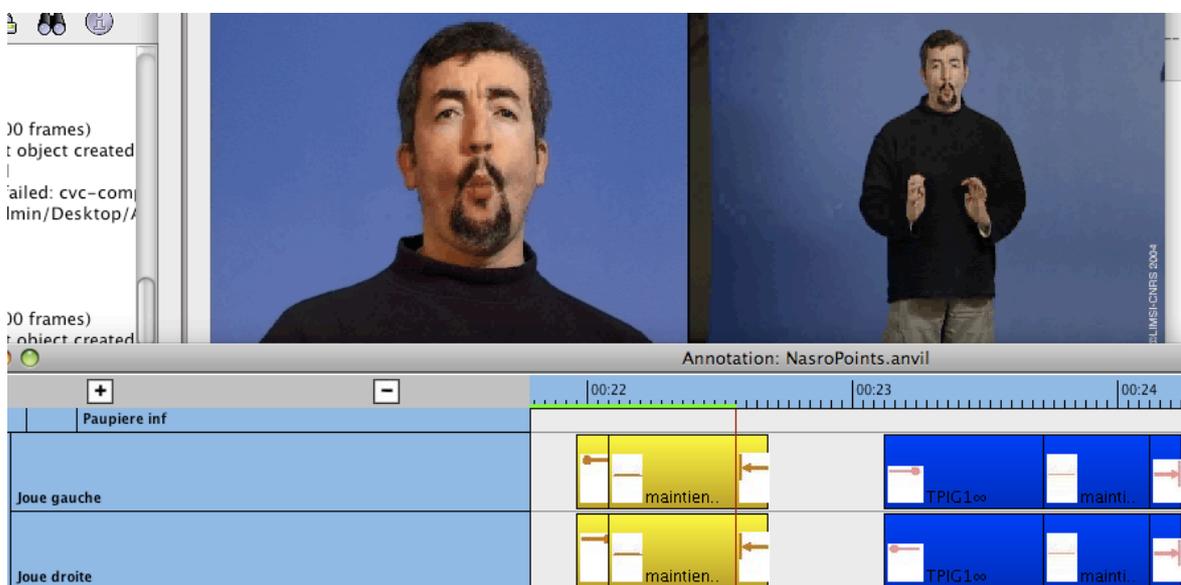


Figure 33 Illustration de l’annotation des joues, extraite de notre transcription du corpus LS-COLIN.

⁶¹ On distingue donc la couleur des symboles (qui informe sur l’intensité des mouvements) de la couleur qui entoure les symboles (les fonds, qui correspondent aux labels et permettent la mise en évidence ou la vérification de corrélation de certains phénomènes).

Le premier bloc de la Figure 33 montre des joues creuses standards (symboles sur fond clair) tandis que le second montre des joues faiblement gonflées (symboles clairs sur fond foncé).

La Figure 34 ci-dessous illustre l'annotation des épaules. Ici, comme les couleurs sont identiques sur les deux pistes (une pour chaque épaule), cela indique un mouvement du haut du corps tout entier, sur la gauche puis sur la droite.

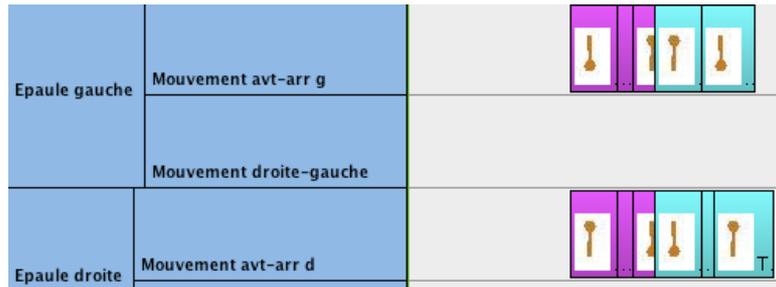


Figure 34 Illustration de l'annotation des épaules, extraite de notre transcription du corpus LS-COLIN.

3.1.3. Bilan

La méthodologie proposée repose sur une décomposition des éléments à annoter et des mouvements complexes. Ces décompositions permettent d'utiliser peu de symboles pour décrire tous les mouvements avec économie. Ainsi, tout en respectant la notion d'amplitude, il est possible de distinguer les phases de réalisation d'un GNM. La dynamique est placée au cœur de la méthodologie, ce qui permet de disposer de descriptions formelles qui renseignent sur la structure interne des GNM. Au final, le caractère visuel de la formalisation favorise l'analyse qualitative et permet de mettre rapidement en évidence les phénomènes récurrents. Cependant, la méthodologie ne répond pas encore à la totalité de nos objectifs. En effet, comme nous l'avons vu, nous souhaitons que la formalisation puisse être utilisée rapidement pour animer un visage de synthèse notamment dans le but de valider notre méthodologie. De plus, vu le travail long et laborieux demandé par l'annotation, il serait pertinent d'envisager une automatisation. Or, ces objectifs nécessitent de disposer de définitions numériques de chacun des symboles. Ainsi, nous souhaitons que chaque symbole soit défini par rapport au mouvement standard relatif à chaque signeur. Par exemple, une hausse amplifiée devra correspondre à un pourcentage d'élévation par rapport à une élévation standard pour la personne. Pour calculer le pourcentage d'élévation, il semble nécessaire de disposer de données numériques de chacune des amplitudes de mouvement. Ainsi, nous avons mis en place une nouvelle méthodologie d'annotation (explicité dans la section suivante), directement sur la vidéo,

afin d'obtenir les coordonnées exactes de chacun des éléments souhaités. Ces coordonnées devront permettre, en premier lieu, de valider l'annotation qualitative (ajustement du nombre de symboles, etc.). En second lieu, elles permettront de définir formellement chacun de ces symboles servant ainsi d'une part à leur génération automatique, à partir des données métriques et d'autre part à rejouer immédiatement l'annotation sur un visage de synthèse afin d'évaluer ses performances (mobilité du visage, etc.). Enfin

3.2. Description quantitative et validation

Pour cette partie de l'étude, nous nous sommes centrés sur les mouvements des sourcils uniquement et ce en raison de la quantité de travail demandée pour réaliser ce type d'annotation.

Dans une première partie nous présentons la méthodologie adoptée. Dans une seconde partie, nous détaillons le traitement des données numériques résultantes de l'annotation. Enfin, nous proposons trois utilisations de ces données permettant la validation et l'analyse structurelle et linguistique de l'annotation.

3.2.1. Annotation sur les images de la vidéo

L'objectif visé ici est d'obtenir des données numériques concernant les mouvements des sourcils. Pour ce faire nous avons utilisé une option d'Anvil qui permet de positionner des points directement sur la vidéo et d'en récupérer les coordonnées.

La première étape était de déterminer les emplacements les plus pertinents pour positionner ces points sur les sourcils. Nous nous sommes alors appuyés sur FACS (présenté dans le chapitre précédent) (Ekman et Friesen, 1978).

Pour les sourcils, Ekman et Friesen distinguent quatre muscles, permettant trois actions : hausse de l'intérieur⁶² du sourcil (Action Unitaire 1, ou AU 1), hausse de l'extérieur⁶³ du sourcil (AU 2) et rapprochement des sourcils (AU 4). Le muscle frontal (Figure 35) permet la hausse de l'intérieur et de l'extérieur des sourcils.

⁶² Par intérieur du sourcil nous entendons le point du sourcil le plus proche de l'arête du nez.

⁶³ Par extérieur du sourcil nous entendons le point du sourcil le plus proche de l'oreille.

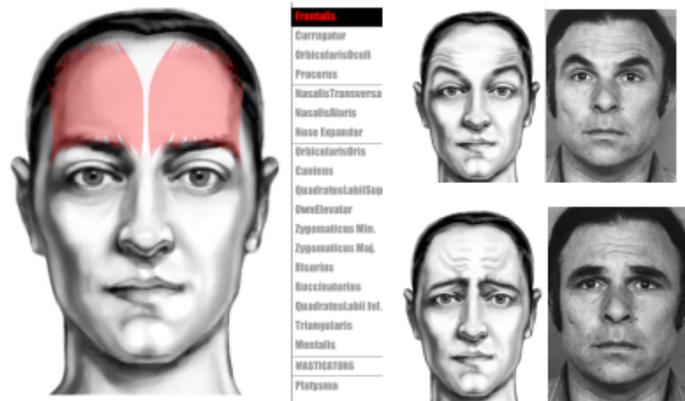


Figure 35 Muscle frontal et ses AUs associées (hausse de l'intérieur (AU 1) et de l'extérieur des sourcils (AU 2)). Images extraites d'Artnatomy⁶⁴ (Contreras Flores, 2005) et du manuel de FACS⁶⁵.

Les muscles sourcilier et orbiculaire des paupières, ainsi que le muscle pyramidal du nez permettent un mouvement latéral des sourcils qui se caractérise par une variation de la distance entre les sourcils par exemple lorsqu'ils sont froncés (Figure 36).

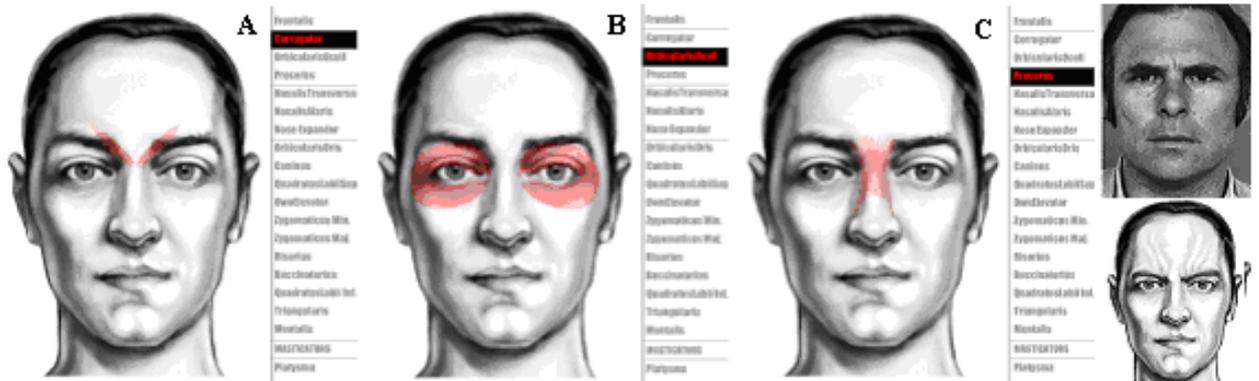


Figure 36 Muscles sourcilier (image A), orbiculaire des paupières (image B) et pyramidale (image C) et leurs AUs associées : rapprochement des sourcils (AU 4). Images extraites d'Artnatomy et du manuel de FACS.

Ces AUs peuvent se combiner entre elles. Les images ci-dessous (Figure 37) nous montrent les combinaisons des AUs 1 et 4 (hausse de l'intérieur des sourcils et rapprochement des sourcils), AUs 1 et 2 (hausse de l'intérieur et l'extérieur des sourcils) et AUs 1, 2 et 4 (rapprochement des sourcils et hausse de l'intérieur et l'extérieur des sourcils).

⁶⁴ www.artnatomia.net

⁶⁵ <http://www.face-and-emotion.com/dataface/facs/manual/TitlePage.html>

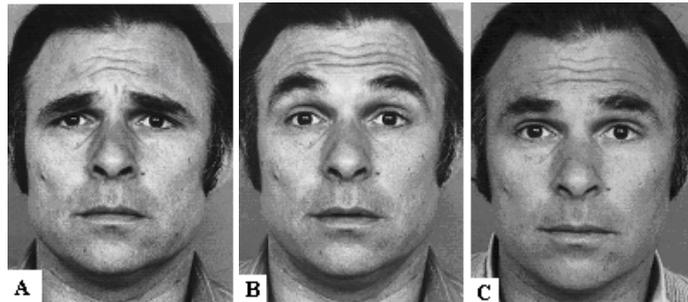


Figure 37 Combinaison d'Aus : Image A : AU 1 + AU 4 ; Image B : AU 1 + AU 2 ; Image C : AU 1 + AU 2 + AU 4

Selon les combinaisons, la longueur des sourcils augmente ou diminue. De plus, il nous semble que lorsque l'extérieur du sourcil monte (AU 2), c'est la hausse du milieu du sourcil qui est perçue et non l'extrémité extérieure.

Le système FACS considère 2 points mobiles pour les sourcils (l'intérieur et l'extérieur) pouvant se mouvoir sur un axe horizontal et latéral. Nous avons retenu ces deux points pour l'annotation de la vidéo. De plus, comme nous l'avons évoqué, le milieu du sourcil semble entraîné par les mouvements des extrémités des sourcils et l'amplitude de son mouvement est parfois plus grande. Nous avons donc considéré également le milieu du sourcil et ajouté un point à cet endroit⁶⁶.

Afin de limiter les imprécisions de l'annotation dues à l'épaisseur des sourcils et au manque de précision de la source des mesures⁶⁷, nous avons doublé les points des extrémités. De même, nous avons triplé le point du milieu du sourcil, qui est plus difficile à positionner avec précision. Enfin, pour déterminer les mouvements des sourcils de manière indépendante de l'orientation de la tête nous avons ajouté des points sur un repère fixe : les deux extrémités de chacun des yeux. Ainsi, ce sont 18 points que nous positionnons sur chaque image de la vidéo, à raison de 25 images par secondes. La Figure 38 récapitule les 18 points positionnés sur chaque image. La Figure 39 montre le résultat d'une annotation des 18 points sur une image extraite de la vidéo.

⁶⁶ Les infographistes positionnent jusqu'à 4 points pour le milieu d'un sourcil.

⁶⁷ Pour obtenir une mesure de la précision de nos données, il aurait fallu soit que les données soient acquises par un système de capture de mouvement, soit qu'elles soient accompagnées d'éléments métriques sur le visage du signeur, information dont nous ne disposons pas. C'est pourquoi nous avons du trouver un autre moyen de limiter les imprécisions.

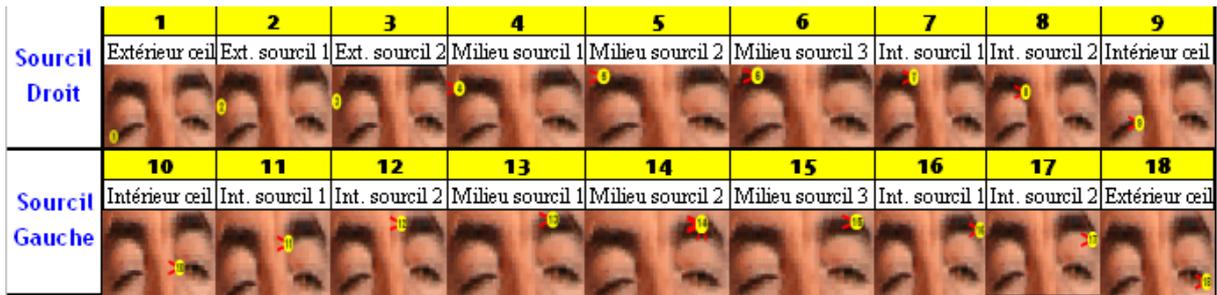


Figure 38 Emplacement des 18 points

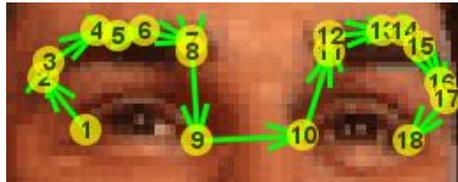


Figure 39 Image extrait de la vidéo avec les 18 points.

Une fois la vidéo entièrement annotée, les coordonnées x,y de chaque point sont exportées dans un environnement de calcul scientifique afin d'établir un certain nombre de données statistiques et numériques explicitées dans la section suivante.

3.2.2. Calcul sur les coordonnées des points

L'objectif des calculs que nous proposons ci-dessous est, à partir des coordonnées des points, de disposer d'un chiffre par partie des sourcils (soit six chiffres en tout : à savoir un pour chaque extrémité des sourcils et un pour le milieu de chaque sourcil) indiquant avec précision la position de la partie du sourcil concernée, et ce pour chaque image.

Pour le traitement de nos données nous avons utilisé le logiciel Scilab⁶⁸, logiciel libre de calcul scientifique, qui nous a permis, en créant un script simple, d'automatiser l'ensemble des calculs nécessaires. L'algorithme prend les données brutes en entrée (les coordonnées des 18 points de chaque image, Figure 40) et effectue une série de calculs détaillés ci-dessous.

⁶⁸ <http://www.scilab.org/>

```

<point time="0.04" x="158" y="120" />
<point time="0.04" x="153" y="105" />
<point time="0.04" x="153" y="103" />
<point time="0.04" x="159" y="96" />
<point time="0.04" x="164" y="96" />
<point time="0.04" x="170" y="96" />
<point time="0.04" x="177" y="100" />
<point time="0.04" x="177" y="105" />
<point time="0.04" x="175" y="116" />
<point time="0.04" x="201" y="115" />
<point time="0.04" x="199" y="101" />
<point time="0.04" x="198" y="96" />
<point time="0.04" x="206" y="94" />
<point time="0.04" x="211" y="94" />
<point time="0.04" x="216" y="95" />
<point time="0.04" x="220" y="101" />
<point time="0.04" x="219" y="102" />
<point time="0.04" x="218" y="116" />

```

Figure 40 Données brutes en entrée : coordonnées des 18 points d'une image

Le premier calcul vise à déterminer les moyennes des 2 points des extrémités et des 3 points du milieu des sourcils dans chaque image. Ainsi, on passe de 7 points à 3 points pour chaque sourcil (Figure 41).

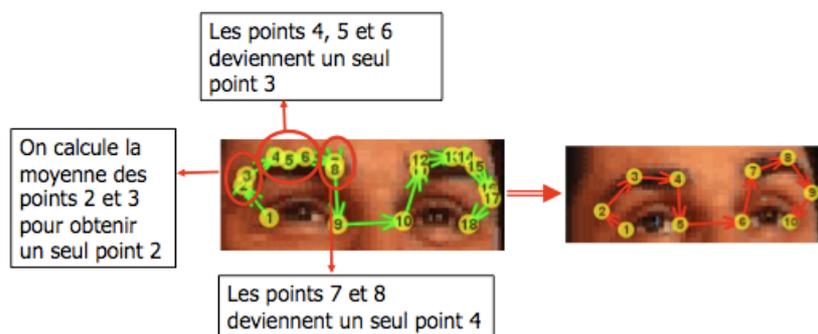


Figure 41 Calcul des moyennes

Les nouvelles coordonnées sont utilisées pour calculer la distance (D) qui sépare chacun des 3 points des sourcils des points des extrémités des yeux. Par exemple, pour calculer la distance entre le point 1(x1, y1) et le point moyen 2 (x2, y2)) (voir Figure 42) ou $D = \text{Racine}((x1-x2)^2 + (y1-y2)^2)$. Pour le point du milieu (point 3), la distance est calculée par rapport au milieu des points des extrémités des yeux (points 1 et 5).

Nous obtenons ainsi des valeurs indépendantes de l'orientation de la tête. Ces valeurs nous permettent de connaître l'amplitude minimale et maximale de chaque signe. Cependant, ces données ne permettent pas de savoir si le sourcil est en hausse ou en baisse. Nous utilisons donc la distance pour de premières analyses mais poursuivons le traitement.

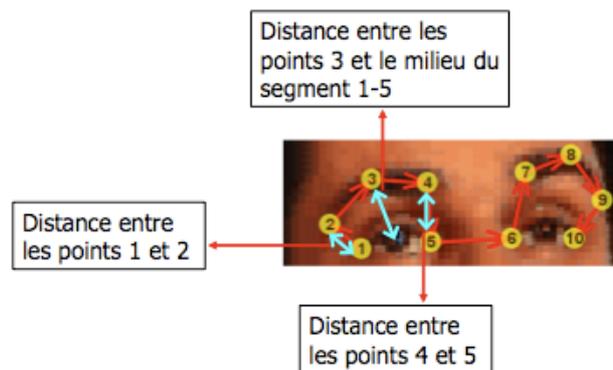


Figure 42 Calcul de la distance entre les points des sourcils et les points des yeux

A partir de la distance (D) nous calculons la variation (V) de la position des points en faisant la différence entre la distance à l'image n et la distance à l'image n-1 : $V(n) = D(n) - D(n-1)$. Cette variation peut donc être positive (hausse) ou négative (baisse) et indique si le sourcil est monté ou descendu par rapport à l'image précédente. La variation nous permet d'observer la structure interne des mouvements des sourcils. Par exemple, elle nous permet d'étudier l'amplitude maximale réalisable en une seule image (soit 40 ms) et de définir ainsi le mouvement standard pour une image. On obtient donc des informations du type : une hausse standard est de X millimètres chez ce signeur. Une hausse supérieure à X est donc considérée comme rapide, etc.

Enfin, la variation (V) nous permet de calculer la position (P) de chaque élément et pour chaque image en calculant la somme de la variation de l'image n et de la position de l'image n-1 : $P(n) = V(n) + P(n-1)$. Il s'agit ici d'obtenir un chiffre indiquant l'amplitude de la hausse ou de la baisse de chaque partie du sourcil indépendamment des images précédentes.

Nous obtenons en sortie les positions des six points des sourcils : intérieur, milieu et extérieur de chaque sourcil et ce pour chaque image comme l'illustre la Figure 43.

```
<point time="0.04" 0.719679 />  
<point time="0.04" 0.903642 />  
<point time="0.04" -0.494185 />  
<point time="0.04" 1.066595 />  
<point time="0.04" 0.005654 />  
<point time="0.04" 0.477885 />
```

Figure 43 Données en sortie : positions des 6 points des sourcils d'une image

Les données obtenues sont ensuite utilisées afin d'évaluer l'annotation et de faciliter les analyses.

3.2.3. Évaluation intermédiaire

Les valeurs calculées sont tout d'abord utilisées pour générer automatiquement l'animation des sourcils sur un visage synthétique à partir des données issues de l'annotation. L'animation a été générée à l'aide du logiciel Xface⁶⁹ (Balci, 2004). Xface est une tête parlante en 3D conçue pour la production vocale, et non pour la production de LSF. Son objectif n'est pas de produire des GNM mais de la langue vocale sur un visage dynamique.

Cependant, la génération automatique sur cette tête parlante nous permet de procéder à une première évaluation qualitative de l'annotation en vérifiant si l'ensemble des phénomènes observés sont présents. En effet, en jouant simultanément le corpus vidéo et l'animation du visage synthétique, il est possible d'ajuster la méthodologie au fur et à mesure (rajout d'un point par exemple) pour que l'annotation finale soit la plus fidèle possible.

Les Figure 44 et Figure 45 illustrent deux configurations des sourcils et des yeux. L'image de gauche est extraite du corpus vidéo, tandis que celle de droite est extraite de Xface.

⁶⁹ <http://xface.itc.it/>

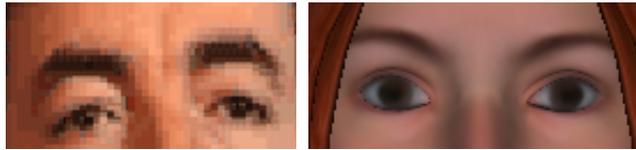


Figure 44 Position neutre



Figure 45 Diminution de la distance entre les sourcils et baisse de la paupière supérieure.

Jouer simultanément les deux vidéos nous permet également d'évaluer la qualité du visage synthétique, d'identifier ses défauts, afin de prévoir des conditions minimales de réalisme pour les visages des personnages de synthèse utilisés dans la production automatique de LSF. Par exemple, la génération sur le logiciel Xface a permis d'observer un manque d'amplitude au niveau des sourcils (il est possible de l'augmenter mais le rendu est irréaliste car il n'est pas conçu pour une telle amplitude) ainsi que l'absence de ride. La combinaison de ces deux défauts induit une perception dégradée du mouvement des sourcils. D'autres mouvements ne sont pas possibles sur le visage synthétique de XFace (par exemple l'abaissement de la paupière du bas, indispensable pour ouvrir grand les yeux) ce qui, lors de l'évaluation, engendre des problèmes de perception. Il est ainsi possible d'établir une liste des éléments essentiels que doivent posséder les visages synthétiques et les systèmes de génération automatique de LSF afin de prendre en compte les GNM et augmenter la qualité visuelle et linguistique des messages générés.

Cette première utilisation des données a permis une évaluation rapide de la méthodologie et son ajustement. Les données peuvent alors être utilisées pour l'analyse des GNM.

3.2.4. Analyse structurelle des GNM

Les données numériques sont exploitées pour produire des courbes afin d'analyser la structure des mouvements. Le graphique ci-dessous (Figure 46) nous informe par exemple⁷⁰ sur l'amplitude du mouvement de l'intérieur des sourcils et permet de distinguer différents types de hausses.

⁷⁰ Nous présentons ici une analyse dans le but d'illustrer cette utilisation des données. Nous avons toutefois procédé à plus d'analyses dont les résultats sont présentés dans le chapitre 4.

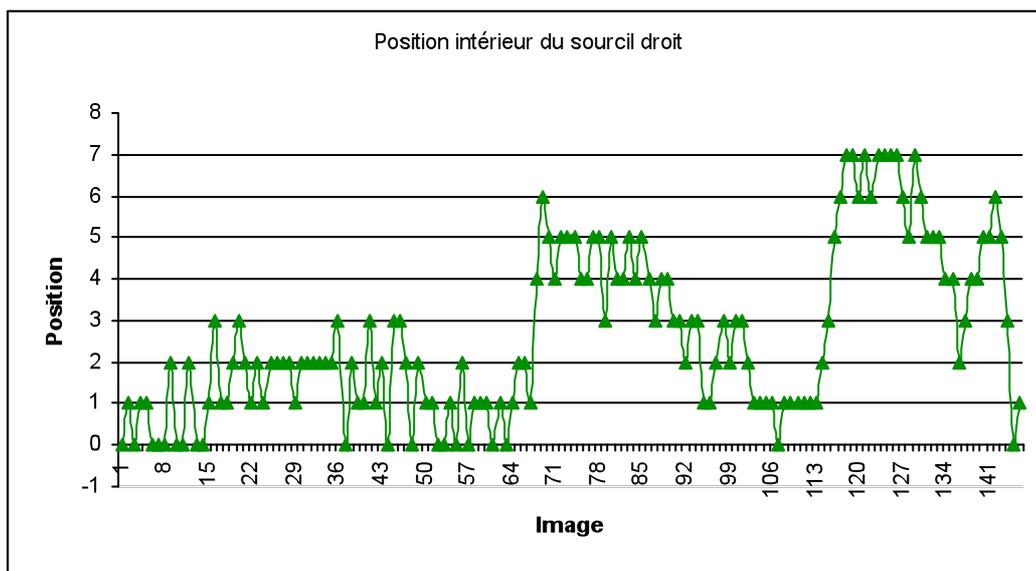


Figure 46 Positions de l'intérieur du sourcil droit.

Les différents niveaux d'amplitudes sont déterminés en observant les variations entre deux positions successives dans la courbe. Nous avons noté trois amplitudes de hausse pour l'intérieur du sourcil : une petite (de 1 unité chez ce signeur), une moyenne (2 unités) et une forte (3 unités). La distance maximale entre deux positions est donc de 3 unités. Il ne s'agit donc pas de considérer que la hausse maximale est une hausse de 3 unités, mais qu'adopter une position plus élevée (par exemple 7 unités sur ce graphique), se fait par étapes : plusieurs hausses successives de différents degrés avec des tenues d'au moins une image pour chacune d'entre elles.

Ces 3 hausses peuvent être définies de façon précise par rapport à la hausse de niveau 1 (petite hausse) : une hausse de niveau 2 est 2 fois plus haute qu'une hausse de niveau 1 tandis qu'une hausse de niveau 3 est 3 fois plus haute qu'une hausse de niveau 1. Si la valeur numérique des amplitudes des hausses varie d'un signeur à l'autre (jusqu'à 1,5 cm), leur nombre et leur proportion les unes par rapport aux autres restent identiques. Le constat est le même concernant les baisses.

Les courbes peuvent donc nous permettre d'analyser finement la structure du mouvement des GNM.

3.2.5. Bilan

Notre objectif était de disposer de données numériques des positions des sourcils et ce pour chaque image.

Nous avons donc proposé une nouvelle méthodologie, directement sur la vidéo, qui permet de fournir la position exacte de l'extérieur, le milieu et l'intérieur de chaque sourcil. Cette nouvelle méthodologie d'annotation permet de disposer de données fines sur la structure des mouvements. De plus, elle permet de valider les choix et hypothèses de notre première annotation. En effet, nous avons utilisé trois flèches représentant les amplitudes de réalisation. Le choix du nombre reposait sur des tests effectués au cours de l'élaboration de la méthodologie et sont aujourd'hui validés par les données numériques de cette seconde annotation. Enfin, chacun des symboles utilisés va ainsi être défini de façon numérique ce qui permettra, à terme, une génération automatique de notre annotation symbolique à partir des données numériques obtenues par l'annotation des points sur la vidéo.

4. Conclusion

A partir des contraintes et des objectifs que nous nous sommes fixés, nous avons proposé une nouvelle structure d'annotation basée sur la description de la trajectoire. Grâce à la décomposition des mouvements et des éléments à annoter, l'annotation utilise un nombre restreint de symboles ce qui la rend plus accessible. De plus, celle-ci s'est avérée efficace pour la description de tout type de mouvement pour chaque élément à annoter⁷¹.

Nous avons ensuite proposé un autre type d'annotation qui consiste à positionner des points directement sur la vidéo afin d'en récupérer les coordonnées. Celles-ci ont été utilisées afin de calculer la position exacte des trois points de chaque sourcil, image par image. Cette annotation permet de disposer de données numériques indispensables à la validation des symboles de la formalisation et à leur définition numérique.

Une première annotation a révélé la nécessité de procéder à quelques ajustements de la structure afin de s'adapter aux phénomènes découverts lors des premières analyses. En effet, d'après l'observation des premiers résultats, certaines dissociations supplémentaires semblaient être nécessaires. De plus, le fichier de spécification initial ne permettait pas d'afficher des courbes ou de positionner les points, nous avons donc du apporter quelques améliorations à ce fichier. Une fois les améliorations apportées, nous avons procédé aux analyses. Les premiers résultats sont présentés dans le chapitre suivant

⁷¹ A noter cependant que nous avons annoté uniquement les images où le signeur faisait face à la caméra : à chaque mouvement de rotation de la tête l'annotation a été suspendue pour les deux annotations.

Chapitre 4 : Résultats des analyses

Ce chapitre présente les résultats de notre étude. Nous proposons tout d'abord d'apporter quelques précisions quant à l'annotation et une synthèse du traitement de chacun des éléments non manuels, puis de présenter les données au cœur des analyses. La seconde partie de ce chapitre détaille les résultats de nos observations et propose en particulier une typologie et une représentation formelle des mouvements des sourcils et des clignements des yeux.

1. L'annotation

Les résultats présentés dans ce chapitre proviennent directement de l'analyse des données de l'annotation. La première section apporte quelques précisions quant à l'annotation. Cette section nous permet d'appréhender la difficulté et la lenteur de la tâche ce qui explique que l'étude se soit focalisé ensuite sur certains éléments non manuels en particulier. La seconde section propose une synthèse des éléments traités lors de l'annotation et de ceux sur lesquels porteront les analyses. Enfin, dans une troisième section, nous présentons le type de données sur lesquelles se sont basées les analyses.

1.1. L'annotation en quelques chiffres

Les énoncés de quatre locuteurs du corpus LS-COLIN ont été annotés (trois hommes et une femme d'origines et âges divers⁷²). Chaque intervention dure environ deux minutes, ce qui représente un total d'environ 13 000 images (25 images par seconde). L'annotation s'est déroulée en deux phases : la première phase consiste en l'annotation à l'aide du système de symboles que nous avons créé et la seconde phase consiste à positionner les points directement sur la vidéo.

L'ensemble de l'annotation a demandé un temps considérable, à savoir par exemple que pour la première phase, chaque piste a été annotée séparément ce qui représente, juste pour les sourcils et les yeux, le visionnage et l'annotation de plus de 120 000 images à raison d'au moins trois secondes passées sur chaque image. Pour la dernière phase, il a fallu environ 15h pour positionner les points sur une seule minute d'une seule vidéo (1' -> 15h).

⁷² Le choix des locuteurs est empirique, les styles de chacun sont assez différents.

La précision et l'attention demandée par l'annotation expliquent en grande partie la lenteur de la tâche. Nous estimons que ce temps était nécessaire, afin d'obtenir des données fiables pour qu'à terme elles puissent servir à calibrer des logiciels de reconnaissance du visage, et donc permettre l'annotation automatique (ou partiellement automatique) des GNM. De plus, ces annotations nous ont permis de valider notre méthodologie en ajustant le nombre de points et de symboles ou en affinant les pistes. Enfin, notre objectif était de connaître la structure des GNM. Or, cette annotation permet de fournir des renseignements suffisamment fins et précis pour l'analyse souhaitée.

Le corpus a été traduit par un interprète, qui a inséré sa traduction directement dans le logiciel, indépendamment de notre annotation. Ainsi, lorsque les frontières de ses traductions (début et fin d'un signe) concordent avec celles de nos segments, il ne s'agit pas d'une influence de l'un sur l'autre, mais réellement de frontières communes. A noter qu'il n'existe pas de traduction "standard". Deux traductions peuvent donc se distinguer d'un interprète à l'autre. Pour notre étude, l'interprète choisi possède l'avantage de connaître la tâche d'annotation et le logiciel Anvil.

1.2. Synthèse des éléments traités

La méthodologie a été conçue pour traiter les yeux, les sourcils, les joues, la bouche, la langue (uniquement lorsqu'elle déforme les joues⁷³), la tête, les épaules et le buste. Tous ces éléments ont été annotés partiellement pour affiner la méthodologie. Cependant, la difficulté et la lenteur de la tâche nous ont contraint à nous concentrer sur certains éléments uniquement, à savoir les sourcils et les clignements des yeux, pour lesquels nous avons annoté l'ensemble des quatre productions.

⁷³ La langue est dans ce cas-là incluse dans la description des joues. Lorsque la langue apparaît, donc ne déforme pas les joues, nous ne l'avons pas annotée car il ne s'agit alors pas d'un mouvement visible que l'on peut décrire mais d'une posture dont on ne voit pas les transitions. Ainsi, ce type de description sera ajouté ultérieurement à notre méthodologie.

Le choix de ces éléments a essentiellement été motivé par la première annotation. Celle-ci a en effet révélé que certains éléments, comme par exemple la langue et les joues, ne présentaient pas suffisamment de mouvements dans notre corpus pour être analysés. Les mouvements de la bouche étaient très présents. Cependant, considérant l'influence de la LV sur la production de certains signeurs (par exemple certaines labialisations), nous avons estimé qu'une étude sur les mouvements de la bouche nécessitait l'annotation de davantage de productions, afin de distinguer les phénomènes buccaux spécifiques aux LS des phénomènes dues à l'influence des LV. Par ailleurs, l'étude des mouvements des sourcils et des yeux a très vite révélé de nombreux phénomènes à étudier, qu'il s'agisse de structures récurrentes, comme la présence simultanée de deux éléments, ou de phénomènes surprenants tels que des passages de l'énoncé où les sourcils n'avaient pas le mouvement attendu (par exemple une baisse des sourcils là où une hausse était attendue). C'est pourquoi nous avons orienté notre étude sur l'analyse de ces deux éléments. A noter que si les mouvements des paupières ont bien été intégralement annotés, seuls les clignements des yeux ont été analysés. Les mouvements des paupières autres que les clignements feront l'objet d'une étude ultérieure.

Ainsi, tous les éléments non manuels listés ci-dessus ont été annotés au moins dans une production mais seuls les mouvements des sourcils et des yeux ont été annotés complètement, avec les deux types d'annotation (qualitative et quantitative) et pour les quatre signeurs. Nous proposons dans le tableau ci-dessous une synthèse des traitements effectués pour chacun des éléments non manuels.

Éléments	Annotation				Analyse	Résultat
	Signeur 1	Signeur 2	Signeur 3	Signeur 4		
Tête	X	/	/	/		
Sourcils	X	X	X	X	/	/
Yeux	X	X	X	X	/	/
Joues	/	/	X			
Bouche	/					
Langue	/		X	/		
Épaules	X		/			

Tableau 6 Synthèse des traitements effectués sur chaque élément

Dans ce tableau nous avons distingué deux types de traitements : un traitement partiel (/) et un traitement complet (X). Si les mouvements des sourcils et des yeux ont été complètement annotés, nous considérons que l'analyse menée à partir de ces annotations peut encore fournir davantage de résultats que ceux que nous présentons dans ce chapitre, d'où le symbole (/).

1.3. Les données utilisées pour les analyses

Les analyses ont porté sur deux types de données : l'observation de l'annotation (telle que visible dans le logiciel) et les données brutes exportées sous tableur.

L'observation de l'annotation permet un parallèle entre la formalisation symbolique et des courbes basées sur les coordonnées des points positionnés sur la vidéo. La Figure 47 illustre le résultat de l'annotation sur quelques pistes. Comme nous pouvons le voir, les courbes sont réparties sur neuf pistes:

- Trois pistes pour les trois points des sourcils (une piste par point : intérieur, extérieur et milieu) : les pistes en questions contiennent chacune deux courbes à savoir une pour chaque sourcil.
- Trois pistes permettent d'observer la différence entre les deux sourcils : la première courbe représente la différence entre l'extrémité extérieure du sourcil gauche et celle du sourcil droit, la seconde représente la différence entre le point du milieu de chaque sourcil puis la dernière pour la différence du point de l'intérieur de chaque sourcil.
- Trois pistes proposent la moyenne entre les deux sourcils pour chacun des trois points.

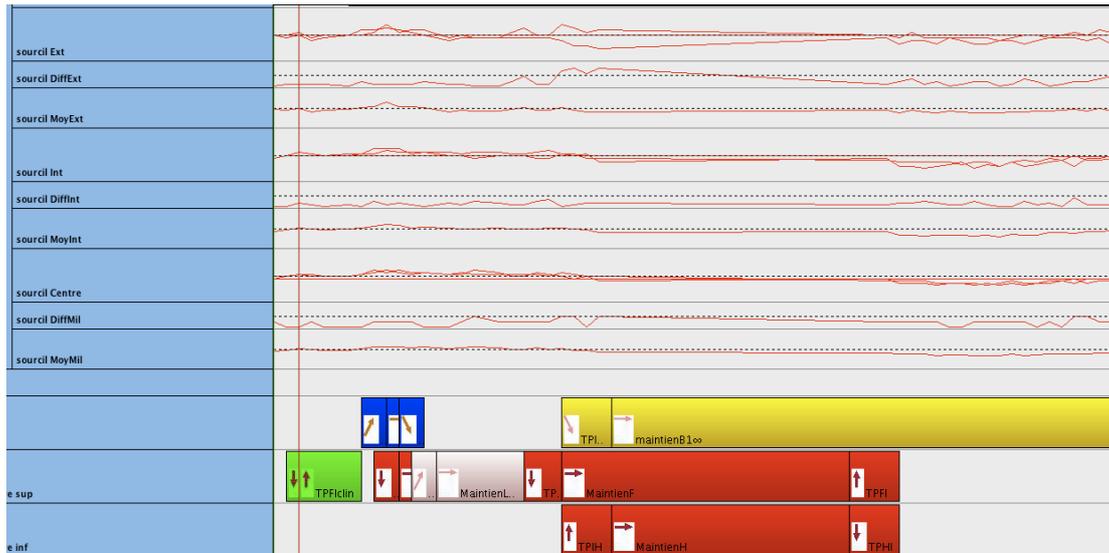


Figure 47 Extrait de l'annotation

Nous avons choisi de visualiser ces mesures afin d'étudier la relation entre les deux sourcils, le lien entre le milieu des sourcils et les extrémités (par exemple entre l'extrémité intérieure et extérieure, laquelle entraîne le milieu du sourcil) mais également afin d'étudier le rôle des différents points des sourcils (à savoir par exemple la signification est-elle la même lorsqu'un signeur lève un sourcil que s'il lève les deux ? S'il lève les extrémités extérieures et pas les intérieures ?). A noter que lorsque la main était devant le visage ou que celui-ci était trop tourné pour voir les yeux, aucun point n'a été positionné jusqu'à ce que les conditions redeviennent optimales. C'est pourquoi, dans la section suivante, des illustrations peuvent parfois ne pas montrer les courbes ou montrent une courbe sans évolution. Dans le premier cas, cela s'explique soit parce que les courbes ne sont pas pertinentes à ce moment là, soit en raison des limites du logiciel Anvil. Dans le second cas, il s'agit d'une partie d'énoncé où les points n'ont pas été positionnés (mains devant le visage ou visage tourné).

Les données sont exportées sous la forme d'un tableau répertoriant tous les segments de la (les) piste(s) sélectionnée(s), leur durée, ainsi que le temps de début et de fin de chacun d'eux. Le Tableau 7 illustre un export des données issues de la formalisation à l'aide de symboles. Les données issues de l'annotation directement sur la vidéo, à l'aide de points, n'ont pas été utilisées au sein d'un tableur mais ont été traitées par un logiciel de calcul scientifique (comme explicité dans le chapitre précédent).

ID	start	end	duration	Yeux
1	0,64	0,72	0,08	TPIFclin
2	0,72	0,88	0,16	TPFIclin
3	1,68	1,76	0,08	TPIFclin
4	1,76	1,88	0,12	TPFIclin
5	2,76	2,84	0,08	TPIFclin
6	2,84	2,96	0,12	TPFO1-
7	2,96	3,28	0,32	maintienB2-
8	3,28	3,64	0,36	TPBI2-

Tableau 7 Extrait de l'exportation de l'annotation des yeux sous Excel.

Nous avons ainsi exporté les pistes relatives aux sourcils et aux yeux.

1.4. Conclusion

Nous avons proposé une synthèse des traitements effectués sur chacun des éléments non manuels et avons expliqué pourquoi la suite de notre étude s'est centrée sur les sourcils et les clignements des yeux. Les données des annotations qualitatives et quantitatives ont été traitées par un tableur ou par un logiciel de calcul scientifique. Ceci nous a permis d'obtenir des résultats numériques et statistiques.

La section suivante présente les premiers résultats de nos analyses.

2. Présentation des résultats

Nous tenons à rappeler que l'annotation porte uniquement sur quatre signeurs. Les analyses que nous proposons ci-dessous étant basées sur notre annotation, les résultats qui en découlent sont à relativiser et restent à confirmer sur davantage de locuteurs. Cependant, nous ne présentons que les phénomènes constatés plusieurs fois et pour l'ensemble des signeurs. Les cas litigieux (doute, phénomène non présent chez tous les locuteurs, etc.) ont été mis de côté pour une analyse ultérieure.

Dans une première et une seconde partie nous présentons nos résultats relatifs spécifiquement aux sourcils puis aux clignements des yeux. Dans une troisième partie, nous proposons de situer nos résultats dans les connaissances actuelles. Enfin, la dernière partie vise à formaliser nos résultats afin de les rendre exploitables pour générer les GNM.

2.1.Observation des mouvements des sourcils

Au regard du grand nombre de données dont nous disposons, nous nous sommes concentrée dans cette étude sur les mouvements des sourcils les plus marqués (environ 260 mouvements étudiés au total).

La sélection des mouvements des sourcils étudiés est le résultat de trois phases successives d'observation : la première phase a porté sur les données quantitatives fournies par les coordonnées des points et traitées par le logiciel de calcul scientifique afin d'obtenir des courbes illustrant les mouvements de chaque partie des sourcils. A partir de ces courbes, les mouvements les plus marqués ont été isolés (les sommets notamment). Nous avons ensuite comparé ces mouvements isolés à l'annotation qualitative, afin de vérifier si chacun de ces mouvements avait été perçu ou non lors de l'annotation (ceci afin de ne garder que les mouvements les plus marqués et d'exclure les erreurs de positionnement des points). Enfin, chaque mouvement a fait l'objet d'une validation sur le corpus grâce à un nouveau visionnage image par image des séquences correspondantes.

Une fois la sélection des mouvements terminée, nous avons répertorié un maximum d'informations sur chacun de ces mouvements : les données quantitatives nous ont par exemple renseigné sur la durée des mouvements et sur l'amplitude maximale et minimale tandis que les données qualitatives nous ont permis d'obtenir des informations sur le contexte des mouvements sélectionnés. Toutes ces informations ont été regroupées dans un tableau (Tableau 8) indiquant :

- Le temps au moment où le mouvement commence (Deb dans le Tableau 8) ;
- La position de chacune des trois parties du sourcil au début du mouvement (DegD : ici les chiffres indiquent que l'extérieur du sourcil est en position de repos (0), le milieu se situe à 1 millimètre au dessous de la position de repos et l'intérieur du sourcil se situe à 1 millimètre au dessus) ;
- Le temps à la fin du mouvement (Fin) ;
- La position de chacune des trois parties du sourcil à la fin du mouvement (DegF : ici, l'extérieur se situe à 3 millimètres au dessus de sa position de repos, le milieu à 2 millimètres et l'intérieur à 6 millimètres) ;
- Le temps à la fin du maintien (s'il y a eu maintien, Maint)⁷⁴ ;
- Le temps à la fin du mouvement de retour à la posture de départ (Ret). S'il n'y a pas un retour à la posture de départ mais un mouvement conduisant à une nouvelle posture, ce mouvement est noté dans une nouvelle ligne car est considéré comme un autre mouvement (d'où la nécessité d'indiquer la position des parties des sourcils au début du mouvement, on sait ainsi si le mouvement part de la posture de repos ou d'une autre posture) ;
- La position à la fin du mouvement de retour (DegR) ;
- Les durées de la première phase (Durée Phas1), du retour à la posture initiale (Durée Phas2) et de la totalité du mouvement (Durée totale) ;
- Le contexte, à savoir les signes impliqués dans le mouvement qu'ils se situent juste avant, pendant ou juste après le mouvement (Contexte) ;

⁷⁴ A noter que le temps Maint indique la fin du maintien et donc le début du mouvement suivant.

- La position du clignement par rapport au mouvement dans le cas où un clignement se situe à proximité (CCL : ici le clignement commence une image après la fin du mouvement).

Deb	DegD	Fin	DegF	Maint	Ret	DegR	Durée Phas1	Durée totale	Durée Phas2	Contexte	CCL
19,72	0_1_1	19,84	3_2_6		20	1_1_0	0,12	0,28	0,16	Entre "à mon avis" / "quoi": Hausse atteinte pour "quoi"	C li ap

Tableau 8 Recouplement des données quantitatives et qualitatives d'un mouvement des sourcils

L'observation de ces tableaux a rapidement révélé des similitudes quant aux comportements des sourcils. Celles-ci constituent les propriétés générales des mouvements des sourcils que nous présentons dans la section 2.1.1. Dans un second temps, l'analyse des tableaux nous a conduit à regrouper des mouvements présentant des caractéristiques communes et à en proposer une typologie (section 2.1.2).

2.1.1. Structure des mouvements : propriétés générales

Contrairement aux clignements des yeux, les mouvements des sourcils sont plus difficilement identifiables. C'est pourquoi les proportions données en pourcentage dans cette section ne sont qu'indicatives et portent sur les mouvements des sourcils les plus marqués que nous avons relevés, à savoir environ 260. Les propriétés générales concernent la mobilité des extrémités, la dynamique des hausses, la corrélation avec les clignements des yeux et la distance entre les sourcils.

Mobilité des extrémités

L'extrémité intérieure des sourcils est plus mobile que l'extrémité extérieure. Cela concerne :

- 1) l'amplitude maximale⁷⁵ des sourcils constatée chez tous les signeurs (Tableau 9).
Par exemple, l'extrémité extérieure du sourcil du signeur varie, de la posture la plus basse à la posture la plus haute

⁷⁵ A distinguer l'amplitude maximale des sourcils (à savoir la différence entre la posture la plus haute et la posture la plus basse) et l'amplitude maximale d'une hausse (à savoir le moment où, lors d'une hausse, les sourcils atteignent la position la plus haute de la hausse en question).

- 2) la fréquence de leur participation à une hausse : seules 29% des hausses montrent un mouvement significatif de l'extérieure des sourcils. A noter qu'un des signeurs étudiés a une mobilité très particulière et surprenante des sourcils. Si nous ne tenons pas compte de ses mouvements des sourcils, 17% seulement des hausses montrent un mouvement significatif de l'extérieur du sourcil.
- 3) Concernant les baisses de sourcils : l'extérieur et l'intérieur interviennent ensemble mais la mobilité de l'extérieur est moindre.

	S1		S2		S3		S4	
	Extérieur	Intérieur	Extérieur	Intérieur	Extérieur	Intérieur	Extérieur	Intérieur
Amplitude maximale en mm	14	18	11	19	12	17	12	20

Tableau 9 Amplitude maximale de l'extérieur et l'intérieur des sourcils pour les quatre signeurs

Dynamique des hausses

- L'amplitude maximale des hausses est atteinte exactement au début d'un signe dans 65% des hausses étudiées (et commence donc en fin de signe précédent). Dans 20% des cas, la hausse commence et atteint son amplitude maximale au milieu d'un même signe. La Figure 48 illustre un tel cas. A noter que la première courbe (a) de la figure indique le mouvement de l'intérieur du sourcil, la seconde (b), la différence entre deux sourcils (donc non pertinente ici) et la dernière (c), l'évolution des points des milieux de chaque sourcil et une graduation correspond à 1 image, soit 40 ms.

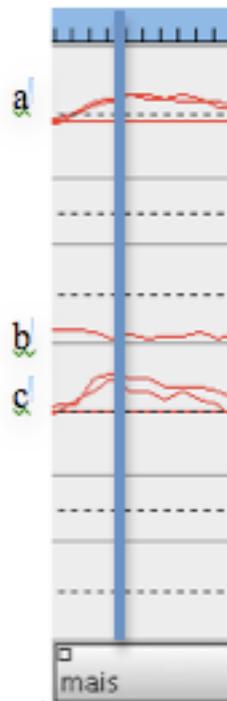


Figure 48 Illustration d'un début de hausse et son amplitude maximale durant un même signe.

- L'amplitude maximale est atteinte, à partir du début du mouvement, en moins de 6 images (250 ms) dans 89% des cas (généralement 3 ou 4 images soit 120 ms ou 160 ms). Dans la Figure 48, l'amplitude maximale est atteinte en moins de 4 images (Trait vertical sur la figure).
- Dans la plupart des hausses, l'amplitude maximale atteinte au début ne sera pas maintenue et diminuera progressivement. Les courbes (c) de la Figure 48 illustrent bien ces propriétés : la hausse est atteinte rapidement puis diminue progressivement.

Corrélation avec le clignement des yeux

- Lorsqu'un mouvement de sourcil se produit pendant un clignement, il commence dans 91% des cas exactement en même temps que la seconde phase du clignement (l'ouverture).
- Lorsqu'un mouvement de sourcil est proche d'un clignement, sans pour autant être en même temps, il est la plupart du temps situé juste après celui-ci (à moins de 3 images soit 120 ms d'écart).

Distance entre les sourcils

- Si l'on considère qu'au repos, la distance entre les sourcils est égale à d et que la distance maximale est égale à $d+X$, alors la distance minimale observée (rapprochement des sourcils) est exactement égale à l'opposé de la distance maximale observée soit $d-X$. Chez nos signeurs, ces distances, exprimées en millimètres, sont égales à : $d-7/d+7$ (pour S1) ; $d-7/d+7$ (pour S2) ; $d-8/d+8$ (pour S3) et $d-5/d+5$ (pour S4) (Image 26, Image 27 et Image 28).

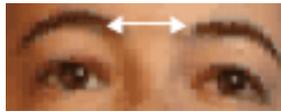


Image 26 Distance entre les sourcils au repos (d)

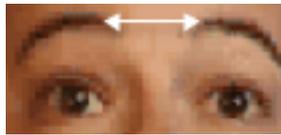


Image 27 Distance entre les sourcils maximale ($d+5$)



Image 28 Distance entre les sourcils minimale ($d-5$)

Une première observation des données nous a permis d'établir la liste des propriétés générales de la structure des mouvements des sourcils. Nous proposons dans la section suivante, une typologie des mouvements des sourcils.

2.1.2. Typologie des mouvements des sourcils

Nous présentons ci-dessous une catégorisation des phénomènes récurrents observés chez les quatre signeurs de notre étude, des plus fréquents au moins fréquents. Nous proposons pour chaque catégorie, si cela est possible, un parallèle avec les mouvements des sourcils observés par Cuxac (2000) (pour les mouvements concessifs, de problématisation et de fausses questions). Nous avons relevé six catégories de mouvements : la segmentation d'énoncés, les concessifs, la mise en évidence du rhème, les transferts personnels, la problématisation et les fausses questions.

2.1.2.1. La segmentation d'énoncés (23,44% des mouvements étudiés)

Il s'agit de brèves hausses des sourcils très prononcées qui marquent une légère pause dans l'énoncé: le signeur marque une pause (de 2 à 4 images soit de 80 à 160 ms) durant laquelle il produit ce mouvement, puis il enchaîne sur la suite de l'énoncé.

Ces mouvements ressemblent aux concessifs (au niveau de leur structure et de leur forme, voir section 2.1.2.2) mais n'apportent aucune restriction à ce qui précède et sont donc à distinguer. Par exemple, [AVANT LE 11 SEPTEMBRE 2001] – pause (de 3 images) – [ETATS UNIS SITUATION] (Figure 49). La hausse se situe durant la pause et se termine pendant le début du signe (ou juste avant selon la durée de la pause). Ces mouvements sont très soudains et très marqués et sont accompagnés d'un clignement des yeux (voir section 2.2.3.1). Contrairement aux mouvements concessifs, les mouvements hausses/baisses (par baisses nous entendons retour à la posture de départ) sont symétriques et il n'y a aucun maintien de la hausse. Dans la Figure 49, ce mouvement est montré par la flèche en diagonale qui se trouve entre [UN] et [ETATS-UNIS] et la flèche à sa droite (on constatera l'absence de maintien représentée par une flèche horizontale). Celles-ci sont foncées et épaisses ce qui signifie une amplitude maximale.

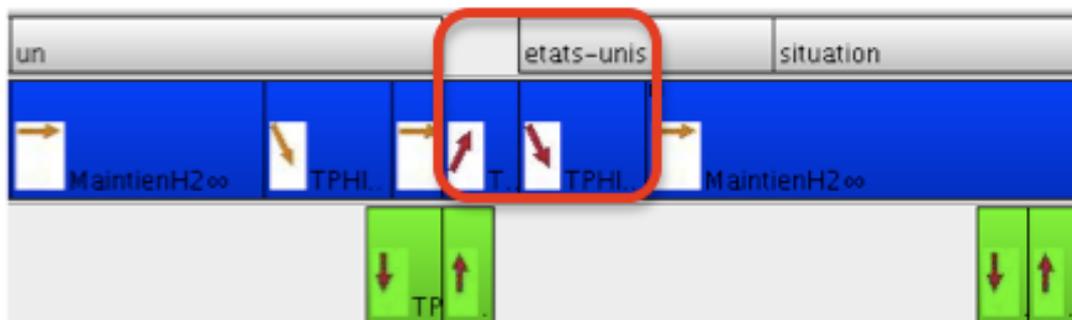


Figure 49 Illustration d'un mouvement de sourcil qui segmente un énoncé

2.1.2.2. Les concessifs (22,48% des mouvements étudiés)

On observe une hausse de sourcil brève qui semble être assimilable à ce que Cuxac nomme le "concessif" à savoir "un relèvement marqué des sourcils est associé à un changement de thématique, apportant une restriction par rapport au contenu d'énoncé émis précédemment" (Cuxac, 2000. p. 232). Ces hausses sont ou avoisinent toujours l'amplitude maximale relevée pour chacun des signeurs. Nous distinguerons trois catégories:

- Les "concessifs" en milieu d'énoncé (57,44% des hausses concessives, 10,34% de l'ensemble des mouvements étudiés): Les mouvements de

hausse ne portent pas sur l'énoncé entier mais uniquement sur le signe qui indique que la suite est une restriction. Par exemple, [JE ME SUIS FOCALISEE SUR EXPLOSION TOULOUSE MAIS – accompagné d'une hausse des sourcils- LES INFORMATIONS] (Figure 50). La hausse ne porte que sur [MAIS]. Ces mouvements de sourcils sont très brefs, de l'ordre de 5 à 10 images soit 200 à 400 ms. Leur durée est la même que celle du signe : le début de la hausse, son amplitude maximale et la fin de la hausse sont comprises dans la durée de la production du signe. A noter que l'amplitude maximale de la hausse est atteinte au début, en moins de 4 images soit 160 ms et diminue progressivement (en général avec un pallier entre deux baisses d'amplitude moyenne puis faible).

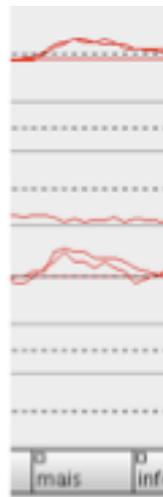


Figure 50 Illustration d'une hausse concessive

- Les “ concessifs ” en début d'énoncé, après une pause (27,65% des hausses concessives et presque 5% des mouvements totaux des sourcils) : Là encore, il s'agit d'apporter des restrictions à ce qui a été dit avant. Cependant, ces restrictions sont apportées après une pause et sont souvent introduites par un pronom interrogatif qui n'a pas de fonction interrogative. Les mouvements de hausse des sourcils sont là encore très brefs : ils commencent en même temps que le pronom interrogatif ou que le premier terme qui apporte la restriction et se terminent avec lui ou 1 image soit 40 ms plus tard. Par exemple, Figure 51: [pause QUOI MOI UN PEU PAS D'ACCORD] que l'on pourrait traduire par “ cependant ” ou “ enfin moi je ne suis pas vraiment d'accord ” ou encore [COMMERCE SOLIDE LA] pause + hausse [QUOI UN AVION PERCUTE] traduisible par “ les commerces étaient

solides, pourtant (ou mais) un avion etc.”. L’amplitude maximale est également atteinte très rapidement, en moins de 3 images soit 120 ms pour diminuer ensuite par étapes (deux en général : une première baisse de faible amplitude, suivie d’une baisse d’amplitude moyenne)

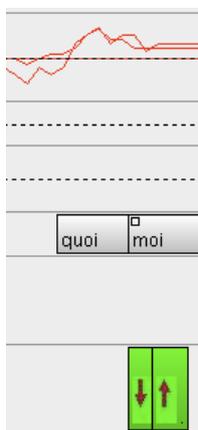


Figure 51 Illustration d’une hausse concessive introduite par un interrogatif après une pause.

- Les “ concessifs ” portant sur l’ensemble de l’énoncé (14,8% des hausses concessives et 2,68% des mouvements totaux des sourcils) : moins présents dans notre corpus, les cas décelés montrent une forte hausse qui commence avant la restriction pour atteindre son amplitude maximale au début de celle-ci. Par exemple : [NEW YORK JAMAIS AVION AU DESSUS JAMAIS] – hausse - [A CE MOMENT LA NEUF HEURE REGARDE EN L’AIR L’AVION S’ECRASE DANS LA TOUR]. La hausse se termine avec [TOUR].

2.1.2.3. La mise en évidence du rhème (14,83% des mouvements étudiés)

Il s’agit d’une hausse des sourcils qui commence à la fin du signe qui précède le rhème et dure tout le rhème. Par exemple : [EXPLOSION] - hausse (dont l’amplitude est atteinte exactement pour le début du signe suivant) - [TOULOUSE A Z F] et qui dure tout le temps de cette séquence. Ces mouvements sont toujours accompagnés d’un clignement des yeux (voir section 2.2.3.2).

A noter que ces hausses se rapprochent des hausses de fausses questions (cette hausse rappelle celle associée à la question “ où ? ”, voir section 2.1.2.6) mais sont moins marquées et concernent surtout l’intérieur des sourcils (Figure 52, première courbe). De plus, elles ne sont pas introduites par un interrogatif.

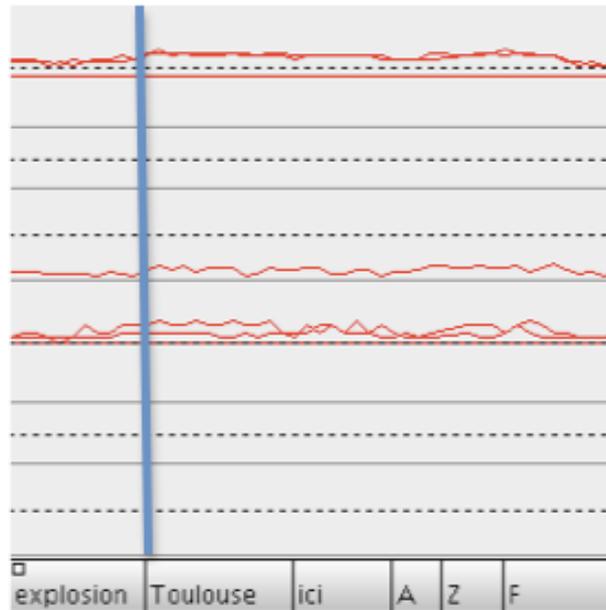


Figure 52. Illustration du mouvement des sourcils juste avant le même.

2.1.2.4. Les transferts personnels (14,35% des mouvements étudiés)

Lors d'un transfert personnel, nous pouvons constater qu'à chaque fois que le signeur passe d'un personnage à l'autre il y a une très nette modification de la posture des sourcils. Cette modification se passe juste avant le changement de personnage afin d'atteindre l'amplitude maximale ou presque au moment où il "devient" le nouveau personnage. A noter que le signeur n'attribue pas forcément un mouvement de sourcils à un personnage de manière fixe. Autrement dit, parfois un personnage va être caractérisé par les sourcils levés et la fois suivante par les sourcils froncés. Ainsi, c'est l'alternance des postures qui semble significative lors du changement de personnage (Figure 53).

Si les sourcils du signeur changent de posture alors qu'il n'y a pas de changement de personnage, ce changement sera moins marqué, plus progressif que lors d'un changement de personnage et nécessite une longue durée. Ensuite, lors du changement de personnage suivant, les mouvements des sourcils (de hausse à baisse ou de baisse à hausse) seront inversés. Dans notre corpus, cette alternance de hausse et de baisse des sourcils dans le passage d'un locuteur à l'autre est systématique.

La Figure 53 montre un transfert personnel : le signeur rapporte un dialogue qu'il a eu avec sa femme et alterne donc en transfert personnel entre son propre discours et celui de sa femme.

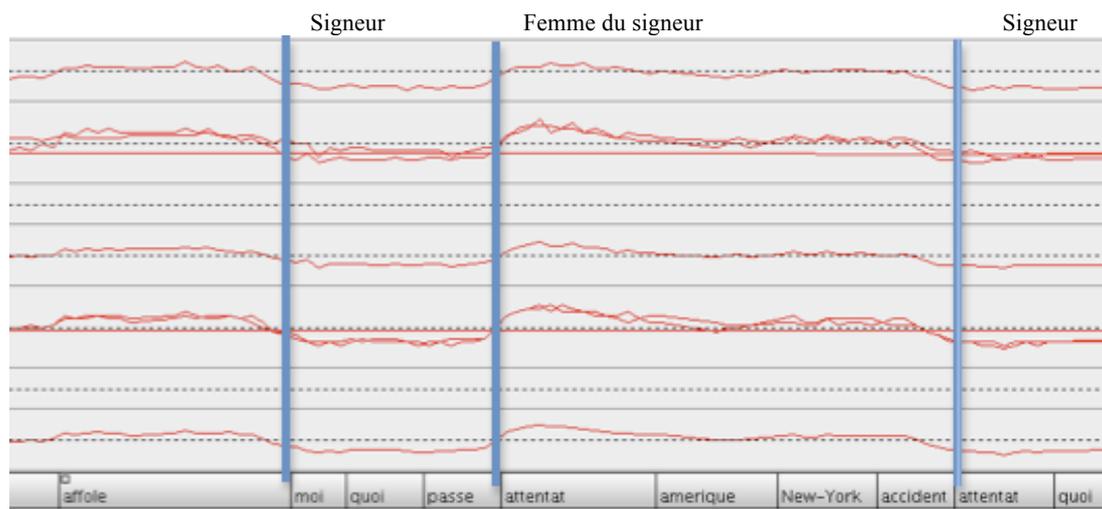


Figure 53 Illustration des mouvements des sourcils lors d'un changement personnage dans un TP. (Signeur 1).

Comme nous l'avons déjà mentionné section 2.1.1, nous pouvons observer que les amplitudes maximales ne sont pas maintenues : elles seront atteintes très rapidement après la prise de rôle et vont progressivement diminuer.

2.1.2.5. La problématisation (13,39% des mouvements étudiés)

Nous relevons des baisses de sourcils de “ problématisation ” que Cuxac définit de la façon suivante : “ Le froncement de sourcils est associé à une problématisation (dans le sens ou ce qui est dit fait problème) du contenu de l'énoncé par l'énonciateur. On pourrait traduire en français, sans trahir, par “ le problème est que ”, “ ... et cela pose problème ” (Cuxac, 2000 ; p. 232). Ces baisses comprennent toujours les mêmes étapes : les trois parties des sourcils, avant le début de la problématisation, quelques soient leurs positions précédentes, atteignent une posture de hausse d'amplitude faible (la plus petite des hausses constatée pour chacun des signeurs), et ce pour une durée d'une image. Ce mouvement n'est pas perceptible mais est révélé par les données fournies par les coordonnées des points positionnés sur la vidéo. Les trois points baissent ensuite ensemble (baisse de faible amplitude) puis, dès l'image suivante, l'intérieur baisse deux fois plus vite que le point du milieu et l'extérieur des sourcils. La baisse maximale (d'amplitude faible ou moyenne) est atteinte en 3 ou 4 images (soit 120 ou 160 ms), ne sera tenue que quelques images et diminuera ensuite progressivement. Si la problématique dure, ces mouvements de baisses successives vont se répéter (toutes les 7 images environ), comme le montre la dernière courbe de la Figure 54.

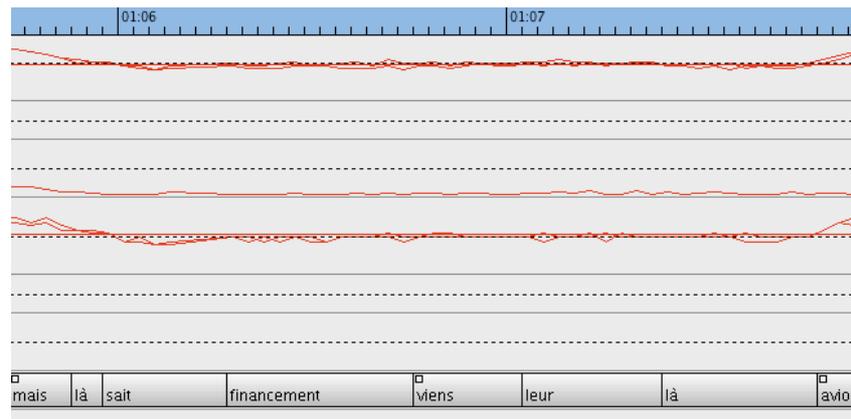


Figure 54 Illustration d'une problématisation.

Les sourcils sont en réalité très proches de leur posture de repos tout au long de l'énoncé (à l'exception des courtes baisses évoquées ci-dessus) et c'est au niveau de la distance entre les sourcils que le mouvement va être perceptible. Celle-ci est “ normale ”⁷⁶ au tout début de la problématisation et va très rapidement (en 2 à 3 images soit 80 à 120 ms) diminuer pour atteindre une distance égale aux trois quarts de la distance minimale possible (voir exemple à la fin de ce paragraphe). Nous pouvons à chaque fois constater que la distance va alors augmenter, en 3 ou 4 images soit 120 ou 160 ms, pour atteindre un tiers de la distance minimale possible, puis à nouveau diminuer pour atteindre à nouveau les trois-quarts de la distance minimale possible. Par exemple, dans l'exemple 55, la distance passe de 0 (distance en position de repos) à -5 en 2 images (80 ms) (pour une distance minimale de -7 constatée chez ce signeur sur l'ensemble du corpus) puis va augmenter pour atteindre les -2 en 4 images, soit 160 ms, et va enfin à nouveau diminuer et se stabiliser à -5 jusqu'à la fin de l'énoncé. La courbe représentant la distance entre les sourcils n'est pas présente dans notre annotation.

⁷⁶ Par normale nous entendons qu'elle est égale à la distance mesurée lorsque le visage est au repos.

2.1.2.6. Les fausses questions (11,48% des mouvements étudiés)

Lorsqu'une question est posée par un pronom interrogatif (quand, où, quoi et qui) et que la réponse est apportée juste après (par exemple les fausses questions fonctionnant comme introducteur de subordonnées ou comme marqueurs thématiques relevés par Cuxac), nous constatons une hausse du sourcil en fin de signe. L'amplitude de la hausse est maximale un peu avant ou exactement là où va commencer la réponse, diminue sur la fin, pour s'achever en même temps que la réponse ou un peu avant. La Figure 55 et la Figure 56 illustrent ces mouvements. Dans le premier exemple, [QUESTIONNER ? 11 SEPTEMBRE 2001], le terme "questionner" correspond, selon le traducteur, à "je pose la question", que l'on pourrait remplacer par "quand?". Le signeur pose donc une question à laquelle il répond immédiatement. La hausse débute en fin du signe [QUESTIONNER] et diminue progressivement. Le second exemple, [OU ? NEW YORK] suit le même procédé mais la hausse ne sera pas maintenue sur tout le signe de réponse. Il semble que la hausse n'est maintenue que si la réponse est composée d'une séquence de signes et ne l'est pas si la réponse est apportée par un seul signe.

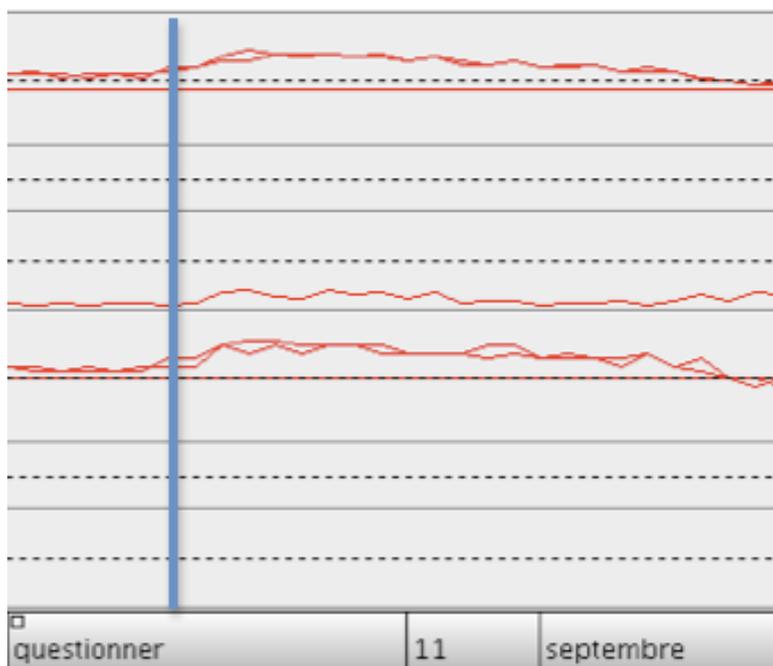


Figure 55 Illustration d'une hausse en fin d'interrogatif, durant toute la réponse.

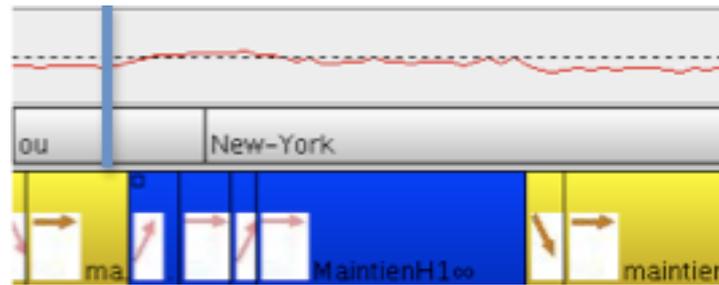


Figure 56 Illustration d'une hausse en fin d'interrogatif, non maintenue durant toute la réponse.

Ce type de hausse est régulièrement accompagné d'un clignement qui sera situé exactement au début de la hausse, en fin d'interrogatif et se terminera avant la réponse (Figure 57). Ce type de clignement est abordé plus en détail dans la section 2.2.

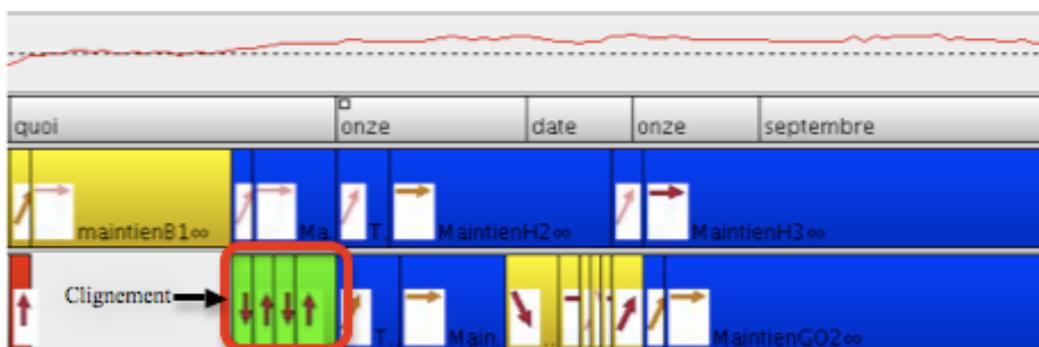


Figure 57 Illustration d'une hausse de fin d'interrogatif avec clignement.

A noter que nous n'avons pas, dans notre corpus, de vraies questions avérées, excepté peut-être lors des transferts, mais le seul cas où les GNM interrogatifs semblent être présents n'est pas analysable car la tête est tournée. Dans tous les autres cas, les mouvements des sourcils sont contraints par le transfert et ne correspondent pas à ceux attendus pour une interrogation. En effet, deux des signeurs racontent en transfert le moment où ils apprennent qu'il y a eu des attentats. Dans ces extraits, nous pouvons supposer que les questions du type "Quoi?", "Qu'est-ce que tu veux?" ou "Qu'y a-t-il?" sont de vraies questions qu'il se posait à ce moment là et donc, en toute logique, que l'interrogation devrait être présente dans le transfert. Cependant, le passage à ce personnage s'est fait par un froncement des sourcils, la hausse interrogative n'est alors pas présente. L'exemple de la Figure 58 illustre ce phénomène. La question est "Quoi? Qu'est-ce qu'il se passe?" il s'agit donc réellement d'une question que la personne transférée pose à son interlocuteur. Or, pour "devenir" ce personnage, le signeur a froncé les sourcils (levés pour "être" l'autre personnage) et les maintient froncés alors qu'il pose ses questions.



Figure 58 Illustration d'une question lors d'un transfert.

2.1.3. Bilan

Notre analyse a permis de définir quelques grandes propriétés des mouvements des sourcils et de dégager une catégorisation des mouvements qui sont apparus dans notre corpus. A noter que la plupart des mouvements portent sur l'ensemble d'un énoncé (donc plusieurs signes). Tel est le cas des mouvements de mise en évidence du rhème, liés aux transferts personnels, de problématisation et liés aux fausses questions tandis que seuls les mouvements de segmentation d'énoncés et la majorité des concessifs ne portent que sur un signe. Cependant, même lorsque le mouvement porte sur l'ensemble de l'énoncé, il ne sera pas maintenu. Il va rapidement atteindre son amplitude maximale puis va diminuer très progressivement sur l'ensemble de l'énoncé.

L'analyse est fine mais incomplète et nécessite une analyse ultérieure portant sur la totalité des mouvements (section 3.2). Elle permet toutefois de disposer de premières informations quantitatives et qualitatives sur les mouvements des sourcils, nous permettant également d'observer le lien entre les mouvements des sourcils et les clignements des yeux. La section suivante présente une synthèse des résultats provenant des analyses effectuées sur les clignements des yeux.

2.2.Observation des clignements des yeux

Nous définissons tout d'abord ce que nous appelons clignement des yeux puis rappelons les connaissances dont nous disposons actuellement concernant leur durée et leur fréquence. Dans une seconde section, nous comparons ces connaissances aux données issues du corpus permettant ainsi de suggérer que les clignements jouent en LSF un rôle plus important que dans les LV. Enfin, nous proposons une classification des différents types de clignements relevés dans notre corpus. Chacune des catégories est définie aussi bien en ce qui concerne sa fonction qu'en ce qui concerne la structure et l'emplacement dans l'énoncé des clignements qui la composent.

2.2.1. Définition et propriétés

Dans la présente section sera considéré comme clignement toute fermeture des yeux ne dépassant pas une image du corpus (soit 40 ms). Un clignement se compose donc uniquement de deux phases : une phase de fermeture de l'œil et une phase d'ouverture quelque soit leur durée, seule la fermeture complète est contrainte à ne pas dépasser une image. Si une pause est faite à la fermeture (donc les yeux fermés pendant plus d'une image) il ne s'agira pas d'un clignement mais d'une fermeture des yeux.

Un adulte (Entendant) cligne des yeux 15 à 20 fois par minute (soit toutes les 3 secondes environ) et chacun de ses clignements dure environ 160 millisecondes. Les clignements volontaires, eux, ont une durée d'environ 250 millisecondes (Charbonnier, 1995). Collet rappelle que “ les clignements volontaires peuvent servir d'actionneur pour l'interaction par le regard d'autant qu'ils sont suffisamment plus longs que les clignements involontaires pour que l'on puisse les détecter sans ambiguïté ” (Collet, 1999 ; p. 110).

2.2.2. Observation de la structure des clignements du corpus

Au regard de l'annotation (presque 200 clignements), il est apparu que la durée type de la première phase d'un clignement (la fermeture) est de deux images (soit 80 millisecondes). Comme nous pouvons le voir sur la Figure 59, la durée des fermetures peut prendre trois valeurs : une que l'on pourrait qualifier de standard (70% des clignements) de deux images (80 ms), une plus marquée et plus rare (24%) de trois images (120 ms) et enfin une beaucoup plus rare (6%) et brève d'une image (40 ms). Cette dernière est utilisée essentiellement lorsque les yeux sont déjà pratiquement fermés.

Répartition des durées de fermeture des clignements

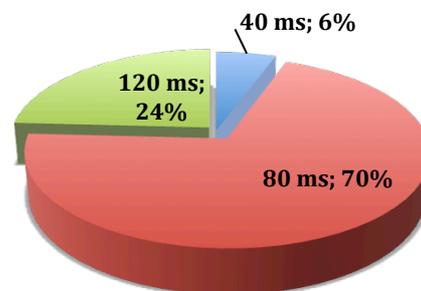


Figure 59 Répartitions des durées de la première phase des clignements (fermeture) pour l'ensemble du corpus.

La durée de la seconde phase, l'ouverture, est plus variable mais nous distinguons tout de même trois durées privilégiées : 46% des ouvertures durent 3 images (soit 120 ms), 37% durent 2 images soit 80 ms, 14% durent 4 images soit 160 ms et 3% des clignements ont une ouverture de 5 images ou plus (Figure 60).

Répartition des durées d'ouverture des clignements

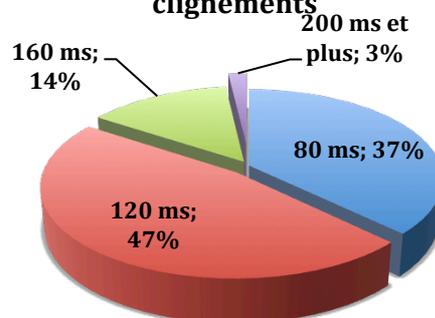


Figure 60 Répartitions des durées de la seconde phase des clignements (ouverture) pour l'ensemble du corpus.

La durée totale d'un clignement, incluant les deux phases, est de 5 images (200 ms) ou plus pour 68% des clignements et 4 images (160 ms) pour 29% des clignements. La Figure 61 ci-dessous récapitule la répartition des clignements du corpus en fonction des durées totales (2 phases comprises).

Répartition des durées totales des clignements

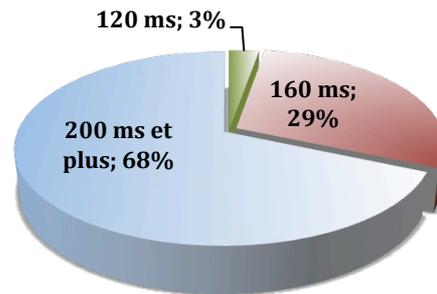


Figure 61 Répartition des durées totales (2 phases comprises) des clignements du corpus.

Nous pouvons observer d'après ce graphique que 68% des clignements sont supérieurs ou égaux à 5 images (200 ms) dont plus de la moitié supérieurs à 6 images (240 ms). Nous pouvons supposer qu'il s'agit de clignements volontaires selon les connaissances dont on dispose sur les clignements des Entendants.

Nous constatons également que la fréquence des clignements est nettement plus élevée chez les quatre signeurs de notre étude que la moyenne des Entendants (qui est de 15 à 20 clignements par minute) à savoir respectivement pour chacun des signeurs de notre étude : 24, 42, 30 et 47 clignements par minute soit une moyenne de 36 clignements par minute ce qui signifie un clignement toutes les 1,40 secondes environ. En observant la Figure 62 ci-dessous, représentant les emplacements des clignements dans le temps (en abscisse), il semble que l'emplacement des clignements ne soit pas aléatoire.

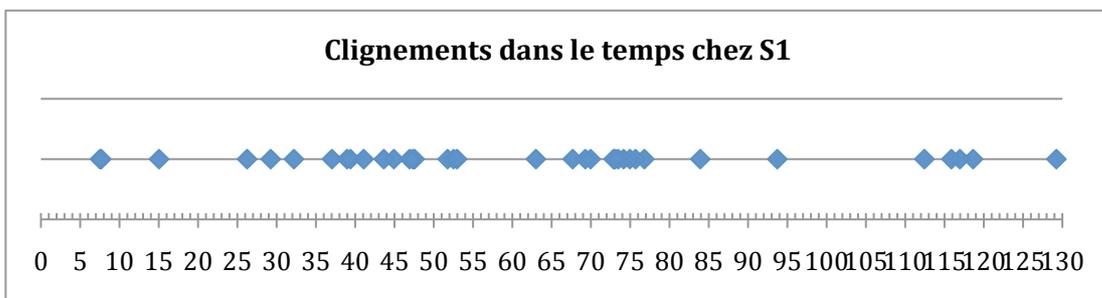


Figure 62 Étendue des clignements dans le temps chez le signeur S1.

En effet, le signeur S1 par exemple peut cligner des yeux jusqu'à 5 fois en à peine plus d'une seconde alors qu'aucun clignement n'est observé durant plus de 18,5 secondes. Nous avons donc recherché les similitudes entre les clignements, à savoir les durées et les emplacements dans l'énoncé par rapport aux signes avoisinants, afin de tenter de catégoriser les clignements et d'en connaître leurs fonctions.

2.2.3. Proposition de typologie des clignements

En observant les 200 clignements du corpus, nous avons dégagé sept types de clignements⁷⁷ en fonction de leur emplacement, du contexte et de leurs durées. La répartition des clignements en fonction de chaque catégorie est représentée par la Figure 63.

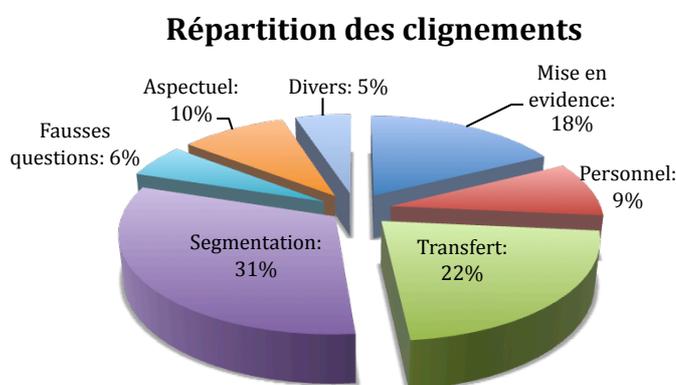


Figure 63 Répartition de l'ensemble des clignements

2.2.3.1. La segmentation (31,5% des clignements)

Ce type de clignement sert à structurer les propos. Nous distinguons quatre sous-catégories de clignements affiliées aux clignements de segmentation : la première regroupe les clignements qui se situent avant l'énoncé et en annoncent le début ; deux sous-catégories regroupent les clignements qui se situent pendant l'énoncé et structurent les propos ; enfin la dernière sous-catégorie regroupe les clignements qui se situent entre deux énoncés afin de les délimiter.

⁷⁷ 7% des clignements ont fait l'objet de doutes et n'ont donc pas été pris en compte dans cette étude. Ils feront l'objet d'une prochaine étude.

- Avant l'énoncé : *Le clignement pré-prise de parole* (15,87% des clignements de segmentation) : Situé avant la prise de parole (Figure 64), ce clignement dure soit au moins 6 images (240 ms) soit uniquement 5 images (200 ms) mais est alors suivi d'un second clignement. La présence de ce phénomène a également été constatée chez d'autres signeurs du corpus non pris en compte pour cette étude. A noter qu'il existe également le clignement de fin d'énoncé qui se termine exactement en même temps que le signe et dure lui aussi au moins 5 à 6 images (200 à 240 ms).

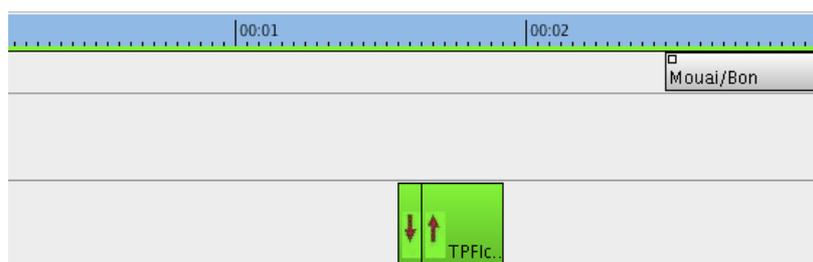


Figure 64 Illustration d'un clignement de pré-prise de parole.

- Pendant l'énoncé : *Le clignement de délimitation de groupes syntaxiques* (31,74% des clignements de segmentation) : Au sein d'un même discours, il s'agit de clignements qui segmentent et structurent la production. Par exemple, [ÇA ARRIVE, SAVOIR QUI FAIT] où le clignement se situe à la fin de [ARRIVE] (Figure 65). Ces clignements sont toujours situés en fin de signe, se terminent en même temps que le signe ou 1 image (40 ms) avant et durent systématiquement 4 images soit 160 ms (2 images pour chacune des phases).

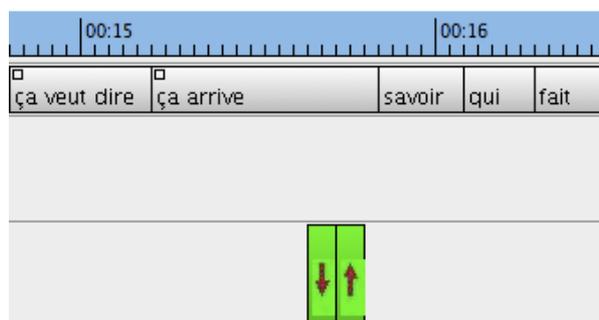


Figure 65 Illustration d'un clignement de délimitation de groupes syntaxiques.

- Pendant l'énoncé : *Le clignement de délimitation de pauses* (25,39% des clignements de segmentation) : Le signeur cligne des yeux pour marquer une pause durant laquelle il réfléchit, puis cligne des yeux au moment de reprendre (Figure 66). La durée de chaque clignement est bien définie et correspond exactement à 6 ou 7 images (240 ms) pour le premier clignement et 4 ou 5 images (160 ms ou 200 ms) pour le second.

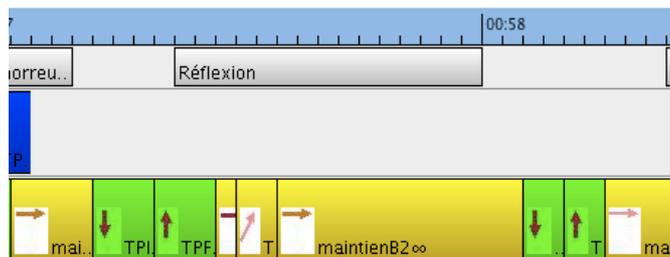


Figure 66 Illustration de clignements de délimitation de pauses.

- Entre 2 énoncés: *Le clignement de segmentation d'énoncés* (26,98% des clignements de segmentation) : Il s'agit d'un clignement qui place une frontière qui indique que l'idée précédente est à archiver et que le signeur enchaîne sur une nouvelle idée. Par exemple : “ [BUILDINGS BEAUX] – clignement – [NEW YORK DEUX TOURS] ”. Ici le signeur explique d'abord que New York est une grande et belle ville avec de beaux buildings puis dans un second temps qu'à New York il y avait deux tours etc. La durée des phases du clignement est stable : 3 images pour la fermeture et 2 images pour l'ouverture soit un clignement de 5 images (200 ms) quelque soit le signeur. La fin de la première phase du clignement coïncide exactement avec la fin du signe précédent (ou une image avant) et le signe suivant commence au plus tôt à la fin du clignement, voire même, dans la majorité des cas, après une brève pause. Ce type de clignement accompagne souvent un léger mouvement de recul arrière de la tête et un mouvement particulier des sourcils. Selon la durée de la pause, le mouvement très bref de hausse des sourcils est situé juste avant (pour une longue pause) ou pendant (pour une pause plus brève) le clignement. Dans ce second cas, la hausse commence exactement à l'ouverture des yeux (seconde phase d'un clignement). La Figure 67 illustre ce type de clignements. Nous pouvons y observer l'emplacement du clignement ainsi que la durée de chacune de ces phases.

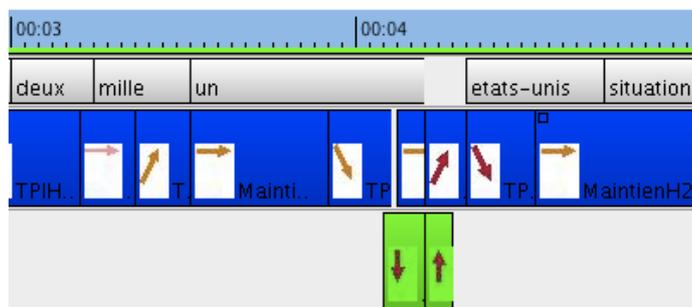


Figure 67 Illustration du clignement de segmentation d'énoncé avec mouvement des sourcils.

Dans cette même catégorie, nous relevons le clignement qui termine une réflexion et commence une nouvelle idée. Il commence à la fin du signe de réflexion du type [BON] et se termine juste avant ou au début du signe suivant. Ce clignement accompagne tout le temps une modification visible des sourcils. Par exemple, l'extrait: “ [BON] - clignement sur la fin du signe - [MOI] ”. Ces clignements durent entre 6 et 7 images (soit 240 à 280 ms) : 3 images (120 ms) pour la fermeture et 3 ou 4 images (120 ou 160 ms) pour l'ouverture. La Figure 68 illustre, à l'aide de deux extraits de l'annotation, l'emplacement, la taille et la simultanéité avec le mouvement des sourcils, de ce type de clignement. Les clignements sont représentés par les blocs A et B, les sourcils sont représentés par la première, les deuxièmes et la dernière courbes. Le trait vertical indique le début du clignement et donc le début du mouvement des sourcils. Dans ces exemples, le premier clignement dure 6 images (240 ms) et le second 7 images (280ms).

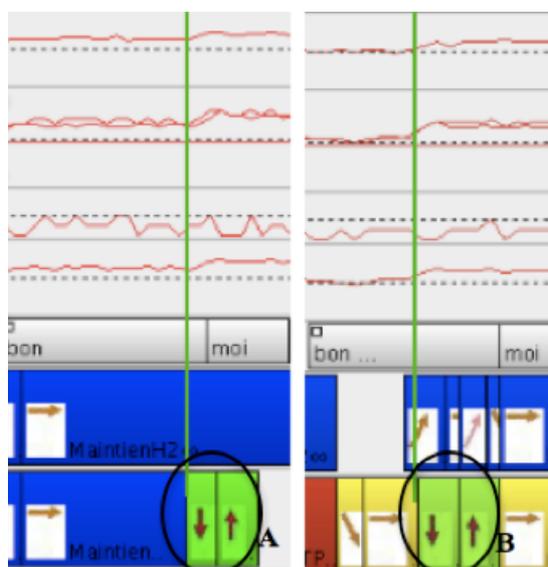


Figure 68 Illustration du clignement de délimitation de fin de réflexion et de nouvel énoncé.

2.2.3.2. La mise en évidence (17,5% des clignements)

Ce type de clignement permet de mettre en évidence un élément du discours. Nous distinguons la mise en évidence du rhème, la mise en évidence de l'objet pointé et la mise en évidence par isolement. Ces clignements semblent être un procédé de mise en relief et peuvent s'apparenter aux topicalisations ou aux focalisations.

- *La mise en évidence du rhème* (40% des clignements de mise en évidence): ce type de clignement se situe avant un rhème et le met en relief comme une topicalisation. Il pourrait se traduire par un “ quoi ” non signé. Par exemple, dans un énoncé comme “ New York est belle “, la production sera “ [NEW YORK] – clignement - [BELLE] ” que l'on

pourrait traduire par “New York est quoi ? Belle ” ou un énoncé tel que “ [AVANT] – clignement – [11 SEPTEMBRE] ” (“ avant quoi ? avant le 11 septembre ”). Le clignement se situe soit entre le thème et le rhème, soit à la fin du thème et va se terminer en même temps que lui (Figure 69 et Figure 70). De plus, il accompagne toujours un mouvement de la tête voire même souvent des sourcils (comme dans nos exemples). Ces mouvements commencent toujours 40 ms avant ou après le début du clignement.

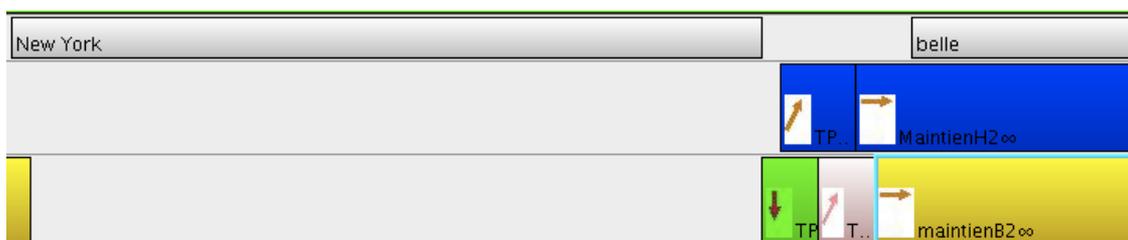


Figure 69 Illustration d'un clignement de mise en valeur du rhème.

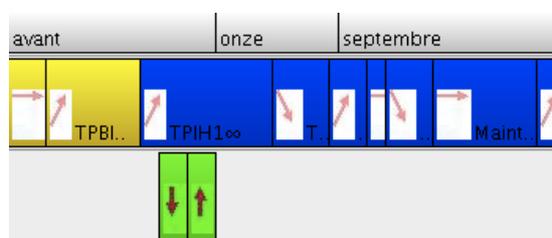


Figure 70 Illustration d'un clignement de mise en valeur du rhème.

- *La mise en évidence de l'objet "pointé"* (31,42% des clignements de mise en évidence): le clignement se situe sur la fin du pointage, qu'il s'agisse d'une personne, d'un lieu ou d'une époque. L'objectif semble être d'insister sur l'objet pointé qui suit le pointage. Par exemple, l'extrait: “ [LA] - clignement - [DATE 11 SEPTEMBRE] ”, que l'on pourrait traduire par “ Là, à cette date, le 11 septembre ”, est illustré Image 29 et Figure 71. Ce clignement a généralement une durée d'au moins 5 images (2 images pour la fermeture, 3 pour l'ouverture), comme dans l'exemple ci-dessous. La fin du pointage se situe soit à la frontière entre la première et la seconde phase du clignement (donc exactement où les yeux sont complètement fermés), soit en même temps que le clignement (à une image, soit 40ms, près). Le début du clignement accompagne toujours un léger recul de la tête et un mouvement des sourcils. Tous les pointages ne nécessitent pas forcément un clignement (et donc de hausse des sourcils). Par exemple, un pointage présent en fin d'énoncé, de spatialisation, auquel on ne fait pas référence immédiatement, n'est pas suivi d'un clignement. Celui-ci n'est présent que pour insister sur l'objet pointé. Ce type de clignement semble se rapprocher des “ mimiques à valeur de démonstratif ” relevées par (Cuxac, 2000 ; p.235). A noter que nous avons relevé de nombreux cas d'auto-pointage, dont certains correspondaient à notre description et d'autres pas. La seconde catégorie

comprend en réalité les clignements de changement de personnes lors d'un transfert personnel, la durée est donc plus courte (4 images soit 160 ms) et la position n'est pas la même (clignement en début de pointage : [MOI] commencera alors en même temps que l'ouverture des yeux).



Image 29 Image correspondante au clignement de l'annotation de la Figure 71

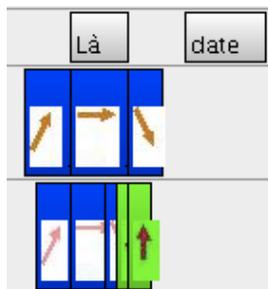


Figure 71 Illustration d'une mise en valeur de l'objet pointé, extraite de notre annotation.

- *La mise en évidence par isolement* (28,57% des clignements de mise en évidence): il s'agit là d'un signe ou de plusieurs signes qui semblent être mis en valeur en étant isolés avant et après leur production par un clignement. Par exemple, pour le signeur S2 : “ [IL A DIT] – clignement - [QUOI ?] - clignement - [IL A DIT QUE] ” (Figure 72). Le premier clignement dure toujours 6 images soit 240 ms (en général 2 images pour la fermeture et 4 pour l'ouverture) tandis que le second clignement ne dure que 4 ou 5 images soit 160 à 200 ms (selon les signeurs, la durée sera toujours de 4 ou toujours de 5 images). Lorsque le signeur souhaite rectifier ou apporter une précision à ce qu'il explique, il va interrompre l'énoncé, cligner des yeux, ajouter l'élément, cligner des yeux, puis reprendre là où il en était.

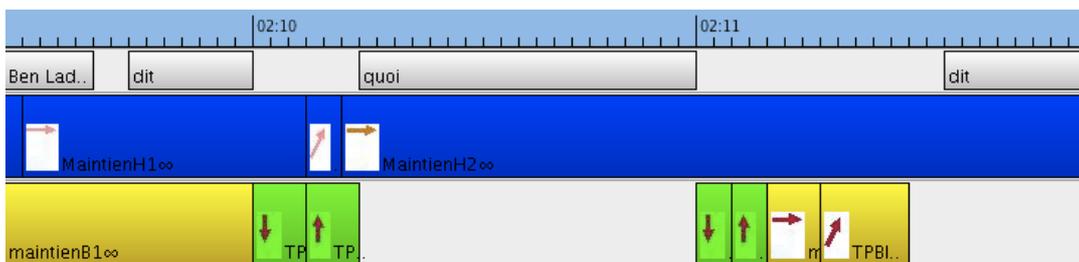


Figure 72 Illustration de la mise en valeur par isolement.

2.2.3.3. Le transfert (22% des clignements) :

Nous distinguons actuellement les clignements liés aux fins de transferts (situationnel ou personnel), le clignement relevé par Cuxac du passage d'une personne à l'autre dans un transfert personnel, le clignement de début de transfert et le clignement à origine sémantique lors d'un transfert situationnel.

- *Les clignements de fin de transfert* (situationnel en général) (27,27% des clignements liés aux transferts) : Ces clignements commencent à la fin du dernier signe du transfert. Le signe suivant commence soit durant la phase d'ouverture des yeux, soit juste après le clignement. Ces clignements ont une durée stable de 5 images (200 ms) sauf s'ils suivent une expression d'émotion, où leur durée s'élèvera à 6 images soit 240 ms (3 images pour l'ouverture et la fermeture), probablement en raison de l'assimilation du clignement personnel (explicité ci-dessus).

- *Les clignements de passage d'une personne à l'autre dans un transfert personnel* (27,27% des clignements liés aux transferts) : Comme nous l'avons déjà évoqué le clignement n'est pas systématique lors d'un changement de personne en transfert personnel dans notre corpus. Nous avons noté que dans le cas de double clignements lors d'un changement de rôle, les durées sont stables : le premier clignement dure 4 images (2 images pour chacune des phases) et le second 5 images (2 images pour la fermeture et 3 pour l'ouverture).

- *Les clignements de début de transfert* (25% des clignements liés aux transferts). A noter qu'au regard de notre annotation, il semble que les clignements de début de transfert sont utilisés pour indiquer le début d'une série de transferts et se situent exactement entre deux signes. Ils ne sont pas présents à chaque transfert mais avant le premier transfert de la série.

- *Le clignement à origine sémantique lors d'un transfert situationnel* (20,45% des clignements) : Il s'agit des clignements nécessaires à la production d'un transfert situationnel exprimant un choc comme [L'AVION QUI PERCUTE LA TOUR]. Ces clignements ont un emplacement bien précis par rapport aux gestes manuels: ici, le clignement se situe au moment où la main représentant l'avion touche le bras représentant la tour, Image 30 et Image 31) et a une durée supérieure à 5 images (200 ms). De plus, la fin de la seconde phase du clignement (l'ouverture) peut déborder sur le signe suivant. Enfin, le clignement peut également être présent dans le cas où la tour n'est pas représentée, au moment où l'avion est supposé percuter la tour (Image 32). A noter que cette expression peut, dans de rares cas, être réalisée sans clignement des yeux. Cependant, l'interprète la traduit alors différemment (par exemple [UN AVION PERCUTE UNE TOUR], [ATTENTAT], etc.).



Image 30 Illustration du clignement lors de [L'AVION QUI PERCUTE LA TOUR].



Image 31 Illustration du clignement lors de [L'AVION QUI PERCUTE LA TOUR].



Image 32 Illustration du clignement lors de [L'AVION QUI PERCUTE] sans le bras de la tour.

2.2.3.4. Le clignement aspectuel de répétition (9,5% des clignements)

Il s'agit de clignements qui semblent indiquer l'abondance, la multitude, la répétition d'une action. Ces clignements interviennent durant la répétition d'un signe, sont généralement très brefs et répétés plusieurs fois avec une faible réouverture des yeux entre les clignements. Ils s'accompagnent systématiquement d'un souffle d'air, d'une labialisation qui pourrait s'apparenter à “ | pa pa pa | “ et d'un mouvement latéral droite-gauche répété de la tête (ce dernier élément n'est pas systématique mais est présent dans 85% des cas). Ces clignements semblent être liés à la “ mimique d'exhaustivité ” et/ou au “ duratif ” relevés par Cuxac.



Image 33 Illustration du clignement de répétition. Ici [LES CAMERAS FILMENT TOUT AZIMUT].



Image 34 Illustration du clignement de répétition. Ici [ILS PARTENT].

2.2.3.5. Le clignement personnel (9% des clignements)

Il s'agit ici de clignements qui permettent d'indiquer que l'énoncé se rapporte au signeur ou d'insister sur le fait que le signeur parle de lui. Par exemple, pour des signes tels que [PENSER] ou [CONNAITRE] qui ne sont pas directionnels, le clignement remplace le pronom personnel “ Je, moi ” (Image 36).



Image 35 Illustration du signe [MOI] accompagné d'un clignement.



Image 36 Illustration du clignement personnel [JE CONNAIS]

A distinguer dans cette catégorie le clignement accompagnant l'expression d'une émotion : l'émotion se rapporte au signeur. Cela est indiqué grâce au clignement qui accompagne systématiquement un recul de la tête et/ou du buste. De plus, la bouche est toujours ouverte en fin de signe (juste après le clignement). Ces clignements se produisent toujours au milieu du signe et sont très longs. Dans certains cas, les clignements sont remplacés par une fermeture des yeux de quelques images. L'Image 37 illustre ce clignement accompagnant une émotion.



Image 37 Illustration du clignement accompagnant une émotion

2.2.3.6. Les fausses questions (5,5% des clignements)

Lorsqu'un pronom interrogatif est utilisé dans le cadre d'une fausse question, il est accompagné d'un clignement des yeux, en fin de production, et d'un mouvement des sourcils comme abordé dans la section 2.1.2.6. Ces clignements commencent 1 à 2 images avant la fin du signe et se prolongent ensuite (très souvent durant une pause). Ils durent généralement 5 images soit 200 ms (Figure 73).

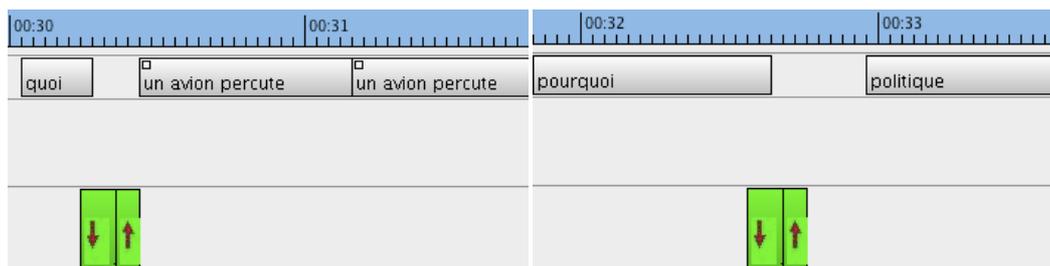


Figure 73 Illustration de clignements liés aux fausses questions.

La Figure 74 ci-dessous illustre la répartition des clignements pour chaque signeur. Elle indique ainsi que si tous les clignements sont bien présents dans l'ensemble des corpus, les proportions sont propres à chaque signeur.

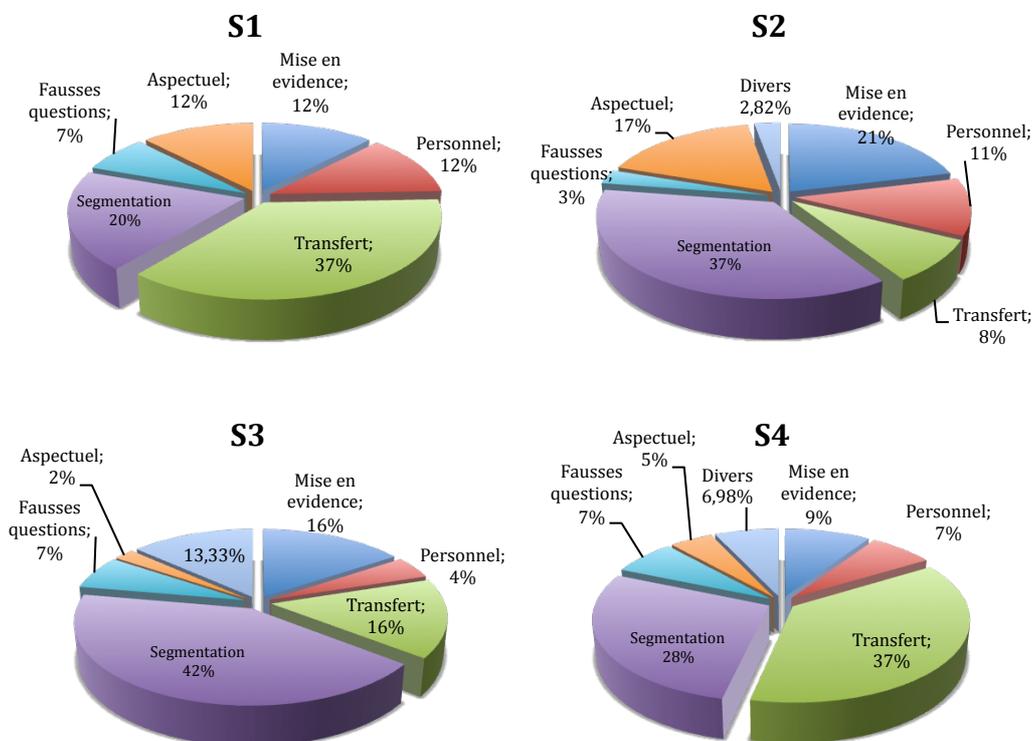


Figure 74 Répartition des types de clignements pour chaque signeur.

2.2.4. Bilan

Nos analyses ont permis de dégager des grandes catégories de mouvements des sourcils et des clignements des yeux. Nous pouvons distinguer les catégories qui fonctionnent en interaction comme les mouvements de segmentation d'énoncés, de mise en évidence du thème et de l'objet pointé et ceux liés aux fausses questions. Toutes ces catégories nécessitent aussi bien un mouvement particulier des sourcils qu'un clignement des yeux. Pour chacune de ces catégories nous avons apporté des informations structurelles. Nous proposons, dans la section suivante, de faire le point sur l'apport de nos travaux aux connaissances actuelles que nous avons présentées dans le premier chapitre.

2.3. Apport de nos résultats aux connaissances actuelles

Lors de notre inventaire des connaissances concernant le non manuel, (chapitre 1, section 2.1), nous avons distingué le niveau lexical du niveau de l'énoncé. Notre apport de connaissances concerne pour l'instant principalement le niveau de l'énoncé.

Les connaissances actuelles mettent en avant le rôle des GNM dans l'expression de la modalité, dans la délimitation des transferts et dans la segmentation d'énoncé. A noter que nous n'avons pas eu l'opportunité de comparer nos travaux aux connaissances linguistiques d'autres LS que la LSF. En effet, celles-ci concernent, comme nous l'avons vu, des éléments non présents dans notre corpus, comme par exemple les interrogations.

Concernant la modalité en LSF, Cuxac a établi une typologie des modalités en proposant pour chaque catégorie un descriptif des comportements non manuels. Nous avons, pour chacune des catégories trouvées dans notre corpus, apporté des précisions quant à la structure et à l'emplacement des mouvements des sourcils concernés. De plus, nous avons expliqué le rôle des clignements des yeux impliqués dans certaines de ces catégories et avons décrit leurs structures.

Vergé a montré que la posture adoptée par les sourcils pour un mode précis sera maintenue durant toute la production. Nous précisons qu'il ne s'agit pas d'un maintien mais d'une très lente diminution (qui s'accélère quelques images avant la fin de la production). De plus, le concessif (assimilables aux modes selon Cuxac) ne porte pas systématiquement sur l'ensemble de l'énoncé mais parfois uniquement sur le signe introductif.

Concernant les topicalisations, Coulter précise que le syntagme qui porte le levé de sourcil fonctionne toujours comme topique. Nous précisons que le levé de sourcil n'est pas l'unique procédé de mise en relief d'un élément mais que le clignement des yeux est également impliqué, voire même remplace parfois la hausse de sourcil.

Au sujet des transferts, Cuxac et Sallandre ont abordé l'importance du non manuel en particulier pour délimiter les transferts personnels et distinguer les types de transferts. Nos travaux ont montré que dans le cas des transferts personnels, les sourcils étaient d'excellents indicateurs des changements de personnages. De plus, les clignements des yeux permettent de délimiter les débuts et fins de transferts.

Enfin, Vergé a indiqué l'importance des clignements des yeux pour la segmentation de l'énoncé en précisant qu'ils délimitaient des "groupes de sens". Nos travaux sur les clignements des yeux ont permis de préciser le rôle de segmentation des clignements tout en apportant des informations quant à leurs structures, leurs emplacements et la durée de chacune des phases qui composent un clignement de segmentation.

Notre étude a donc permis d'apporter des connaissances aussi bien linguistiques que structurelles concernant les clignements des yeux et les mouvements des sourcils. Ces résultats sont une première étape. Cependant, l'objectif étant de générer les GNM, il est nécessaire de proposer une formalisation des différentes catégories de mouvements. Nous présentons dans la partie suivante une tentative de modélisation de chacune de ces catégories.

2.4. Vers une formalisation de la typologie

Pour formaliser nos résultats, nous nous inspirons du modèle Zebedee (Filhol, 2008 ; 2009). La section 2.4.1 en présente rapidement les grands concepts et en explique les limites actuelles en termes de description des GNM. Dans une seconde section (2.4.2), nous proposons quelques ajustements du modèle ainsi que de nouvelles conventions de description afin de répondre aux spécificités du non manuel. Enfin, nous définissons les structures type des mouvements des yeux et des sourcils (section 2.4.3) et récapitulons les attributs propres à chaque catégorie, à ajouter à la structure type afin d'obtenir le mouvement souhaité (section 2.4.4).

2.4.1. Modèle Zebedee : notions et discussion

(Filhol, 2008) propose une modélisation formelle des signes reposant sur la temporalité proposée dans le modèle linguistique de Johnson et Liddell (à paraître). L'auteur explique que “ les signes se découpent en une alternance de postures clés et de transitions, où les premières correspondent aux postures motivées de la signation, et les autres aux passages de chacune à la suivante ” (Filhol, 2008, p. 74). Ainsi, son modèle ne décrit pas un signe globalement mais propose une description de toutes les postures clés grâce à des spécifications. Les mouvements, ou transitions, ne sont eux pas décrits mais sont la conséquence du passage d'une posture clé à l'autre.

Ce découpage temporel correspond à notre distinction des postures et des mouvements qui permettent de passer d'une posture à l'autre. Le modèle de description de Filhol semble donc particulièrement intéressant pour notre étude.

Cependant, si ledit modèle permet la description de l'ensemble des articulateurs, il ne dispose pas encore des primitives pour le visage. Ainsi, nous nous inscrivons dans ce modèle afin d'en proposer une adaptation pour visage. En effet, ce dernier ne répond pas aux mêmes exigences que les articulateurs décrits actuellement (les mains, la tête, etc.) et nécessite donc quelques ajustements. Par exemple, la spatialisation est très problématique pour la description des mains. Concernant les sourcils et les paupières, les degrés de liberté sont beaucoup plus restreints et la question se résume souvent à savoir si l'élément est levé ou pas mais aucunement de savoir où il est placé dans l'espace. Durant un clignement, nous pouvons nous interroger en effet pour savoir si les yeux sont ouverts ou fermés mais ils restent toujours sur le visage et la spatialisation n'est ici pas pertinente (bien entendu nous ne parlons pas du regard mais bien des paupières). L'orientation n'est pas non plus requise pour décrire les mouvements des paupières et des sourcils. En revanche, il est nécessaire de pouvoir synchroniser le mouvement avec le signe manuel, dans le cas où les deux sont distincts. De même, la posture neutre acquiert une place beaucoup plus importante pour nos descriptions que pour les articulateurs, où la posture neutre (ou de repos) n'est quasiment jamais présente (à part peut-être entre deux énoncés) et dans tous les cas difficilement définissable. En effet un sourcil en position de repos est sans conteste un sourcil qui n'est ni levé ni baissé. En revanche, la question d'une main en position de repos soulève davantage de polémique et semble beaucoup moins pertinente.

Le modèle nécessite donc quelques ajustements afin de prendre en compte ces spécificités. Ainsi, nous n'expliquons pas la totalité du modèle, la plupart des éléments étant spécifiques aux mouvements manuels, mais en retenons le découpage en unités temporelles à savoir les postures clés et les transitions.

2.4.2. Conventions de description

Nos descriptions concernent les sourcils et les clignements des yeux, donc les mouvements de la paupière supérieure. Pour chaque mouvement de ces deux éléments, nous distinguons en général (sauf exception abordée ci-dessous), 5 unités temporelles. En effet, comme nous le montre la Figure 75, il y a 3 postures clés (la posture de départ, la posture adoptée et la posture finale) et 2 transitions (représentées par les flèches sur la figure) lors d'une hausse de sourcils. Concernant une baisse, le nombre d'unités temporelles est exactement le même (la posture adoptée étant simplement une posture basse et les transitions sont les mouvements inverses).

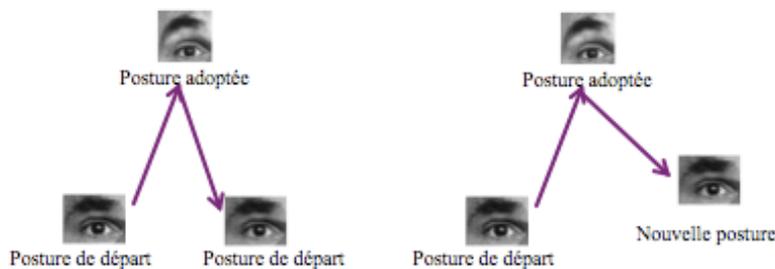


Figure 75 Postures clés et transitions lors d'une hausse de sourcil

A noter cependant que certaines hausses de sourcils (par exemple les hausses concessives) n'ont pas 3 mais 4 postures clés. En effet, comme nous l'avons évoqué, le retour à la position initiale se fait alors en deux fois avec une étape (donc 1 posture clé) entre les deux. Ces hausses ont donc 7 unités temporelles : 4 postures clés et 3 transitions.

Concernant les clignements des yeux, le nombre d'unités temporelles est également de 5 à savoir 3 postures clés (œil ouvert, œil fermé et œil ouvert) et 2 transitions (passage de œil ouvert à œil fermé et passage de œil fermé à œil ouvert), comme l'illustre la Figure 76.



Figure 76 Postures clés et transitions lors d'un clignement

Nous souhaitons préciser, pour chacune de ces unités, sa durée et la posture. De plus, la posture de repos doit pouvoir être utilisée sans besoin de la définir.

Nous tenons à préciser que les postures clés de Zebedee et les nôtres ne se situent pas au même niveau. En effet, Zebedee propose une description des signes, soit une description qui se situe au niveau lexical, voire sous-lexical⁷⁸. Les mouvements des sourcils et les clignements des yeux quant à eux se situent au niveau de l'énoncé et leur description implique plusieurs signes. Lorsque nous définissons une posture clé, il ne s'agit donc pas des mêmes postures clés que Zebedee. Cependant, il est nécessaire de lier les deux niveaux, d'autant plus que la plupart des mouvements étudiés impliquent un signe. Nous proposons donc une synchronisation, lorsque cela est rendu possible, de nos postures clés et de celles des signes de Zebedee. Pour notre description, nous avons donc établi les conventions suivantes :

- Les éléments sont nommés “ P ” (paupières) et “ S ” (sourcils).
- Concernant les sourcils, il est possible de distinguer l'intérieur, le milieu et l'extérieur en ajoutant les lettres “ i ”, “ m ” et “ e ”.
- “ P : 0 ” et “ S : 0 ” indiquent les postures de départ des yeux et des sourcils. Il s'agit, à défaut de spécifications antérieures, des postures de repos. En revanche, si ces postures ont été définies pour un mouvement précédent, et non spécifiées depuis, ce sont elles qui sont les postures de départ et non les postures de repos.
- Le contact entre les deux paupières lors d'un clignement (fin de la première phase) est noté “ contact ”.
- L'amplitude des mouvements est indiquée par les signes “ + ” et “ - ” qui peuvent être doublés ou triplés (pour des amplitudes 1, 2 ou 3).

⁷⁸ Il s'agit de descriptions des postures clés qui composent un signe, d'où le niveau sous-lexical.

- La durée des mouvements (ou transitions) est indiquée en nombre d'images. Dans les cas où les mouvements se terminent en même temps que le signe ou que l'énoncé, nous utilisons "FS" (Fin Signe) ou "FE" (Fin Énoncé).
- Pour indiquer l'emplacement de ces mouvements par rapport aux signes, autrement dit synchroniser nos postures clés avec celles des signes avoisinants, nous devons distinguer le premier signe et le signe suivant. Nous utilisons "PS" (pour Premier Signe) et "SS" (pour Signe Suivant). De plus, afin de préciser l'emplacement du GNM par rapport à un de ces signes, nous ajoutons les lettres "D" (Début) et "F" (Fin). Ces lettres sont situées soit avant, pour indiquer précisément le début ou la fin du signe, soit après, pour indiquer globalement la première ou la dernière partie du signe. Par exemple, pour indiquer qu'un clignement des yeux commence exactement en même temps que le premier signe, nous utilisons "DPS" (exactement le Début du Premier Signe). En revanche, pour indiquer que le clignement commence vers la fin du premier signe (pour se terminer en même temps que lui par exemple), nous utilisons "PSf" (vers la fin du Premier Signe) ("FPS" indiquera la fin du clignement en même temps que le signe). Enfin "OS" indique une pause, donc aucun signe (0 Signe). La Figure 77 récapitule l'utilisation de ces symboles.

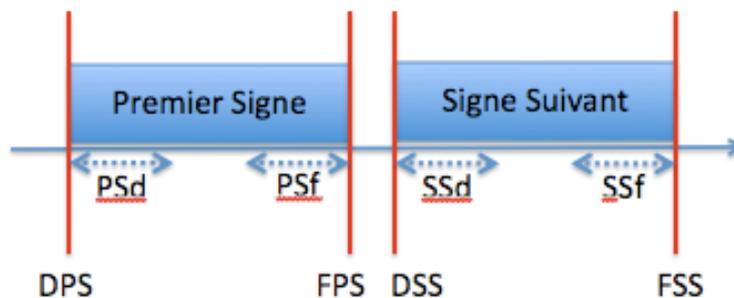


Figure 77 Illustrations des symboles utilisés pour synchroniser nos postures clés à celles des signes

Exemple : le **mouvement de délimitation d'énoncés** (explicité précédemment section 2.1.2.1 et 2.2.3.1, il s'agit de clignement et de brèves, mais prononcées, hausses des sourcils qui marquent une légère pause dans l'énoncé):

```

KEY_POSTURE          /// La durée non spécifiée79
Sync                /// Annonce la synchronisation
PSf                 /// Se situe sur la fin du signe
{
  P: 0              /// La paupière au repos ou dans une posture précédente
}
TRANSITION (3) {}  /// Durée de la transition : 3 images
KEY_POSTURE (1)    /// Durée de la 1ère key_posture : 1 image
Sync
FPS                 /// La fin du signe coïncide avec la fermeture de l'œil
{
  P : contact      /// La paupière supérieure se ferme (contact des paupières)
  S : 0            /// Le sourcil est au repos ou en posture précédente
}
TRANSITION (2) {}
KEY_POSTURE (1)
Sync
DSS                 /// Début du Signe Suivant pendant la transition
{
  P : 0            /// La paupière est revenue à sa posture initiale
  S : +++          /// Le sourcil est en hausse d'amplitude 380
}
TRANSITION (3)
{
  S : ---
}
KEY_POSTURE (FE)    /// Jusqu'à la fin de l'énoncé
{
  S : 0            /// Le sourcil est revenu à sa posture initiale
}

```

⁷⁹ Cette key_posture représente en effet la dernière image stable avant le mouvement. Sa durée dépend du contexte et n'est pas pertinente ici.

⁸⁰ A noter que cette hausse n'est pas absolue mais relative : il s'agit d'une hausse par rapport à sa posture initiale.

2.4.3. Les structures type :

Les structures types regroupent toutes les composantes communes à tous les clignements puis à tous les mouvements des sourcils. L'objectif est de disposer d'un " patron " et de ne définir ensuite chaque catégorie que par ses spécificités. Nous nommons attribut tout ce qui compose une unité temporelle (la durée, l'emplacement par rapport au signe, etc.).

Les attributs variables d'une catégorie à l'autre sont précisés, à titre informatif, entre crochets et en italique, et ce dans le but de répertorier les valeurs possibles. Ces valeurs entre crochets ne font donc pas partie du patron mais permettent d'indiquer ce qu'il est possible d'ajouter, à quel emplacement, et s'il est possible de ne rien ajouter (symbole " X "). Lorsque qu'aucun attribut n'est systématique dans une catégorie, nous ne proposons pas de valeurs entre crochets. Les attributs qui ne varient pas (comme par exemple la durée de la deuxième posture clé d'un clignement) sont indiqués. De plus, afin de faciliter la réutilisation de ces structures, nous avons nommé chacune des unités temporelles par un sigle en italique entre parenthèse (par exemple KPC1 pour " Key_Posture Clignement 1 " soit la première posture clé d'un clignement).

Nous proposons trois structures types pour les mouvements des sourcils : deux pour les hausses et une pour les baisses. Les deux structures de hausses distinguent les hausses dont la seconde transition (retour à la position initiale) se fait par étapes des hausses dont le retour à position initiale ne se fait pas par étapes (il s'agit alors soit du retour très rapide des mouvements de segmentation d'énoncés, soit d'un retour très lent sur l'ensemble d'un énoncé).

Structure type d'un clignement :

```
(KPC1) KEY_POSTURE
Sync
[0S / PSf / FSP / X]          /// En fin de signe exclusivement81
{
    P : 0
}

(TC1) TRANSITION ( [2 / 3 / X ] ) {} /// Si durée connue : 2 ou 3 images.
(KPC2) KEY_POSTURE (1)          /// La durée est toujours de 1 image
Sync
[0S / FPS / X]
{
    P : contact
}

(TC2) TRANSITION ( [2 / 3 / 4 / X ] ) {}
(KPC3) KEY_POSTURE
Sync
[DSS / 0S / FSP / X]
{
    P : 0
}
```

⁸¹ Aucune des valeurs proposées ne contient un 'D', la KPC1 se trouve donc toujours en fin de signe. A noter qu'il est possible de ne rien indiquer ce qui est signifié par la valeur 'X'.

Structure type d'une hausse des sourcils :

```
(KPH1) KEY_POSTURE
Sync
[PSf / FPS / OS / X]
{
    S : 0
}

(TH1) TRANSITION {}
(KPH2) KEY_POSTURE ( [1 / FE / FS / X] )
Sync
[DSS / DSP / OS / X]
{
    S : [+ / ++ / +++]
}

(TH2) TRANSITION ( [FE / FS / X] ) {}
(KPH3) KEY_POSTURE
{
    S : 0
}
```

Structure type d'une hausse des sourcils avec retour à la posture initiale par étapes:

```
(KPH-1) KEY_POSTURE
Sync
[DPS / X]
{
  S : 0
}

(TH-1) TRANSITION {}
(KPH-2) KEY_POSTURE
{
  S+++
}

(TH-2) TRANSITION {
}

(KPH-3) KEY_POSTURE {
  S : [+ / ++]
}

(TH-3) TRANSITION {
}

(KPH-4) KEY_POSTURE {
  S : 0
}
```

Structure type d'une baisse des sourcils :

```
(KPB1) KEY_POSTURE {
S : 0
}

(TB1) TRANSITION {
}

(KPB2) KEY_POSTURE {
  S : [- / -- / ---]
}

(TB2) TRANSITION {
}

(KPB3) KEY_POSTURE {
  S : 0
}
```

Ces structures permettent de décrire tous les mouvements des sourcils et les clignements. Pour chacune des catégories, il suffit d'ajouter les attributs spécifiques. Ces attributs sont regroupés, dans les deux tableaux ci-dessous, par catégorie.

Le premier tableau (Tableau 10) spécifie les attributs propres aux différentes catégories de clignement de chaque unité temporelle en interaction avec les mouvements des sourcils, si ceux-ci sont présents. Ainsi, pour obtenir le clignement propre à une catégorie choisie, il suffit d'ajouter les attributs de la ligne correspondante de ce tableau à la structure type de clignement des yeux. Pour connaître l'emplacement de chacun des attributs, nous avons utilisé les noms attribués à chaque unité de la structure type. Si rien n'est spécifié pour une unité, cela signifie qu'aucun élément n'est à ajouter.

A noter que dans les cas où les postures clés des clignements ne coïncidaient pas avec les postures clés des mouvements des sourcils (par exemple dans le cas de la mise en évidence du rhème où le clignement a entièrement lieu durant la hausse de sourcil), nous avons noté en italique le comportement des sourcils dans le tableau des clignements, pour indication, et réciproquement dans le tableau des mouvements des sourcils. Ces éléments ne sont donc pas à ajouter à la structure type. Dans les cas où les unités coïncident, les catégories ne sont présentes que dans un des tableaux.

		Clignement										Sourcils	
		KPC1		TC1		KPC2		TC2		KPC3			
		Signe	Sourc	Durée	Sourc	Signe	Sourc	Durée	Sourc	Signe	Sourc	Durée	Sourc
Segmente	Pré-prise de parole	OS		2		OS		3 ou 4		OS			
	Groupe syntaxique	PSf		2				2		FSP			
	Pauses	FPS		3 puis 2		OS		3 puis 2		DSS			
	Énoncés	PSf		3		FPS	KPH1	2	TH1	DSS (+1)	KPH2	3	THP2
Mise en évidence	du rhème	FPS	<i>KPH1 ou TH1</i>		<i>TH1</i>		<i>TH1</i>		<i>TS1</i>		<i>KPH2</i>	<i>FE</i>	
	Objet pointé	PSf ou FPS		3		FPS ou OS		2		OS			
Lié aux fausses questions		PSf	<i>KPH1</i>	3	<i>TH1</i>	OS	<i>TH1</i>	2	<i>TH1</i>	DSS	<i>KPH2</i>	FE	

Tableau 10 Caractéristiques de chaque catégorie de clignements

Le Tableau 11 spécifie les attributs propres aux différentes catégories de mouvements des sourcils.

		Sourcils									
		KPH1 / KPH-1		TH1	KPH2 / KPH-2				KPH3 / KPH-3 KPH-4		
		Signe	ccl	Durée	Durée	sourcil	Signe	ccl	Durée	Sourcil	ccl
Concessif	Milieu énoncé	DPS			1	+++			FS	- puis 0	
	Début énoncé	DPS		- de 3	1	+++			FS (ou +1)	-- puis 0	
	Tout l'énoncé	PSf			1	+++	DSS		FE	0	
Mise en évidence	du rhème (1)	FPS	TC1, KPC2 TC2	5 ou 6	FE	++	OS et DSS	TC2 et KPC2		0	
	du rhème (2)	PSf	TC1, KPC2, TC2 et KPC2	8 ou 9	FE ou FS	++				0	
	Objet pointé	OS			FS	+++	DSP	KPC1+ TC1+KPC2		0	TC2
Lié aux fausses questions		PSf	TC1 -> KPC3		- de 3	++	DSS		FE	0	

Tableau 11 Caractéristiques de chaque catégorie de mouvements des sourcils

Dans ce tableau, la seconde ligne contient des doubles sigles, par exemple “ KPH1 / KPH-1 ”. Cela signifie qu’il faut se reporter soit à la structure des hausses types, soit à la structure des hausses avec retour à la posture initiale par étapes. Pour distinguer les deux, il suffit d’observer la troisième posture clé (“ KPH3 / KPH-3 KPH-4 ”) et de voir si une seule posture est notée ou deux. Dans le cas où deux postures sont notées, cela signifie que la première posture est adoptée via la TH-2 tandis que la deuxième est atteinte via la TH-3. Tel est le cas des deux premiers concessifs. Le détail de la formalisation de l’ensemble des catégories est proposé en annexe 1 et 2.

Exemple : Le clignement de segmentation de groupes syntaxiques :

Le clignement commence en fin de signe (“ PSf ”), les transitions durent 2 images et le clignement se termine avec le signe (“ FPS ”). Nous savons également grâce à la structure type d’un clignement que la seconde image clé dure 1 image. La Figure 78 représente ce type de clignement.

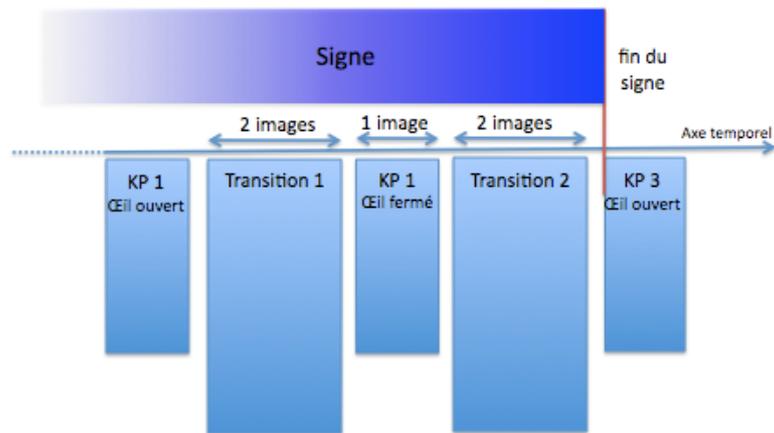


Figure 78 Représentation du clignement de segmentation de groupe syntaxique.

La description est donc la suivante :

```
(KPC1) KEY_POSTURE
Sync
SPF
{
P : 0
}
(TC1) TRANSITION (2) {
}
(KPC2) KEY_POSTURE (1) {
P : contact
}
(TC2) TRANSITION (2) {
}
(KPC3) KEY_POSTURE
Sync
FSP
{
P : 0
}
```

3. Conclusion

Ce chapitre a permis de présenter les données et de fournir des précisions sur l'annotation afin d'appréhender les résultats. Ces derniers concernent aussi bien les mouvements des sourcils que les clignements des yeux. Nous avons présenté tout d'abord des propriétés générales de chacun de ces éléments puis une typologie permettant de distinguer des grandes catégories que nous avons définies aussi bien au niveau structurel qu'au niveau sémantique. Nous avons alors confronté nos résultats aux connaissances actuelles et avons spécifié les apports de nos travaux. Enfin, nous avons proposé une formalisation des catégories des clignements et des sourcils.

Cependant, nos analyses n'ayant porté que sur quatre signeurs, il est nécessaire de poursuivre l'étude et d'étendre l'annotation sur un corpus plus vaste. De plus, les typologies que nous proposons sont à affiner et à confirmer. Nous proposons dans le chapitre suivant des pistes de poursuite de notre étude.

Conclusion et perspectives

1. Conclusion

Cette thèse visait à accroître les connaissances disponibles sur les GNM et à en proposer une formalisation dans le but de les générer automatiquement.

Nous avons proposé une nouvelle méthodologie d'annotation comportant deux phases : une annotation qualitative puis une annotation quantitative. Cette méthodologie a été appliquée aux mouvements des sourcils et aux clignements des yeux mais se veut générique (possibilité de traiter l'ensemble des GNM). Elle a la particularité de se centrer sur le mouvement et non sur la posture. Elle permet ainsi de tenir compte de la temporalité des LS et d'obtenir des données fines et précises concernant les GNM. Ces données sont symboliques (pour l'annotation qualitative) et numériques (pour l'annotation quantitative). Nous proposons ensuite un traitement de ces données dans un logiciel de calcul scientifique afin de disposer de données numériques indépendantes des mouvements latéraux de la tête et des courbes montrant l'évolution des mouvements des sourcils.

Les données, regroupées dans un tableur, ont été analysées séparément (les mouvements des sourcils d'une part et les clignements des yeux d'autre part) d'abord, puis en interaction.

Nous avons proposé, pour les mouvements des sourcils et pour les clignements des yeux, une liste des propriétés générales puis une typologie regroupant les phénomènes similaires récurrents. Pour chacune des catégories, nous avons précisé le rôle du mouvement, son emplacement par rapport aux signes avoisinants et la durée de chacune de ces phases. Nous avons ensuite situé nos résultats par rapport aux connaissances dont nous disposions en début d'étude et constaté que nous avons apporté des précisions quant à la structure des GNM. L'objectif de ces travaux étant de générer automatiquement des GNM, nous avons proposé une formalisation de chacune des catégories des clignements des yeux et des mouvements des sourcils.

Nos travaux ont donc permis d'élaborer et de tester une nouvelle méthodologie d'annotation des GNM mais également d'accroître les connaissances dont nous disposons à leur sujet. Il s'agit toutefois d'une première étude sur un corpus limité et où nous nous sommes restreints aux mouvements des sourcils et des yeux. Nous proposons dans la partie suivante des pistes de prolongement de cette étude.

2. Perspectives de poursuite de l'étude

Cette partie propose des perspectives de poursuite de notre étude. Il s'agit tout d'abord de poursuivre l'annotation, ce qui nécessite de procéder au choix des corpus en fonction des objectifs de l'étude (section 2.1) puis d'automatiser une partie du processus (section 2.2). Nous proposons enfin de procéder à des tests de perception afin de valider l'ensemble de nos observations (section 2.3).

2.1. Annotation

Nous proposons d'étendre l'annotation non seulement à un corpus d'Entendants mais également à d'autres corpus de sourds ciblés afin de valider les observations de notre étude. Nous présentons ensuite brièvement comment nous pourrions diminuer le temps d'annotation en proposant d'automatiser une partie du processus. Enfin, nous apportons quelques précisions concernant l'annotation des éléments non manuels autres que les sourcils et les yeux.

2.1.1. Extension des corpus annotés

Trois types de corpus nous semblent intéressants à prendre en compte par la suite.

Étude simultanée sourds Entendants

Comme nous l'avons déjà évoqué les descriptions actuelles ne permettent pas de connaître les structures internes et temporelles des éléments non manuels. Ainsi, il serait intéressant de procéder à une annotation de corpus vidéo d'Entendants en utilisant notre méthodologie. Cela nous permettrait d'avoir les mêmes données afin de comparer réellement les mouvements des éléments non manuels chez les sourds et les Entendants.

Cela permettrait en outre de connaître la part de mouvements spécifiques aux LS, la part de mouvements naturels de la communication en général et la part de mouvements influencés par les LV.

De même, une telle annotation permettrait de comparer la mobilité des sourcils chez les sourds et chez les Entendants et ainsi savoir si les phénomènes particuliers que nous avons noté sont propres aux sourds ou pas.

Enfin, il serait intéressant de comparer l'évolution des éléments non manuels dans les récits avec prise de rôle chez les Entendants et les mouvements des sourds lors des TP.

Validation des GNM lors d'un transfert

Afin de valider l'ensemble des observations concernant les clignements des yeux et des sourcils lors d'un transfert personnel, l'annotation devrait se poursuivre sur le corpus d'histoires de LS-COLIN où les transferts sont très présents, ou sur un autre corpus présentant les mêmes caractéristiques. Il serait ainsi également possible de continuer l'analyse du comportement des GNM lors d'autres transferts et en particulier d'observer leur évolution lors des transitions d'un type de transfert à un autre.

Validation des catégories de clignements

Il s'agit ici de valider les catégories existantes et peut-être d'en ajouter. Nous envisageons d'utiliser le corpus du passage à l'Euro et des recettes (LS-COLIN) pour cette nouvelle étude. En effet, le passage à l'Euro est, comme le 11 septembre, peu contraint dans sa forme, seul le sujet est imposé ce qui permet d'observer la présence des clignements en expression libre. Les recettes de cuisine sont plus contraintes et comportent des listes, des étapes dans la préparation, etc., tout en étant sur un mode d'expression explicatif. Ce corpus permettra d'observer en particulier les clignements de nouvelle idée, à savoir par exemple, entre deux étapes de préparation. De plus, ce type de discours nécessite la mise en valeur d'éléments du discours, comme par exemple la température du four, ou tout autre élément du discours sur lesquels il est nécessaire d'insister car indispensable à la recette.

2.1.2. Automatisation d'une partie de l'annotation

Dans l'idéal, un corpus pourrait être créé afin de positionner des capteurs ou tout du moins des marques visibles, aux endroits où nous positionnons nos points. Cela permettrait de disposer directement des coordonnées de chaque point.

Par automatisation partielle nous entendons soit utiliser les données recueillies par cette étude pour calibrer des logiciels de détection et suivi du visage afin que les points soient positionnés de manière automatique, soit, tout du moins, que les coordonnées des points génèrent l'annotation à l'aide de symboles. Pour ce faire, chacun des symboles a été défini de la façon suivante :

- La position de repos est propre à chaque signeur ainsi que le mouvement standard. Ce dernier est le mouvement le plus fréquent constaté chez un signeur.
- Les mouvements de faible amplitude mesurent deux fois moins qu'un mouvement standard.
- Les mouvements de forte amplitude mesurent le double d'un mouvement standard.
- Les tenues sont caractérisées par une immobilité d'au moins 40 ms.

Un nouveau symbole doit être généré à chaque fois que le mouvement est stoppé soit pour un maintien, soit pour un changement de mouvement (de baisse à hausse par exemple). Générer l'annotation symbolique implique d'élaborer un filtre qui devra éliminer une variation inférieure à 0,3 millimètres environ, qui représente le " bruit " de la mesure.

2.1.3. Précision pour l'annotation d'autres GNM

Le système de symboles que nous avons présenté est utilisable tel quel pour l'annotation de la bouche et des joues. Concernant la tête, nous avons créé d'autres symboles, basés sur les mêmes principes, afin de faciliter la lecture de l'annotation. La tête peut s'incliner, se tourner et se pencher. Elle évolue sur trois axes qui sont symbolisés dans le Tableau 12. Chacun de ces symboles permet également la distinction de trois amplitudes de mouvement.

Fonctions	Symboles
Avant / Arrière	
Gauche / Droite	
Inclinaison Gauche / Droite	

Tableau 12 Illustration des symboles utilisés pour la tête

Concernant le buste, nous avons utilisé les mouvements des épaules pour caractériser les mouvements de rotation du buste. En effet, si le buste est tourné sur la droite, nous constaterons que l'épaule gauche est reculée tandis que l'épaule droite est avancée. Le buste est donc uniquement annoté pour ses mouvements de bascule avant-arrière, grâce aux mêmes symboles que ceux utilisés pour les mouvements frontaux de la bouche.

Le positionnement des points est quant à lui envisagé en particulier pour la bouche où huit points seront positionnés : un à chaque extrémité, trois au milieu de la lèvre supérieure et trois au milieu de la lèvre inférieure. Concernant la tête, les joues et les épaules, il ne nous semble pas nécessaire à priori de disposer d'une telle précision.

Enfin, concernant le regard, nous avons imaginé représenter l'espace de signation dans un cube, représenté par la Figure 79, découpé en 27 cubes de taille égale. Il est ainsi possible de faire référence à l'endroit regardé en utilisant une lettre et un chiffre (comme dans une bataille navale). Nous n'avons toutefois pas testé ce système.

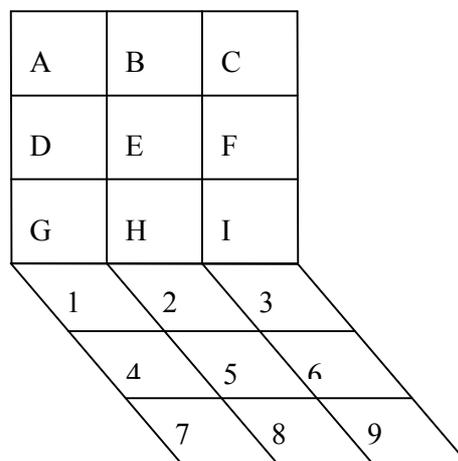


Figure 79 Représentation de l'espace de signation

2.2. Analyse

Une analyse ultérieure des nouvelles données permettra d'affiner les typologies proposées concernant aussi bien les clignements des yeux que les mouvements des sourcils. En plus de vérifier ou de préciser le nombre de catégorie, il s'agira également d'affiner les propriétés de chacune d'elles. Concernant les clignements, nous suggérons de vérifier en particulier la présence d'autres GNM. Concernant les sourcils, il s'agira en particulier d'observer les baisses dont l'étude n'a été que superficielle, ainsi que les mouvements de plus faible amplitude et l'interaction avec les mouvements des paupières que l'étude n'a pas abordée. Enfin, il s'agira de considérer l'amplitude de chacun des mouvements, pas suffisamment détaillée dans cette étude.

Nous avons débuté une étude sur le rapport entre la distance entre les sourcils et les mouvements des trois parties des sourcils. Celle-ci a permis de montrer que, de manière générale, la corrélation entre ces éléments était très forte. Cependant, elle a surtout permis de mettre en évidence certains phénomènes particuliers qu'il conviendrait d'analyser. Par exemple, une hausse des extrémités intérieures des sourcils engendre la plupart du temps une augmentation de la distance entre les sourcils. Cependant, dans de rares cas, cette distance diminue. Une étude ultérieure permettrait d'isoler l'ensemble de ces cas afin de les analyser.

De même, une première étude portant sur le lien entre l'extérieur et l'intérieur des sourcils a montré que l'extérieur n'était mobile que dans certains cas qu'il s'agira de définir avec précision. Enfin, concernant le milieu des sourcils, nous avons constaté que les mouvements du point central suivaient généralement l'intérieur des sourcils. Nous avons relevé uniquement deux cas où seul l'intérieur du sourcil était en mouvement et quatre cas où le milieu des sourcils suivait les mouvements de l'extérieur, et non pas l'intérieur, des sourcils. Les deux cas de non mobilité du milieu des sourcils étaient des cas de transfert personnel. Une étude ultérieure sur les transferts pourrait permettre de généraliser ou pas ces observations.

Nous avons également mené une première observation sur la relation entre la paupière inférieure et les mouvements des sourcils. Il semble que la paupière monte (donc yeux plissés) lorsque les sourcils sont rapprochés (d'où la présence de rides sur le front et entre les sourcils). La distance entre les sourcils serait donc corrélée aux mouvements des paupières inférieures.



Image 38 Paupière inférieure levée en même temps que la hausse et le rapprochement des sourcils.

Par ailleurs, nous avons commencé une analyse sur les mouvements de sourcils quantifiant et en particulier “grande quantité de, grand nombre de” (Image 39) suffisamment présent pour être analysés en détail. Cuxac caractérise cette expression par “gonflement des joues, plissement des yeux, souffle d’air en continu” (Cuxac, 2000 ; p.35). En examinant en détail cette production, il semble que la distance entre les sourcils diminue juste avant le signe ([PLUSIEURS] la plupart du temps dans notre corpus), afin d’atteindre une distance équivalente à la moitié de la distance minimale possible. Les yeux se plissent au début du signe. Les joues se gonflent au milieu du signe et tous les GNM sont maintenus pendant environ 160 ms (on notera une légère augmentation de la distance entre les sourcils durant le maintien). A la fin du maintien, la distance entre les sourcils se réduit encore et est maintenue jusqu’à la fin de l’énoncé concerné (à savoir ce qui est en “grand nombre de”). A noter également que l’extérieur des sourcils est autant en mouvement que l’intérieur et marque un léger froncement des sourcils. Ces observations demandent à être validées sur davantage de locuteurs.



Image 39 Illustration de “ grande quantité de ”, image extraite de notre corpus.

De même, nous avons relevé des GNM qui semblent être des qualificatifs, dont la traduction diverge et dont l'interprétation et la définition des caractéristiques nécessitent davantage d'analyses. Par exemple, pour les signes [TOUR] et [GRANDE TOUR] les GNM sont difficilement catégorisables et la distinction entre les deux reste problématique comme le montrent l'Image 40 et l'Image 41.



Image 40 [TOUR] chez trois de nos signeurs.



Image 41 [GRANDE TOUR] chez deux de nos signeurs.

En effet, si l'on observe l'Image 40 et l'Image 41, nous retrouvons de nombreuses similitudes quant à la bouche, l'inclinaison de la tête, le froncement des sourcils (qui est en réalité davantage une diminution de la distance entre les sourcils qui atteint presque son minimum qu'une réelle baisse des sourcils) et le plissement des yeux. La seule exception étant la dernière image qui se distingue complètement des autres.

Enfin, nous avons initié une analyse portant sur la façon dont les hausses étaient atteintes. Nous avons constaté que lorsque le signeur souhaitait accentuer progressivement ses propos, la hausse est atteinte par étape. Autrement dit, le signeur lève les sourcils, maintient, hausse à nouveau etc. Chacune des hausses correspondant à un argument supplémentaire. Ce phénomène demande à être confirmé et précisé.

2.3. Validation à l'aide de tests perceptifs

Pour valider nos observations, nous souhaitons procéder à des tests perceptifs. Pour ce faire, les GNM seront produits par un visage de synthèse. Ce visage devra posséder un certain nombre de propriétés que nous proposons de définir dans une première section. La seconde section liste les observations qu'il conviendra de valider.

2.3.1. Propriétés du visage de synthèse

Nous avons, comme déjà évoqué, procédé à des tests sur le visage synthétique de XFace. Cette phase nous a permis de mettre en avant qu'un visage synthétique, pour produire des GNM, devait posséder les propriétés suivantes (liste non exhaustive) :

- La distance entre les sourcils doit être modulable : chez les quatre signeurs de notre étude, la distance entre les sourcils varie du simple au double (ou presque) : 14 à 28 millimètres pour S1 ; 15 à 31 mm pour S2 ; 16 à 32 mm pour S3 et 19 à 35 mm pour S4.
- Une amplitude plus grande de l'intérieur des sourcils que de l'extérieur (comme déjà expliqué) en particulier en ce qui concerne la baisse des sourcils. En effet, chez nos quatre signeurs, l'amplitude de l'extérieur et de l'intérieur en cas de baisse varie jusqu'à être deux fois plus petite pour l'extérieur du sourcil.

- Des rides au front et entre les sourcils : Lors de notre étude, un des signeurs n'avait pas, ou presque pas, de ride sur le front ni entre les sourcils. Or, nous n'avons pas perçu la totalité des mouvements des sourcils. En effet, l'annotation de ce signeur a été celle qui a montré le plus de divergence entre l'annotation symbolique et l'annotation sur la vidéo. La première étant basée sur la perception, elle n'a pas relevé un grand nombre des mouvements des sourcils⁸².
- Le milieu des sourcils doit pouvoir suivre les mouvements de l'intérieur des sourcils. Après l'observation de nos données, il semble en effet que la mobilité du milieu dépende principalement de l'intérieur des sourcils. Cependant, il est également nécessaire que ces deux parties des sourcils ne soient pas dépendantes, afin de produire certaines hausses uniquement de l'intérieur des sourcils.
- Les paupières inférieures devront posséder une mobilité d'au moins 3 millimètres afin de distinguer les yeux plissés des yeux baissés (par exemple regard vers le bas).

Les trois derniers points cités ici sont problématiques sur le visage synthétique de XFace, ce qui engendre des difficultés de perception et des ambiguïtés.

2.3.2. Objectifs des tests de perception

Nous proposons ci-dessous une liste des différents éléments que nous souhaitons confirmer grâce aux tests de perception :

- L'emplacement et la structure (durée) de chacun des types de clignement.
- L'emplacement et la structure de chacun des types de mouvement des sourcils.
- L'interaction entre les mouvements des sourcils et les clignements des yeux.
- L'importance des deux éléments à savoir est-ce qu'un de ces éléments peut être facultatif dans certains cas ?

⁸² A noter également que ce corpus a été celui qui a posé le plus de problèmes de traduction à l'interprète. N'en connaissant pas les causes exactes, il serait intéressant de déterminer la part de l'absence de ride dans ces difficultés.

Pour ce faire, nous proposons de présenter des énoncés signés par un signeur virtuel en faisant varier les durées et les emplacements des sourcils et des clignements puis en les supprimant et de demander à des sujets sourds ce qui est perçu.

2.4. Résumé des perspectives

Nous avons proposé des perspectives de poursuite de cette étude aussi bien au niveau du choix des corpus à étudier qu'au niveau de l'extension à d'autres éléments non manuels tels que les joues et la tête. L'annotation étant lente et laborieuse, il semble pertinent de l'automatiser. La génération automatique de l'annotation qualitative à partir de l'annotation quantitative est rendue possible par la définition formelle de chacun des symboles utilisés. Ainsi, les coordonnées fournies par l'annotation quantitative permettent de générer un symbole de l'annotation qualitative.

L'annotation quantitative, quant à elle, peut être automatisée soit grâce à un logiciel de reconnaissance du visage suffisamment fin pour suivre des mouvements difficilement perceptibles (comme chaque point du sourcil), soit grâce à un corpus créé dans ce but, où des capteurs sont placés sur des points spécifiques du visage.

Enfin, il nous paraît particulièrement intéressant de procéder à la même annotation sur un corpus d'Entendants afin de comparer le rôle et les structures des clignements des yeux chez les sourds et chez les Entendants et de savoir si ces phénomènes sont propres aux LS ou si certains sont universels.

Références

1. Bibliographie

- AZNAR, G. ; DALLE, P. (2005). “ Variations dans la représentation écrite d'un signe en Signwriting ”. In actes de *TALN 2005*, 6-10 juin 2005, Dourdan
- BAHAN, B. (1996). *Non-Manual Realization of Agreement in American Sign Language*.
Doctoral dissertation, Boston University, Boston, MA.
- BAKER-SHENK, C. (1985). The facial behavior of deaf signers : évidence of a complex language. *Am Ann Deaf*, 130(4), 297-304.
- BAKER, C. (1979). NonManual components of the sign language signal. Paper presented at the NATO Advanced Study Institute, Copenhagen.
- BALCI, K. (2004). MPEG-4 based open source toolkit for 3d facial animation. In *AVI04, working conférence on Advanced Visual Interfaces*. Gallipoli, Italie, 25-28 mais 2004.
- BANGHAM, J.A ; COX, S. J. ; ELLIOTT, R. ; GLAUERT ; J.R.W ; MARSHALL, I ; RANKOV, S. ; WELLS, M. (2000). *Virtual Signing : Capture, Animation, Storage and Transmission - An Overview of the VisiCast Projet*, dans IEE Seminar on "Speech and language processing for disabled and elderly people", Londres, 2000
- BEBIAN, A. (1825). *Mimographie, ou essai d'écriture mimique, propre à régulariser le langage des sourds-muets*. L. Colas.
- BELL, A. G. (1883). *Memoir upon the formation of a deaf variety of the human race*. National Academy of Sciences. New Haven.
- BERGMAN, B. (1984). “Non-manual components in signed language: Some sentence types in Swedish Sign Language”. In Loncke F., Boyes-Braem P., Lebrun Y. (eds), p. 49-59, Lisse : Swets & Zeitlinger.

- BOLOT, L. ; BRAFFORT, A. ; FILHOL, M. (2006). *ELSI balbutie ! Vers une plateforme d'animation d'avatar signant*, WACA 2006, Toulouse, France.
- BOUALLEGUE, M. ; MARAOUI, M. ; MARS, M. ; ZRIGUI, M (2008). *Un système de génération et étiquetage automatique de dictionnaires linguistiques de l'arabe*. TALN-Recital 2008, Avignon.
- BOUILLON, P. ; VANDOOREN, F. (1998). *Traitement Automatique des langues naturelles*. Université francophones, De Boeck Université, 1998.
- BOUTORA, L (2008). *Problématique de la formalisation d'un niveau phonologique dans les langues signées : le cas de la configuration manuelle en Langue des Signes Française*. Thèse de Doctorat, Université Paris 8.
- BOUVET, D. (1992). Classification articulatoire des configurations de la main dans la Langue des Signes Française : Portée heuristique de cette classification pour la recherche des unités distinctives. *Protée*, 20(2-3), 23–32 ; 87–99.
- BOUVET, D. (1996). *Approche polyphonique d'un récit en Langue des Signes Française*. Ed. ARCI. Presses Universitaires de Lyon.
- BRAFFORT, A. ; DALLE, P. (2008). *Sign language applications: preliminary modelling*. Dans : International journal *Universal Access in the Information Society* (UAIS), Special issue 6/4 *Emerging Technologies for Deaf Accessibility in the Information Society* (Guest Editors: Eleni Efthimiou, Evita Fotinea, John Glauert), Springer.
- CHARBONNIER, C. (1995). *La commande oculaire : étude et validation expérimentale d'interfaces homme-machine contrôlées par la direction du regard*. Grenoble, France. Mémoire de doctorat Université Joseph Fourier/ LETI CEA.
- CHETELAT-PELE, E.; BRAFFORT, A. ; VERONIS, J. (2008). "Sign Language Corpus Annotation: Toward a New Methodology". In LREC 2008. Conference on Language Resources and Evaluation : Workshop Construction and Exploitation of Sign Language Corpora 2008 . Marrakesh.
- COERTS, J. (1992). *Nonmanual grammatical markers. An analysis of interrogatives, negations, and topicalisations in sign language of the Netherlands*. 234 p. Amsterdam University, Dissertation.

- COLLET, C. (1999). *Capture et suivi du regard par un système de vision dans le contexte de la communication homme-machine*. Mémoire de thèse, Doctorat de l'École Normale Supérieure de Cachan.
- COOPER, H M. ; BOWDEN R, (2009). *Sign Language Recognition: Working with Limited Corpora*. HCI (7) Universal Access in Human-Computer Interaction. Applications and Services, 2009: 472-481. DOI 10.1007/978-3-642-02713-0.
- CORINA, D.P. ; BELLUGI, U. ; REILLY, K. (1999). Neuropsychological studies of linguistic and affective facial expressions in deaf signers. *Language and Speech*, 2, 307-31.
- COULTER, G. R. (1978) "Raised eyebrows and wrinkled noses : The grammatical fonction of facial expression in relative clauses and related constructions", in F. Caccamise et D. Hicks (éd.), p. 65-74.
- COX, S. ; LINCOLN, M. ; TRYGGVASON, J. ; NAKISA, M. ; WELLS, M. ; TUTT, M. ; ABBOTT, S. (2002). *Tessa, a system to aid communication with deaf people*, dans 5th International ACM SIGCAPH Conference on Assistive Technologies (ASSETS 2002), 2002, Edimbourg.
- CUXAC, C. (1996). *Fonctions et structures de l'iconicité des langues des signes. Analyse descriptive d'un idiolecte parisien de la langue des signes française*. Université René Descartes - Paris V. Thèse de Doctorat d'Etat.
- CUXAC, C (2000). *La Langue des Signes Française : Les voies de l'iconicité*. Faits de Langues. Ophrys, Paris
- CUXAC C. (2002). "langage et cognition" , rapport de fin de recherche du projet ls-colin.
- CUXAC, C. (2004). Phonétique de la LSF : une formalisation problématique. *Silexicales*, 4, 93–113.
- DALLE-NAZEBI, S. ; LACHANCE, N. (2008). "sourds et médecine : impact des représentations sur les conditions d'accès aux soins. Regards croisés France-Québec" dans *?Interrogations?* n° 6.

- DELAIS-ROUSSARIE, E. ; G. CAELEN-HAUMONT ; D. HIRSt ; P. MARTIN et P. MERTENS (2006). *Outils d'aide à l'annotation prosodique de corpus*. In Bulletin PFC, numéro édité par A.C Simon et G. Caelen
- DELORME, M. FILHOL, A. BRAFFORT (2009). "An architecture for Sign Language Synthesis", Gesture Workshop, Bielefeld, Allemagne.
- DUBUISSON, C. (1996). "Rôle linguistique ou para-linguistique des sourcils dans les interrogatives en LSQ". In C. Dubuisson et D. Bouchard (éd), p47-58, Montréal : ACFAS.
- DUBUISSON, C. ; LELIEVRE, L. ; PARISOT, A.-M. ; RANCOURT, R. (1999). *Analyse du comportement non manuel "sourcils relevés" dans les interrogatives, les conditionnelles et les topicalisations en langue des signes québécoise (LSQ)*. In *Actes de l'ACL*. pp. 123-134.
- EFTHIMIOU, E. ; FOTINEA, S-E. ; VOGLER, C. ; HANKE, T. ; GLAUERT, J. ; BOWDEN, R. ; BRAFFORT, A. ; COLLET, C. ; MARAGOS P. ; SEGOUAT, J. (2009). Sign Language Recognition, Generation and Modelling : A Research Effort with Applications in Deaf Communication, HCI Int. 2009, juillet 2009, San Diego, USA.
- EKMAN P., FRIESEN W. V. (1978). Facial Action Coding System (FACS). Manuel Palo Alto : Consulting Psychologists Press.
- ELLIOTT, R. ;GLAUERT J.R.W. ; KENNAWAY, J.R. ; MARSHALL, I. (2000). *The development of language processing support for the ViSiCAST project*. In 4th International ACM SIGCAPH Conference on Assistive Technologies (ASSETS 2000), Washington, Novembre 2000
- ELLIOTT, R. ;GLAUERT J.R.W. ; KENNAWAY, J.R. (2004). *A Framework for Non-Manual Gestures in a Synthetic Signing System*. In actes de Cambridge Workshop on Universal Access and Assistive Technology (CWUAAT), March 2004, Cambridge, U.K.
- ESSER, C-L. (1831). *Mémoire sur les Fonctions des diverses parties de L'Organe Auditif*. Archives générales de médecine, Béchet et Migneret.

- FILHOL, M. (2009). *A descriptive model of signs for Sign Language modelling*, thesis abstract for the Sign Language and Linguistics journal vol. 12:1, direction J. Quer et R. Pfau, John Benjamins publishing, to be published.
- FILHOL, M. (2008). *Modèle descriptif des signes pour un traitement automatique des langues des signes*, PhD thesis, Université Paris-11 (Paris sud), Orsay.
- FUSELLIER I. ; BOUTORA L. (2005). "Travail contrastif sur les moyens d'annotation de corpus de LSF (partition et Sign Writing) visant l'analyse linguistique du domaine référentiel." Actes du colloque TALN-RECITAL'05, atelier Traitement Automatique des Langues des Signes, Dourdan, 6-10 juin.
- HANKE, T.; STORZ, J. (2008). "iLex : A Database Tool for Integrating Sign Language Corpus Linguistics and Sign Language Lexicography", 3rd Workshop on the Representation and Processing of Sign Languages, LREC 2008.
- JEMNI, M. ; ELGHOUL, O. ; MAKHLOUF, S. (2007). "A Web- Based Tool to Create Online Courses for Deaf Pupils", IMCL Conference, 18 -20, 2007, Amman, Jordan.
- JOHNSON, R. E. ; LIDDELL, S. K. (in preparation) *Sign Language Phonetics: Architecture and Description*
- JOUISSON, P. (1995). *Ecrits sur la langue des signes française*. L'Harmattan. Edité par B. Garcia.
- KENNAWAY, R. (2001). Synthetic Animation of Deaf Signing Gestures. Gesture Workshop 2001: 146-157.
- KENNAWAY, R. (2003). Experience with and Requirements for a Gesture Description Language for Synthetic Animation. Gesture Workshop 2003: 300-311.
- KIPP M. (2004). "Gesture Generation by Imitation - From Human Behavior to Computer Character Animation." Boca Raton Florida: Dissertation.com
- KLIMA, E. ; BELLUGI, U. (1979). *The Signs of Language*. Cambridge, Massachusetts, London, England : Harvard University Press.
- KOLOD, E. (2004). "How does Learning sign language affect perception ?". Intel Science Talent Search.

- LACERTE, L. ; PINSONNEAULT, D. ; FOURNIER, R. (1989). “ Le signe non manuel en LSQ ”. In actes du colloque de l'Acfas, Université du Québec à Montréal.
- LANE, H. (1979). *Histoire Chronologique de la répression de la langue des signes en France et aux états unis*. Northeastern University Volume 13, Numéro 56. Pp92-124.
- LEFEBVRE-ALBARET, F. ; DALLE, P. (2009). *Analyse de vidéo en langue des signes : méthodes et stratégies*. Dans : *ORASIS, Trégastel, 08/06/2009-12/06/2009*, 2009
- LEON, J. (2001). “ Conceptions du mot et débuts de la traduction automatique “, *Histoire Épistémologie Langage*, vol.23-1 : 81-106.
- LEON, S. (2009). *Un système modulaire d'acquisition automatique de traductions à partir du Web*. Conférence sur le Traitement Automatique des Langues Naturelles 2009, Senlis.
- LEON, S. (2008). *Acquisition automatique de traductions d'unités lexicales complexes à partir du Web*. PhD thesis, Université de Provence, 2008.
- LEUPRECHT, P. (1999). “ Médecine et droits de l'homme “, dans Association pédagogique nationale pour l'enseignement de la thérapeutique et Réseau européen des enseignants de thérapeutique, *Ethique et thérapeutique*, J.-M. Mantz, P Grandmottet et P Queneau eds., Presses universitaires, Strasbourg, pp. 485-490.
- LIDDELL, S. (1980). *American Sign Language Syntax*. The Hague : Mouton.
- LIDDELL, S. ; JOHNSON, R. (1989). American sign language : the phonological base. *Sign Language Studies*, **64**, 195–277.
- LIDDELL, S. (2003). *Grammar, gesture and meaning in American Sign Language*. Cambridge University Press, Cambridge.
- MACFARLANE, J. (1998). “From affect to grammar: Ritualization of facial affect in signed languages”. Paper presented at the *6th International Conference on Theoretical Issues in Sign Language Research*, Washington, DC.
- MARTIN-DUPONT, X. (1995). “ Les modalités d'évaluation objective dans le domaine de la communication non verbale ”. Notes and documents. Orsay : LIMSI/CNRS.

- MARTINET, A. (1960). *Éléments de linguistique générale*, Paris, Armand Colin.
- MCCLAVE, E. (2002). "Non-manual Gestures in American Sign Language". In *Gesture: the living Medium*, June 5-8. 2002, Austin, TX ?
- MEDIN, N. ; SALLANDRE, M. – A. (2006). "Typologie des procès dans les récits en Langue des Signes Française". *Colloque International : syntaxe, interprétation, lexiques des langues signées*. 1er et 2 juin 2006, Lille.
- MERTENS, P. (2004). *Un outil pour la transcription de la prosodie dans les corpus oraux*. In *Traitement Automatique des langues* 45 (2), 109-130.
- MILLET, A. (1997). Réflexions sur le statut du mouvement dans les langues gestuelles : aspects lexicaux et syntaxiques. *Lidil*, **15**, 11–30. Grenoble, Lidilem.
- MILLET, A. (1998). Typologie des signes et structuration du lexique en LSF. Réflexions autour de la notion d'Unité Linguistique Intermédiaire. In *Oralité et Gestualité. Communication multimodale, interaction*, pp. 95–100, Paris : L'Harmattan.
- MINISTERE DE L'INSTRUCTION PUBLIQUE, (1881). *Compte-rendu du congrès international pour l'amélioration du sort des sourds-muets tenu à Milan du 6 au 11 septembre 1880*. Rome, Imprimerie Héritiers Botta.
- MOODY, B. (1986). *La langue des signes - Tome 2: dictionnaire bilingue élémentaire*. International Visual Theatre (I.V.T.). Editions Ellipses. Paris.
- MOODY, B. (1990). *La langue des signes - Tome 3: dictionnaire bilingue élémentaire*. International Visual Theatre (I.V.T.). Editions Ellipses. Paris.
- MUNOT, P. ; NEVE, F.-X. (2002). *Une introduction à la phonétique: manuel à l'intention des linguistes, orthophonistes et logopèdes*, Editions du CEFAL, 2002, Volume 9 de CEFAL sup.
- NEIDLE, C. ; KEGL, J. ; MACLAUGHLIN, D. ; BAHAN, B. ; LEE, R. G. (2000). *The Syntax of American Sign Languages – Functional Categories and Hierarchical Structure*. MIT Press.

- PELACHAUD, C. ; BRAFFORT, A. ; BRETON, G. ; ECH CHADAI, N. ; GIBET, S. ; MARTIN, J.-C. ; MAUBERT, S. ; OCHS, M. ; PELE, D. ; PERRIN, A. ; RAYNAL, M. ; REVERET, L. ; SADEK, D. (2004). "AGENTS CONVERSATIONELS : Systèmes d'animation Modélisation des comportements multimodaux Applications : agents pédagogiques et agents signeurs." Action Spécifique du CNRS Humain Virtuel.
- PIERRE, J. (2007). *La langue au cœur du numérique. Les enjeux culturels des technologies de la langue*. Délégation générale à la langue française et aux langues de France, 2007.
- PRILLWITZ, S. ; LEVEN, R. ; ZIENERT, H. ; HANKE, T. ; HENNING, J. (1989). *HamNoSys Version 2.0, Hamburg Notation System for Sign Languages, an Introductory Guide*. Hamburg : Signum Press.
- REGNARD, A. (1902). *Contribution à l'histoire de l'enseignement des sourds-muets*, Larose, Paris.
- RICHET, C. (1919). *La sélection humaine*, éditions F. Alcan, Paris.
- SALLANDRE, M. - A. (2003). Les unités du discours en Langue des Signes Française. Tentative de catégorisation dans le cadre d'une grammaire de l'iconicité. Thèse de Doctorat, Université Paris 8.
- SALLANDRE, M. - A. (2005). " Les va et vient de l'iconicité en langue des signes française ". Acquisition et Interaction en Langue étrangère. Les Langues des Signes : une perspective sémiogénétique.
- SANDLER, W. (1986). " The spreading hand autosegment of American Sign Language ". *Sign Language Studies*, **50**, 1–28.
- SAUSSURE, F. de (1972 [1916]) : *Cours de linguistique générale*, édition critique préparée par Tullio De Mauro, Payot. p. 13-43, 155-184, Paris.
- SEGOUAT, J. (2009). " A study of sign language coarticulation ". *In Sigaccess Newsletter*, " Accessibility and computing ". pp. 31-38, Issur 93.
- STOKOE, W. C. (1960) : *Sign Language Structure*. Studies in Linguistics. Occasional Papers 8. Buffalo, NY: University of Buffalo Press.
- STOKOE, W. C. (1972). *Semiotics and human sign languages*. The Hague: Mouton.

- SUTTON, V. (2002). "Lessons in SignWriting, Textbook & Workbook", 3rd ed., The Deaf Action Committee for SignWriting.
- SUTTON, V.; GLEAVES, R. (1995). *SignWriter – The world's first sign language processor*. Deaf Action Committee for SignWriting, La Jolla, CA.
- VERGE, F. (2001). *Le regard en Langue des signes française*. Thèse de doctorat de Sciences du Langage, Université de Toulouse-le-Mirail.
- VERLAINE, L.; GUITTENY, P.; LEGOUIS, P. *La Langue des signes*. Manuel du centre d'information sur la surdité d'Aquitaine.
- VERLINDEN, M. ; TIJSSELING, C. ; FROWEIN, H. (2001). *A signing avatar on the WWW*, dans International Gesture Workshop, Londres, avril 2001
- VOGT-SVENSEN, M. (1990) "Eye gaze in Norwegian Sign Language interrogatives, in W. H. Edmondson et F. Karlsson (éd.), p. 153-162, Hamburg : Signum-Press.

2. Webographie :

2.1. Textes :

Aristote (“ Opuscules, plan de traité de la sensation et des choses sensibles ”), traduction de J. Barthélémy Saint Hilaire.

remacle.org/bloodwolf/philosophes/Aristote/tableopuscules.htm

Code Justinien, Livre 1.

www.cis.gouv.fr/spip.php?article1530

Compte-rendu du congrès de l’UNISDA, 8 octobre 2005.

etsf.fnsf.free.fr/informationannonce2.htm

Conférence d’octobre 2007, à l’URAPEDA de Bretagne, de Cécilia Hutter, doctorante en linguistique à l’Université de Rouen.

www.unapeda.asso.fr/article.php3?id_article=776

Discours de Jack Lang du 13 février 2002.

ftp://trf.education.gouv.fr/pub/edutel/actu/2002/02_13_discours_lang_languesignes.pdf

Kaiser S. (2006). “Les expressions faciales”. Support de cours de l’Université de Genève.

dokéos.unige.ch/courses/7217F/document/Historique.pdf?cidReq=7217F⁸³

Loi du 11 février 2005 sur l’égalité des droits et des chances.

www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000809647&dateTexte=

Manuel HamNoSys de la version 4.0]

www.sign-lang.uni-hamburg.de/projekte/hamnosys/hns4.0/hns4.0eng/contents.html

Rapport Gillot, 30 juin 1998.

www.ladocumentationfrancaise.fr/rapports-publics/984001595/index.shtml

Trésor de la Langue Française Informatisé.

atilf.atilf.fr/tlf.htm

⁸³ Le site n’est plus disponible au moment de l’impression de ce manuscrit, nous ne savons pas s’il s’agit d’un problème temporaire ou si le document a été retiré.

2.2. Outils et ressources :

Anvil: <http://www.dfki.de/~kipp/anvil>

Artnatomia : www.artnatomia.net

Deaf Connection : www.deafconnexions.org.uk/dev/

Dicta-Sign : www.dictasign.eu

Elan : www.mpi.nl/tools/elan.html

ESign : www.visicast.sys.uea.ac.uk/eSIGN/index.html

FACS : www.face-and-emotion.com/dataface/facs/manual/TitlePage.html

HamNoSys: www.sign-lang.uni-hambourg.de/projects/HamNoSys.html

Ilex : www.sign-lang.uni-hambourg.de/ilex/

Scilab : www.scilab.org/

SignStream : bu.edu/asllrp/SignStream/

SignWriting : www.signbank.org

VISICAST : www.visicast.co.uk

Websourd : www.websourd.org

XFace : xface.itc.it/

Illustrations

1. Les Images

Image 1	[LIT] _____	16
Image 2	[ECOSSE] _____	17
Image 3	[ESCARGOT] _____	17
Image 4	[ARRETER] _____	17
Image 5	[ARBRE] _____	17
Image 6	Transfert personnel imagé (B. Moody & al. (IVT), 1983, p.79, 80). _____	20
Image 7	[LAQUE] _____	23
Image 8	[VERRE] - [CONSERVE] _____	24
Image 9	[CHOCOLAT] - [VIDE] _____	24
Image 10	[CONTENT] - [MAL AU CŒUR] _____	24
Image 11	[FONDER] - [PUNIR]. Images extraites de (Verlaine) _____	25
Image 12	[APERCEVOIR] - [ATTRAPER]. Images extraites de (Verlaine) _____	25
Image 13	[PIPE] - [FUMER LA PIPE] _____	25
Image 14	[NUIT] - [LA NUIT TOMBE] _____	26
Image 15	TESSA _____	39
Image 16	GUIDO dans deaf connexion _____	40
Image 17	Extrait de la seconde planche de Bébien : « Caractéristique des diverses parties de la tête et du corps » _____	45
Image 18	Secousse de la tête - Hausse de l'épaule gauche _____	47
Image 19	Avant la production : tête droite; Pendant la production : tête tournée; Après la production : tête droite. _____	49
Image 20	[TRISTE] sur le site "LSF, dictionnaire et + " _____	50
Image 21	Exemple de deux hausses de sourcils différentes. _____	55
Image 22	Illustration du logiciel Elan _____	68
Image 23	Illustration du logiciel Ilex. _____	69
Image 24	Interface d'Anvil _____	70
Image 25	Yeux plissés - Regard vers le bas (Images extraites de notre corpus) _____	84
Image 26	Distance entre les sourcils au repos (d) _____	117
Image 27	Distance entre les sourcils maximale (d+5) _____	117
Image 28	Distance entre les sourcils minimale (d-5) _____	117
Image 29	Image correspondante au clignement de l'annotation de la Figure 71 _____	135

Image 30	Illustration du clignement lors de [L'AVION QUI PERCUTE LA TOUR].	137
Image 31	Illustration du clignement lors de [L'AVION QUI PERCUTE LA TOUR].	137
Image 32	Illustration du clignement lors de [L'AVION QUI PERCUTE] sans le bras de la tour.	137
Image 33	Illustration du clignement de répétition. Ici [LES CAMERAS FILMENT TOUT AZIMUT].	138
Image 34	Illustration du clignement de répétition. Ici [ILS PARTENT].	138
Image 35	Illustration du signe [MOI] accompagné d'un clignement.	138
Image 36	Illustration du clignement personnel [JE CONNAIS]	139
Image 37	Illustration du clignement accompagnant une émotion	139
Image 38	Paupière inférieure levée en même temps que la hausse et le rapprochement des sourcils.	163
Image 39	Illustration de " grande quantité de ", image extraite de notre corpus.	164
Image 40	[TOUR] chez trois de nos signeurs.	164
Image 41	[GRANDE TOUR] chez deux de nos signeurs.	164

2. Les figures

Figure 1	Description paramétrique du signe [MUR]	15
Figure 2	rôle du regard lors d'un Transfert personnel	29
Figure 3	Organisation des travaux de M&TALS	41
Figure 4	Résultat d'une transcription de D'Sign (Jouison, 1995, pp. 237)	45
Figure 5	HamNoSys : [WHAT] dans l'énoncé " What quote three bears ? "	46
Figure 6	Extrait des symboles de localisation sur le visage d'HamNoSys	47
Figure 7	Ecriture de [LOI] en SW.- illustration de [LOI]	48
Figure 8	Ecriture de [TRISTE] en SW	49
Figure 9	: Symboles SW pour désigner les postures possibles des sourcils et le front	50
Figure 10	Extrait des symboles de SW utilisés pour désigner l'emplacement de la langue	51
Figure 11	Extrait des symboles de pour désigner les formes que les yeux peuvent adopter	51
Figure 12	Expression du dégoût, analyse de FACS	53
Figure 13	postures contenues dans un GNM	56
Figure 14	Schéma d'une hausse de sourcils normale (traits pointillés) et d'une hausse interrompue suivie d'un froncement (traits pleins)	57
Figure 15	Deux hausses de sourcils à temps de réalisation variable qui peuvent exprimer l'incrédulité (à gauche) et la surprise (à droite)	57
Figure 16	Détail de la transcription de [MONTAGNE] d'après (Martin Dupond, 1995).	58
Figure 17	HamNoSys : tête tournée	58
Figure 18	Extrait de la partition de Cuxac p.352	64
Figure 19	Extrait de la partition de Bouvet p.66	65
Figure 20	Extrait de la partition	66

Figure 21 Sous-partie de la partition Sign Stream	67
Figure 22 Plan des caméras	78
Figure 23 Interface d'Anvil	80
Figure 24 Extrait de l'annotation	82
Figure 25 Les mouvements compris dans un GNM	83
Figure 26 Décomposition d'une diagonale	85
Figure 27 Structure des pistes de l'annotation	86
Figure 28 Symboles les plus fréquents	87
Figure 29 Amplitudes de réalisation	87
Figure 30 Phases de réalisation d'une hausse sans contact suivie d'une tenue puis d'un retour à la posture initiale	88
Figure 31 Extrait de notre annotation	89
Figure 32 Extrait de notre transcription du corpus LS-COLIN	92
Figure 33 Illustration de l'annotation des joues, extraite de notre transcription du corpus LS-COLIN.	92
Figure 34 Illustration de l'annotation des épaules, extraite de notre transcription du corpus LS-COLIN.	93
Figure 35 Muscle frontal et ses AUs associées (hausse de l'intérieur (AU 1) et de l'extérieur des sourcils (AU 2)). Images extraites d'Artnatomy (Contreras Flores, 2005) et du manuel de FACS.	95
Figure 36 Muscles sourcilier (image A), orbiculaire des paupières (image B) et pyramidale (image C) et leurs AUs associées : rapprochement des sourcils (AU 4). Images extraites d'Artnatomy et du manuel de FACS.	95
Figure 37 Combinaison d'Aus : Image A : AU 1 + AU 4 ; Image B : AU 1 + AU 2 ; Image C : AU 1 + AU 2 + AU 4	96
Figure 38 Emplacement des 18 points	97
Figure 39 Image extrait de la vidéo avec les 18 points.	97
Figure 40 Données brutes en entrée : coordonnées des 18 points d'une image	98
Figure 41 Calcul des moyennes	98
Figure 42 Calcul de la distance entre les points des sourcils et les points des yeux	99
Figure 43 Données en sortie : positions des 6 points des sourcils d'une image	100
Figure 44 Position neutre	101
Figure 45 Diminution de la distance entre les sourcils et baisse de la paupière supérieure.	101
Figure 46 Positions de l'intérieur du sourcil droit.	102
Figure 47 Extrait de l'annotation	110
Figure 48 Illustration d'un début de hausse et son amplitude maximale durant un même signe.	116
Figure 49 Illustration d'un mouvement de sourcil qui segmente un énoncé	118
Figure 50 Illustration d'une hausse concessive	119
Figure 51 Illustration d'une hausse concessive introduite par un interrogatif après une pause.	120
Figure 52 Illustration du mouvement des sourcils juste avant le rhème.	121

<i>Figure 53 Illustration des mouvements des sourcils lors d'un changement personnage dans un TP. (Signeur 1).</i>	122
<i>Figure 54 Illustration d'une problématisation.</i>	123
<i>Figure 55 Illustration d'une hausse en fin d'interrogatif, durant toute la réponse.</i>	124
<i>Figure 56 Illustration d'une hausse en fin d'interrogatif, non maintenue durant toute la réponse.</i>	125
<i>Figure 57 Illustration d'une hausse de fin d'interrogatif avec clignement.</i>	125
<i>Figure 58 Illustration d'une question lors d'un transfert.</i>	126
<i>Figure 59 Répartitions des durées de la première phase des clignements (fermeture) pour l'ensemble du corpus.</i>	128
<i>Figure 60 Répartitions des durées de la seconde phase des clignements (ouverture) pour l'ensemble du corpus.</i>	128
<i>Figure 61 Répartition des durées totales (2 phases comprises) des clignements du corpus.</i>	129
<i>Figure 62 Étendue des clignements dans le temps chez le signeur S1.</i>	129
<i>Figure 63 Répartition de l'ensemble des clignements</i>	130
<i>Figure 64 Illustration d'un clignement de pré-prise de parole.</i>	131
<i>Figure 65 Illustration d'un clignement de délimitation de groupes syntaxiques.</i>	131
<i>Figure 66 Illustration de clignements de délimitation de pauses.</i>	132
<i>Figure 67 Illustration du clignement de segmentation d'énoncé avec mouvement des sourcils.</i>	132
<i>Figure 68 Illustration du clignement de délimitation de fin de réflexion et de nouvel énoncé.</i>	133
<i>Figure 69 Illustration d'un clignement de mise en valeur du rhème.</i>	134
<i>Figure 70 Illustration d'un clignement de mise en valeur du rhème.</i>	134
<i>Figure 71 Illustration d'une mise en valeur de l'objet pointé, extraite de notre annotation.</i>	135
<i>Figure 72 Illustration de la mise en valeur par isolement.</i>	135
<i>Figure 73 Illustration de clignements liés aux fausses questions.</i>	139
<i>Figure 74 Répartition des types de clignements pour chaque signeur.</i>	140
<i>Figure 75 Postures clés et transitions lors d'une hausse de sourcil</i>	144
<i>Figure 76 Postures clés et transitions lors d'un clignement</i>	145
<i>Figure 77 Illustrations des symboles utilisés pour synchroniser nos postures clés à celles des signes</i>	146
<i>Figure 78 Représentation du clignement de segmentation de groupe syntaxique.</i>	154
<i>Figure 79 Représentation de l'espace de signation</i>	161

Remerciements

Je tiens à remercier sincèrement :

Mes directeurs de thèse qui, face à mes contraintes et mes choix, m'ont accompagnée et soutenue durant ces trois ans (et plus) avec un œil confiant et bienveillant : Jean Véronis, pour m'avoir poussée dès le début, avoir su avant moi combien cela me plairait et pour sa qualité d'orateur qui m'a convaincue de me lancer, pour sa détermination afin de m'obtenir le nécessaire et plus encore, mais également pour sa confiance et son soutien qui m'ont ouvert de nouveaux horizons... Annelies Braffort qui m'a encadrée, écoutée, conseillée, guidée, motivée, recadrée, boostée et remotivée à chaque pas, chaque jour de cette thèse. Mais également pour sa génération automatique d'humour, de patience et d'encouragements; pour m'avoir appris à « poser mes étiquettes » et mes idées de façon simple; pour ne pas être restée sur une « position figée » mais s'être complètement adaptée à mon « rythme » de travail et mon fonctionnement (interne ?) lors de son fantastique encadrement dont la « mécanique (du mouvement ?) » est si bien huilée; pour avoir su animer nos échanges (en 3D ?) pour que la motivation ne tombe pas ; pour avoir toujours su « harmoniser » les différences afin que chacun trouve sa place ; pour la valeur de sa collaboration, ses nombreuses relectures critiques qui, au delà du retour formel, étaient pleines d'humour, je ne risque pas d'oublier !

Christian Cuxac pour s'être rendu disponible, malgré les contraintes, pour sa bienveillance, pour l'attention qu'il a portée à mon manuscrit, pour la passion dans sa voix lorsqu'il parle de LSF qui est incroyablement contagieuse et pour la patience dont il a fait preuve dans l'élaboration de la liste des manques...

Jean-Marc Colletta pour sa réactivité, sa bienveillance, son professionnalisme, sa compréhension et pour s'être souvenu combien il est difficile d'attendre dans l'ignorance.

Patrice Dalle pour m'avoir encouragée dans la communication scientifique (Toulouse 2007) en m'apportant confiance et conseils. Un grand merci également pour sa compréhension et sa patience qui ont été mises à dure épreuve.

L'équipe des niolus M&TALS du LIMSI qui m'ont offert un cadre de travail idéal, pour les échanges toujours pertinents et source de motivation. Pour les gestes discrets mais toujours attentionnés d'Annick, les regards doux et le sourire apaisant de Laurence, l'humour de Max qui tombe toujours à pic, les suggestions pertinentes de Micka, le soutien discret et bénéfique de Jérémie, la gentillesse de Cyril et la bonne humeur de tous... Ils m'ont montrée que malgré la distance avec l'équipe M&TALS, c'est possible !

L'équipe du CILSH, pour leur disponibilité, leur soutien et leurs conseils. Pour avoir su rire de mes multiples péripéties et les avoir partagées. Tout le monde a vibré avec moi lorsqu'on a entendu la sonnerie... merci en particulier à Marie-Thérèse Ponsonnet et Jean-Luc Peris qui a répondu présent à chacune de mes sollicitations et toujours avec enthousiasme.

Les générateurs de soutien et de bons conseils spécialisés dans les moments critiques à qui je veux adresser une mention spéciale : Stéphanie Léon, avec qui j'ai eu l'honneur de partager tant de choses, qui m'a soutenue et encouragée et dont les regards complices m'ont bien souvent plus apporté que n'importe quel mot ; Alice Carne, qui depuis longtemps déjà m'accompagne et dont la présence a toujours été précieuse, les coups de mains généreux et sans limite et les relectures profondes et attentives ; Christophe Mathieu, l'oreille attentive, toujours présent pour conseiller ou faire rire et dont les réponses sont tellement réconfortantes ; Xavier Tannier, vrai générateur de crise de rire, qui a vraiment subi toutes mes péripéties, toujours présent, toujours disponible et toujours de bons conseils ; et Sonia Testa, sans qui clairement la soutenance n'aurait pas eue lieu, son soutien, sa souplesse et sa gentillesse ont été précieux.

Toutes les personnes qui de près ou de loin m'ont aidée à poser le point final de cet ouvrage : Colette (pour ses encouragements répétés, son soutien et sa compréhension), Gaëlle (pour avoir été si présente malgré la distance), Monsieur Goffredo (pour ses mouvements circulaires qui ont rendu supportable toutes ces heures assises), Stéphanie, Guillaume, Lorraine et Alex (pour leur soutien et leur confiance), Patricia, FX, Thomas Lebarbé et Virginie Zampa (que de conseils et de fous rires), Loïc Kervajan, Alexia, Annabelle Rodriguez (pour son implication à tous les niveaux, son regard réconfortant, sa présence et son amitié précieuse), et la grande rencontre de la dernière ligne droite : Leïla Boutora. Un grand merci également à Claire Maury-Rouen, à qui je dois ma venue à Aix, qui m'a accompagnée et encouragée depuis le début, et à Estelle Véronis qui, au delà de la thèse, m'a offert un cadre et un confort de travail adaptés à mes besoins et pour avoir considéré l'ensemble de mes projets avec bienveillance en m'apportant toujours son soutien.

Enfin, je remercie tout particulièrement Franck pour sa patience sans limite, son inestimable soutien, mais également pour ses attentions de chaque jour qui m'ont permis de traverser toutes les épreuves sereinement et avec le sourire en particulier dans la dernière ligne droite où il a été source de tellement d'apaisement et de confiance.

Les remerciements sont comme le monde, il est difficile d'en faire le tour, pourtant même si certaines personnes n'ont pas été nommées elles sauront ici qu'au même titre que les personnes citées, je les remercie d'avoir apporté leur contribution et leur appui à ce projet.

Annexes

1. Formalisation des catégories de clignements des yeux

1.1. Le clignement Pré-prise de parole

(KPC1) KEY_POSTURE

Sync

0S

{

P : 0

}

(TC1) TRANSITION (2) {

}

(KPC2) KEY_POSTURE (1)

Sync

0S

{

P : contact

}

(TC2) TRANSITION (3/4){

}

(KPC3) KEY_POSTURE

Sync

0S

{

P : 0

}

1.2. Le clignement de segmentation de groupes syntaxiques

(KPC1) KEY_POSTURE

Sync

PSf

{

P : 0

}

(TC1) TRANSITION (2){

}

(KPC2) KEY_POSTURE (1) {

P : contact

}

(TC2) TRANSITION (2){

}

(KPC3) KEY_POSTURE

Sync

FSP

{

P : 0

}

1.3.Les clignements de pauses

1^{er} clignement :

(KPC1) KEY_POSTURE
Sync
FSP
{
P : 0
}
(TC1) TRANSITION (3) {
}
(KPC2) KEY_POSTURE (1)
Sync
0S
{
P : contact
}
(TC2) TRANSITION (3) {
}
(KPC3) KEY_POSTURE {
P : 0
}

2^{ème} clignement

(KPC1) KEY_POSTURE
Sync
0S
{
P : 0
}
(TC1) TRANSITION (2) {
}
(KPC2) KEY_POSTURE (1)
Sync
0S
{
P : contact
}
(TC2) TRANSITION (2) {
}
(KPC3) KEY_POSTURE
Sync
DSS
{
P : 0
}

1.4. Le clignement de mise en évidence du rhème :

(KPC1) KEY_POSTURE

Sync

FSP

{

P : 0

}

(TC1) TRANSITION {

}

(KPC2) KEY_POSTURE (1) {

P : contact

}

(TC2) TRANSITION {

}

(KPC3) KEY_POSTURE {

P : 0

}

1.5. Le clignement lié à l'objet pointé

1^{er} cas

```
(KPC1) KEY_POSTURE
Sync
PSf
{
P : 0
}
(TC1) TRANSITION (3) {
}
(KPC2) KEY_POSTURE (1)
FPS
{
P : contact
}
(TC2) TRANSITION (2) {
}
(KPC3) KEY_POSTURE
Sync
0S
{
P : 0
}
```

2^{ème} cas :

```
(KPC1) KEY_POSTURE
Sync
FPS
{
P : 0
}
(TC1) TRANSITION (3) {
}
(KPC2) KEY_POSTURE (1)
Sync
0S
{
P : contact
}
(TC2) TRANSITION (2) {
}
(KPC3) KEY_POSTURE
Sync
0S
{
P : 0
}
```

1.6. Le clignement lié aux fausses questions

(KPC1) KEY_POSTURE

Sync

PSf

{

P : 0

}

(TC1) TRANSITION (3) {

}

(KPC2) KEY_POSTURE (1)

Sync

OS

{

P : contact

}

(TC2) TRANSITION (2) {

}

(KPC3) KEY_POSTURE

Sync

DSS

{

P : 0

}

2. Les mouvements des sourcils

2.1. Le concessif de milieu d'énoncé

(KPH-1) KEY_POSTURE

Sync

DPS

{

S : 0

}

(TH-1) TRANSITION {

}

(KPH-2) KEY_POSTURE (1) {

S+++

}

(TH-2) TRANSITION {

}

(KPH-3) KEY_POSTURE {

S : +

}

(TH-3) TRANSITION {

}

(KPH-4) KEY_POSTURE {

S : 0

}

2.2. Le concessif de début d'énoncé

(KPH-1) KEY_POSTURE

Sync

DPS

{

S : 0

}

(TH-1) TRANSITION {

}

(KPH-2) KEY_POSTURE (1) {

S+++

}

(TH-2) TRANSITION {

}

(KPH-3) KEY_POSTURE {

S : ++

}

(TH-3) TRANSITION {

}

(KPH-4) KEY_POSTURE {

S : 0

}

2.3. Le concessif sur tout l'énoncé

(KPH1) KEY_POSTURE

Sync

PSf

{

S : 0

}

(TH1) TRANSITION {

}

(KPH2) KEY_POSTURE (1)

Sync

DSS

{

S : +++

}

(TH2) TRANSITION (FE) {

}

(KPH3) KEY_POSTURE {

S : 0

}

2.4.La mise en évidence du rhème

(KPH1) KEY_POSTURE

Sync

FPS

{

S : 0

}

(TH1) TRANSITION {

}

(KPH2) KEY_POSTURE (FE)

Sync

DSS

{

S : ++

}

(TH2) TRANSITION{

}

(KPH3) KEY_POSTURE {

S : 0

}

2.5.L'objet pointé

(KPH1) KEY_POSTURE

Sync

OS

{

S : 0

}

(TH1) TRANSITION {

}

(KPH2) KEY_POSTURE (FS)

Sync

DSP

{

S : +++

}

(TH2) TRANSITION {

}

(KPH3) KEY_POSTURE {

S : 0

}

Résumé : Cette thèse s'inscrit dans le cadre de la génération automatique de Langue des Signes Française et plus particulièrement la génération des Gestes Non Manuels (GNM). La génération nécessite une description précise des GNM. Actuellement, nous ne disposons que de notations de nature symbolique. Nous proposons une nouvelle méthodologie d'annotation permettant une description fine et précise des mouvements, tout en tenant compte de la structure temporelle des GNM. Cette méthodologie comporte une annotation qualitative et une annotation quantitative. Nous avons appliqué notre méthodologie aux mouvements des sourcils et aux clignements des yeux et obtenons ainsi des données numériques et symboliques à partir desquelles nous menons l'analyse. Celle-ci permet de dégager une typologie de ces mouvements. Pour chaque catégorie nous précisons son rôle, l'emplacement du mouvement par rapport aux signes avoisinants et sa structure. Nous formalisons chacune des catégories pour une génération ultérieure et proposons des perspectives de poursuites de l'étude.

Mots clés : Langue des Signes Française ; Gestes Non Manuels ; Génération automatique ; Sourcils ; Clignements des yeux ; Formalisation ; Annotation.

Abstract : These thesis tackles Non Manual Gestures (NMGs) annotation within the context of Sign Language (SL) research and more particularly within the context of automatic generation of NMGs. The generation requires a precise description of NMGs. At the present, only symbolic notations are available. Thus, we propose a new annotation method, allowing precise movements descriptions taking into account the temporal structure of NMGs. This method involves qualitative as well as quantitative annotations and was used to get numerical and symbolical data for analysis on eyebrows movements and eyes blinking. Analysis of this data led to define a typology of these movements. Categories are defined in terms of role, the place of the movement with the surrounding signs and the structure. These categories are formalized in order to allow their generation.

Keywords : French Sign Language ; Non Manual Gestures ; Automatic génération ; Eyebrows ; Blinking ; Formalization ; Annotation.