



HAL
open science

Contribution à l'analyse des mouvements 3D de la Langue des Signes Française (LSF) en Action et en Perception

Mohamed-El-Fatah Benchiheub

► **To cite this version:**

Mohamed-El-Fatah Benchiheub. Contribution à l'analyse des mouvements 3D de la Langue des Signes Française (LSF) en Action et en Perception. Traitement du signal et de l'image [eess.SP]. Université Paris Saclay (COMUE), 2017. Français. NNT : 2017SACLS559 . tel-01839676

HAL Id: tel-01839676

<https://theses.hal.science/tel-01839676>

Submitted on 16 Jul 2018

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.

Contribution à l'analyse des mouvements 3D de la Langue des Signes Française (LSF) en Action et en Perception

Thèse de doctorat de l'Université Paris-Saclay
préparée à l'Université Paris-Sud

École doctorale n°566
Sciences du Sport, de la Motricité et du Mouvement Humain
Spécialité de doctorat : Sciences du Sport et du Mouvement Humain

Thèse présentée et soutenue à Orsay, le 08/11/2017, par

Mohamed-El-Fatah Benchiheb

Composition du Jury :

M. Michel-Ange Amorim Professeur, Université Paris-Sud	Président
M. Thierry Pozzo Professeur, Université de Bourgogne	Rapporteur
M. John McDonald Professeur associé, Université DePaul	Rapporteur
M. Dominique Boutet, Maître de conférences, Université de Rouen	Examineur
Mme. Annelies Braffort Directrice de Recherche, Limsi – CNRS	Directrice de thèse
M. Bastien Berret Maître de conférences, Université Paris-Sud	Co-Directeur de thèse

Remerciements

Une thèse doctorat est loin d'être une entreprise individuelle. Si j'ai réussi à mener à bien mes travaux de recherche, c'est sûrement grâce à la participation et au soutien de plusieurs personnes. Je tiens donc à remercier ceux qui m'ont aidé et accompagné durant ma thèse.

Je désire tout d'abord remercier mes directeurs de recherche. Merci à Annelies Braffort et Bastien Berret de m'avoir permis de travailler sur des problématiques de recherche que je ne connaissais pas. Grâce à eux, j'ai pu faire mes travaux de recherche dans d'excellentes conditions et découvrir le monde de la Langue des Signes Française et du contrôle moteur. À chacune de mes questions, ils ont su m'apporter une réponse avisée.

Comme une thèse est un travail de longue haleine, je tiens à remercier mes collègues qui m'ont permis de travailler au quotidien dans la bonne humeur. Je remercie : Annick et Laurence, pour leur gentillesse, mes deux voisins de bureau Michael et Cyril avec qui j'ai eu le plaisir de travailler durant ces années. Je tiens à remercier aussi les thésards du LIMSI pour nos échanges, mais aussi pour avoir mis en place une atmosphère de travail agréable.

Je remercie mes amis Tahar, Yassine, Nassim et Hamdi pour leur soutien et tous les moments passés ensemble et qui m'ont permis d'oublier le stress du travail. Je remercie également mon ami Moussab qui a pris le temps de relire une partie de ce mémoire.

Je tiens également à remercier ma famille qui m'a toujours encouragé, soutenu et aidé. Je remercie ma mère pour toute son attention, mon père pour l'intérêt qu'il a porté à ma thèse, mon frère et ma sœur pour leurs encouragements. Un grand merci à ma grand-mère à qui je dédie ces travaux.

Un remerciement tout spécial à Djihan, la femme qui partage ma vie et qui a été à mes côtés durant les bons et les mauvais moments. Je la remercie pour son soutien lors des moments de démotivation et pour l'énergie qu'elle m'a apportée.

J'aimerais tout particulièrement remercier et témoigner toute ma reconnaissance à mes chers parents qui tiennent une immense place dans mon cœur pour m'avoir tout donné.

Pour terminer, je voudrais remercier toutes les personnes que je n'ai pas pu citer ici, mais qui m'ont aidé également et que je n'oublie pas.

Table des matières

Liste des symboles et acronymes	4
1 Introduction générale	5
1.1 Les enjeux sociétaux liés à la surdit�.....	5
1.2 Le mouvement 3D en LSF.....	6
1.3 Cr�ation et �tude d'un corpus 3D de LSF	6
1.4 Organisation du manuscrit.....	8
2 La Langue des Signes Fran�aise, le Contr�le Moteur, et la Perception du mouvement humain.....	10
2.1 D�finitions	10
2.1.1 Qu'est-ce que la Langue des Signes Fran�aise ?.....	10
2.1.2 Qu'est-ce que le contr�le moteur ?	13
2.1.3 Qu'est-ce que la cognition motrice et le couplage perception/action?.....	16
2.1.4 Qu'est-ce que la capture de mouvement ?	17
2.2 Etat de l'art	22
2.2.1 Langue des Signes	22
2.2.2 Contr�le moteur.....	27
2.2.3 Le mouvement et la LS	34
2.2.4 La Langue de Signes et les informations pertinentes pour la compr�hension ...	36
2.2.5 Conclusion.....	41
3 Cr�ation et annotation du corpus de la Langue des Signes Fran�aise	43
3.1 Cr�ation du Corpus MOCAP1.....	43
3.1.1 Contenu du corpus.....	43
3.1.2 Mat�riel utilis�.....	47
3.1.3 Les Marqueurs.....	47
3.1.4 Disposition des cam�ras	50
3.1.5 Calibration.....	51

3.1.6	Enregistrement des données (corpus).....	56
3.1.7	Post-traitement	56
3.1.8	Export des données.....	61
3.2	Annotation du Corpus.....	62
3.2.1	Le logiciel ANVIL	62
3.2.2	Création des fichiers BVH	63
3.2.3	Structure de l'annotation	64
3.2.4	La procédure d'annotation	66
3.2.5	L'export de l'annotation.....	68
3.3	Conclusion.....	69
4	Méthode et analyse des données	70
4.1	Méthode de traitement automatique des données.....	70
4.1.1	L'organisation des données	71
4.1.2	Les opérations sur le mouvement.....	72
4.2	Résultats et discussions	80
4.2.1	Principe d'isochronie et relation distance-durée.....	81
4.2.2	La loi des asymétries verticales et horizontales	89
4.2.3	La loi de puissance 2/3	100
4.3	Conclusion.....	102
5	Perception du mouvement sur un signeur virtuel en LSF.....	104
5.1	Méthodologie.....	104
5.1.1	Participants	104
5.1.2	Contenu	105
5.1.3	Procédure.....	108
5.2	Résultats.....	109
5.2.1	Effet du type de vidéo sur la performance	110
5.2.2	Effet principal du Groupe sur la performance.....	110

5.2.3	Interaction Groupe*Type de vidéo sur la performance.....	111
5.2.4	Effet de la catégorie et des Scènes décrites sur la performance.....	112
5.2.5	Effet principal du groupe sur la performance en Condition Avatar.....	112
5.2.6	Effet de la vidéo sur la performance en CA pour l'ensemble des participants	114
5.2.7	Effet d'interaction Vidéo*Groupe en condition Avatar.....	114
5.2.8	Analyse descriptive et apports linguistiques	115
5.3	Discussion.....	116
5.4	Conclusion	121
6	Conclusion générale.....	123
6.1	Contributions de la thèse	123
6.2	Limites de cette thèse et perspectives.....	124
7	Bibliographie.....	128
8	Annexe 1 : Catégories des signes et attributs de l'annotation	148
8.1	Catégories des signes.....	148
8.2	Les attributs de l'annotation du corpus.....	148
9	Annexe 2 : Prétraitement des données du corpus	151
9.1	Importation des données	151
9.1.1	Importation et lecture des données de capture de mouvement.....	151
9.1.2	Importation des données d'annotation	152
9.1.3	Importation des données de synchronisation	152
9.2	Extraction des données de déplacement	153
10	Annexe 3 : Contenu du corpus	155

Liste des symboles et acronymes

Amax : Accélération maximale

ASL : Langue des Signes Américaine

Curv : Indice de courbure global

DA : Durée d'accélération

DD : Durée de décélération

DDL : Degrés De Liberté

Dist : Distance parcourue

DM : Durée du mouvement

Dmax : Décélération maximale

FinSL : Langue des Signes Finlandaise

IRMf : Imagerie par Résonance Magnétique fonctionnelle

LCCM : Lois Classiques du Contrôle Moteur

LS : Langue des Signes

LSF : Langue des Signes Française

Mocap : Motion-Capture (Capture de mouvement)

OMS : Organisation Mondiale de la Santé

SNC : Système Nerveux Central

Vinst : Vitesse instantanée

Vmax : Vitesse maximale (vitesse de pointe)

Vmoy : Vitesse moyenne

1 Introduction générale

1.1 Les enjeux sociétaux liés à la surdité

Selon l'Organisation mondiale de la Santé (OMS) près de 360 millions de personnes dans le monde vivent avec un handicap dû à l'absence d'audition à la naissance ou à la perte de l'audition. Une partie des personnes sourdes utilisent notamment les Langues des Signes (LS) pour communiquer. Les LS sont des langues naturelles, visuelles et gestuelles et chaque pays a sa langue spécifique.

Pour l'OMS, l'absence d'audition et particulièrement l'absence d'éducation adaptée aux enfants sourds peuvent avoir des conséquences importantes chez ces individus notamment pour l'emploi et l'éducation. Selon de nombreuses études, les niveaux d'écriture et de lecture des individus sourds est souvent inférieur à celui des individus entendants (Kyle & Harris, 2006; Luckner & Handley, 2008; Qi & Mitchell, 2012; Wauters, Van Bon, & Tellings, 2006). Avec les avancées technologiques dans la société et la vie quotidienne, les obstacles à la communication, à l'accès aux informations et aux services, peuvent devenir important pour ces personnes.

Cependant, ces avancées technologiques ont apporté des solutions pour diminuer certains de ces problèmes. En effet, au cours de ces dernières années, les recherches sur les LS, dont le but est de rendre l'information et les services accessibles aux personnes sourdes se sont développées.

En France, la Langue des Signes Française (LSF) n'a été reconnue comme langue à part entière qu'en 2005, dans le cadre d'une loi sur le handicap. Cette loi impose aussi aux lieux publics de rendre les informations accessibles, quel que soit le type de handicap. Donc pour les usagers sourds, cela implique de développer des applications permettant la diffusion de messages en LSF à l'aide de « signeurs virtuels ». Ce sont des personnages virtuels qui sont animés et s'expriment en LSF.

La génération de la LSF à l'aide des signeurs virtuels reste très limitée car la LSF est encore peu décrite et les modèles de génération existants sont très simplifiés. Le défi majeur de cette thèse est de contribuer à la description de son fonctionnement et de mieux comprendre ses propriétés en action et en perception dans un but ultérieur de modélisation.

1.2 Le mouvement 3D en LSF

La LSF comme beaucoup d'autres LS est une langue encore peu décrite, en particulier en ce qui concerne le mouvement des articulateurs¹. La recherche sur les LS s'est concentrée sur la compréhension et la modélisation des propriétés linguistiques. Peu d'études ont essayé de comprendre la cinématique et la dynamique du mouvement en lui-même et la façon dont il améliore la compréhensibilité de la LS générée. Ainsi, les rares études qui se sont intéressées au mouvement ont été réalisées sur des corpus vidéo, ce qui ne permet pas d'étudier avec fiabilité le mouvement (Lefebvre-Albaret, 2010; Segouat & Braffort, 2009).

Pour mieux prendre en compte le mouvement de la LS cela nécessite de nouvelles ressources, à savoir :

- L'enregistrement de la cinématique 3D qui permet d'étudier le mouvement avec fiabilité.
- De concevoir des modèles de générations de LS plus naturels.
- D'améliorer les connaissances dans les différentes disciplines liées à la LS.

En effet, des études antérieures ont montré que les corpus 3D ont un intérêt pour les LS dans plusieurs domaines : la génération des mouvements naturels, la segmentation automatique, les études linguistiques, etc. (Heloir, Gibet, Multon, & Courty, 2006; Lu & Huenerfauth, 2010; Malaia, Borneman, & Wilbur, 2008; Malaia, Wilbur, & Milković, 2013; Malaia & Wilbur, 2012). Ainsi, le mouvement est un paramètre prépondérant dans la LS. Des études ont montré qu'il est d'une valeur incontournable. En particulier, la cinématique du mouvement a permis d'identifier les types de verbe dans la Langue des Signes Américaine (ASL) (Malaia et al., 2008; Malaia & Wilbur, 2012). L'analyse en composante principale (ACP) appliquée sur le mouvement de la main a permis de segmenter avec fiabilité la LSF (Heloir et al., 2006).

1.3 Création et étude d'un corpus 3D de LSF

Pour faire progresser les connaissances sur la LSF et créer à terme des modèles de génération, il est nécessaire d'étudier cette problématique sur deux axes : (1) l'axe de la production des mouvements (action) et (2) l'axe de la compréhension de la LSF générée (perception). Ces deux axes sont intimement liés chez l'humain et constituent le couplage perception/action.

¹ Les articulateurs sont les différentes parties du corps (mains, les expressions faciales, les bras, etc.) qui contribuent à transmettre les informations en langue des signes.

La figure 1-1 montre l'articulation entre action, perception et modélisation en LSF.

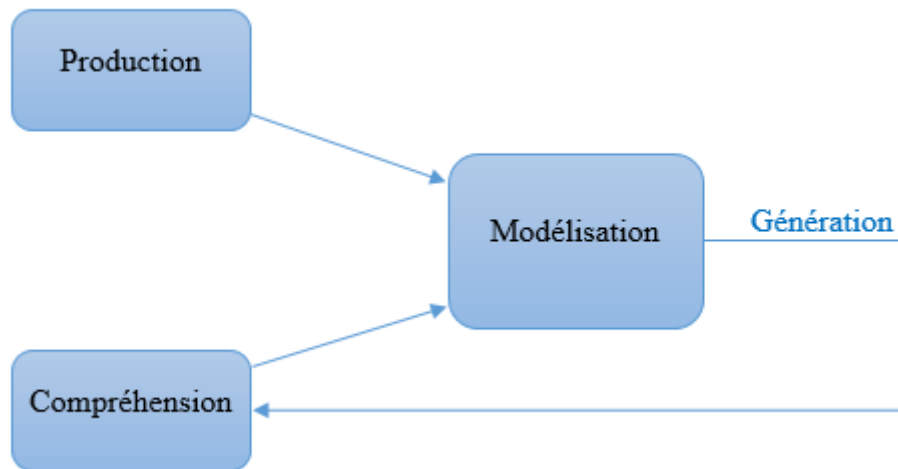


Figure 1-1: Schéma des liens entre production, compréhension et modélisation du mouvement en LSF.

Comprendre comment sont produits les mouvements de LSF et quelles informations perceptives sont cruciales à une bonne compréhension de la LSF permettra de mieux savoir comment modéliser les mouvements LSF. Une fois un modèle élaboré, il faudra tester sa compréhensibilité auprès des utilisateurs.

Nous avons expliqué pourquoi il est important de concevoir des modèles de génération de mouvements de LSF naturels à l'aide d'un signeur virtuel. Pour proposer de tels modèles, un prérequis est d'étudier les mouvements de la LSF afin notamment de comprendre les différences cinématiques entre les deux types de mouvements qui la compose (les signes et les transitions entre les signes) et d'analyser l'importance du mouvement des membres supérieurs dans la compréhension de la LSF.

Pour ce faire, nous proposons d'étudier la validité des lois classiques du contrôle moteur dans les deux types de mouvements de LSF. Ces lois connues et validées dans plusieurs types de mouvements (ex : pointage, geste d'atteinte de cible, dessin de formes, etc.), n'ont pas été étudiées dans le contexte de la LS. Dans cette démarche, nous avons décidé de vérifier trois lois : le principe d'isochronie, la loi des asymétries, et la loi de puissance $2/3$.

Pour cela, l'analyse vidéo n'est pas suffisante et il faut un corpus de données 3D de LSF. Comme il y a un manque de corpus 3D dans la communauté scientifique, et afin d'étudier la cinématique du mouvement dans le contexte de LSF, nous avons décidé de créer et d'annoter un corpus de LSF composé de données 3D issues de capture du mouvement. Ce corpus est

destiné à la recherche pluridisciplinaire, notamment les études linguistiques, informatiques et en sciences du mouvement, et constitue donc une plus-value importante dans le domaine de la LSF.

Afin de comprendre comment des individus sourds perçoivent la LSF en fonction de la quantité d'informations qui leur est disponible, et l'importance du mouvement dans la génération de LSF à l'aide d'un signeur virtuel, nous avons aussi utilisé ce même corpus pour mesurer la compréhension d'individus sourds et entendant, face à deux différents types de vidéos (personne réelle et un signeur virtuel).

En résumé, les motivations de cette thèse se déclinent en **trois objectifs principaux** :

- **Créer et annoter un corpus de LSF de données 3D, destiné à la recherche pluridisciplinaire.**
- **Étudier les mouvements de LSF afin de vérifier les lois classiques du contrôle moteur dans les mouvements portant une information linguistique ou dans ceux qui n'en portent pas.**
- **Examiner comment les individus sourds perçoivent la LSF en fonction de la quantité d'informations qui est disponible, notamment pour ce qui est des membres supérieurs.**

1.4 Organisation du manuscrit

Afin d'apporter des réponses à ces questions complexes, les travaux de cette thèse s'articulent sur trois domaines : la langue des signes, le contrôle moteur et la perception motrice (figure 1-2).

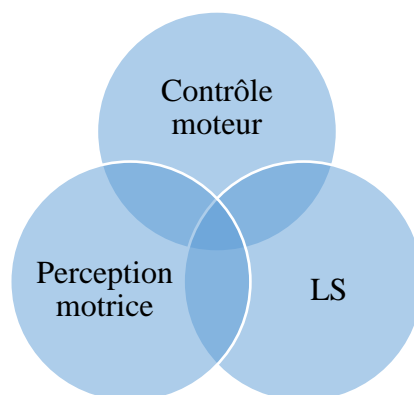


Figure 1-2: Les trois disciplines scientifiques abordées dans cette thèse.

Cette thèse contient six chapitres. Celui-ci est le premier, il pose la problématique générale de l'étude et définit le cahier des charges de la thèse.

Le deuxième chapitre présente les problématiques et l'état des lieux. Il est structuré en deux sections, la première introduit les notions importantes à connaître pour bien comprendre les notions manipulées dans ce manuscrit. La deuxième partie présente les corpus de LS disponibles, les lois du contrôle moteur étudiées dans cette thèse et la perception de la LS.

Le troisième chapitre présente le corpus créé dans le cadre de cette thèse, son contenu et la méthode utilisée durant sa création. Aussi, nous présentons le logiciel et la méthode suivie pour l'annoter.

Le quatrième chapitre est consacré à l'analyse des données. Il est structuré aussi en deux sections. La première section présente la méthode de traitement des données pour étudier le mouvement. La deuxième section présente l'étude des lois de contrôle moteur dans les mouvements de LSF.

Le cinquième chapitre présente l'étude réalisée pour comprendre comment des individus sourds perçoivent la LSF en fonction de la quantité d'informations qui leur est disponible.

Le dernier chapitre conclut l'étude, ouvre des perspectives sur de futurs travaux et valorisations.

2 La Langue des Signes Française, le Contrôle Moteur, et la Perception du mouvement humain

Ce chapitre présente les domaines principaux de cette thèse, à savoir la langue des signes (LS), le contrôle du mouvement et sa perception (cognition motrice).

Il est structuré en deux sections, une première qui introduit les notions importantes et une deuxième qui présente un état de l'art dans ces domaines.

2.1 Définitions

Cette première partie vise à introduire quelques notions fondamentales : la LSF, le contrôle moteur, la perception du mouvement humain, la cognition motrice, et la capture de mouvement.

2.1.1 Qu'est-ce que la Langue des Signes Française ?

Les LS sont des langues naturelles pratiquées au sein des communautés de Sourds et donc la LSF est celle utilisée en France. Cette section présente la LS selon trois aspects, respectivement physiologique (section 2.1.1.1), historique (section 2.1.1.2) et linguistique (section 2.1.1.3).

2.1.1.1 Aspect physiologique

La LSF est une langue gestuelle qui utilise tous les articulateurs de la partie supérieure du corps : Le buste, les épaules, les bras, les mains, la tête, le visage et les yeux. Chaque unité gestuelle est produite en associant la position, la configuration et l'orientation de la main et des doigts avec son emplacement, son mouvement, ainsi que des gestes et des mouvements réalisés par les articulateurs non manuels.

La LS est une langue visuelle, où le cerveau traite les informations linguistiques captés par les yeux. Ce qui implique que les expressions du visage et les mouvements du corps jouent un rôle important pour transmettre les informations.

Ainsi, les LS sont des langues visio-gestuelles qui permettent de véhiculer plusieurs informations simultanément par l'intermédiaires d'un ensemble d'articulateurs qui se déploient dans l'espace.

La figure 2-1 montre un exemple de communication en LS, où l'espace de signation est la zone devant le locuteur où les signes sont réalisés.

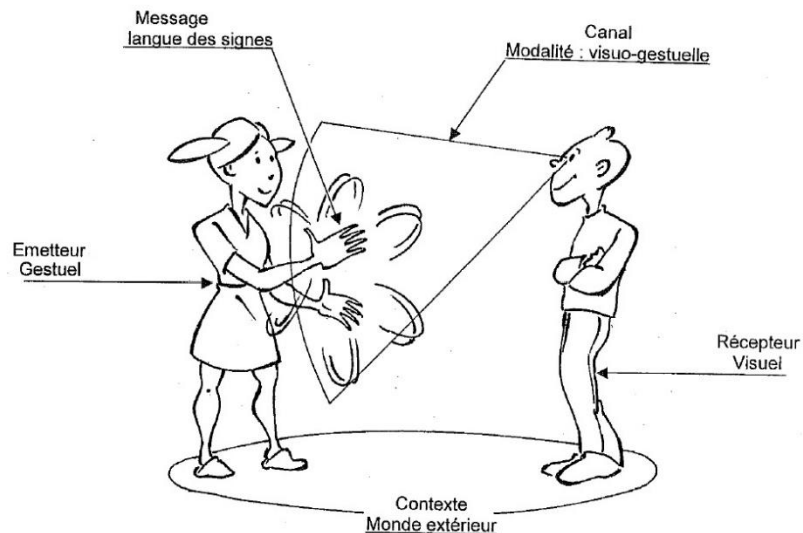


Figure 2-1: Exemple de communication en LS (illustration de Laurent Verlaine).

2.1.1.2 Aspect historique

La LSF a une histoire un peu particulière, elle a été reconnue comme une langue à part entière en 2005.

Durant l'Antiquité, Aristote pensait que quelqu'un qui ne parle pas, ne pense pas. De ce fait, les sourds n'ont bénéficié d'aucune éducation. A partir du 16ème siècle, des éducateurs ont commencé à l'exercer. En Espagne, des enfants sourds issus des familles nobles ont été instruits par des précepteurs. Pedro Ponce de León est l'un d'entre eux, il est souvent considéré comme le premier professeur pour les sourds (Daniels, 1997).

En 1760, le prêtre français Charles-Michel de L'Épée, appelé l'abbé de L'Épée a créé la première école pour les sourds à Paris. De plus, il est le premier entendant qui s'est intéressé à la communication des sourds (Berthier, 1852; Bézagu-Deluy, 1990; PRESNEAU, 1998). En 1779, Pierre Desloges a publié le premier livre écrit par un sourd (Desloges, 1779). Dans ce livre, Desloges recommandait d'utiliser une langue signée pour l'éducation des enfants sourds.

En 1880, le congrès international sur l'éducation des sourds qui s'est tenu à Milan, a conclu à des résolutions affirmant que la « méthode orale » doit être préférée à la « méthode gestuelle ». À la suite de ce congrès, le gouvernement français a décidé de faire appliquer ces résolutions sur tout le territoire : désormais la langue des signes n'est plus utilisée dans les écoles pour sourds, ce sont la parole et la lecture sur les lèvres qui sont principalement enseignées. Ce congrès apparaît comme une mascarade pour la LS, tant par le choix du lieu du congrès, les

Italiens qui étaient défenseurs de l'oral, ainsi que des personnes invitées pour mener les débats et voter (seuls trois sourds étaient présents) (Encrevé, 2008). Cette interdiction a continué jusqu'à la fin du 20ème siècle.

À partir des années 1970, on a assisté au « réveil sourd » : les sourds ont milité pour une reconnaissance de la langue des signes, notamment pour l'éducation des jeunes sourds. Progressivement, cette langue retrouve sa place, avec le linguiste William Stokoe, qui prouve que la LS est une véritable langue. D'autres chercheurs en linguistique et en sociologie tels que Christian Cuxac et Bernard Mottez poursuivent ce travail

Enfin, en 2005 la LSF est devenue une langue à part entière dans le cadre de la loi sur l'handicap du 11 février 2005. Cette loi impose aussi aux lieux publics de rendre les informations accessibles, quel que soit le type de l'handicap. Depuis, les recherches linguistiques se multiplient, ainsi que les formations universitaires et professionnelles menant aux métiers d'enseignant de la LSF, interprète en LS, etc.

2.1.1.3 Aspect linguistique

Les LS sont des langues visuo-gestuelles et leur système linguistique exploite des canaux spécifiques : de nombreuses informations sont exprimées simultanément et s'organisent dans l'espace, et l'iconicité joue un rôle central.

La composante manuelle est la source d'une grande part de l'information émise, mais elle n'est pas la seule, car les expressions faciales et le regard ont un rôle déterminant dans la désambiguïsation de certains signes, et pour d'autres fonctions linguistiques telle que la qualification (gros, petit, etc.). Le mouvement possède aussi de nombreux rôles linguistiques.

L'iconicité est une particularité dans la LSF, elle permet de représenter des signes et des situations de manière iconique (Cuxac, 1996, 2000). Cette particularité permet d'exprimer des idées complexes sans recourir au lexique standard.

Une autre particularité de la LSF est de placer des éléments du discours dans des endroits spécifiques dans l'espace de signation. Pour rappeler ces éléments ultérieurement, il suffit de désigner leur position avec le regard ou avec un pointage.

Selon les théories linguistiques, la notion de signe est très variable. Pour notre étude, nous considérons toutes les unités gestuelles manuelles porteuses de sens sans restriction, qu'elles correspondent à des unités lexicales (équivalentes à des mots des langues vocales) ou à des

unités à visées illustratives telles que définies par Christian Cuxac. Ces unités illustratives permettent par exemple de décrire la taille ou la forme des relations des entités du discours (i.e. figure 2-2).



Figure 2-2: À gauche correspond à un signe illustratif montrant l'emplacement des bouteilles dans la photo et à droite montre le signe lexical vin.

2.1.2 Qu'est-ce que le contrôle moteur ?

Cette partie présente le contrôle moteur, en commençant par donner des définitions (section 2.1.2.1), expliciter la notion de degré de liberté (section 2.1.2.2), puis les notions de synergies ou primitives motrices (section 2.1.2.3)

2.1.2.1 Définition

Le contrôle moteur est le processus par lequel l'humain utilise son cerveau pour activer et coordonner les muscles et les membres participant à l'exécution d'une habileté motrice. Fondamentalement, il est l'intégration de l'information sensorielle, à la fois sur l'environnement et l'état actuel du corps, afin de déterminer l'ensemble approprié de forces musculaires et l'activation des articulateurs pour générer un mouvement ou une action souhaitée. Ce processus nécessite une interaction de coopération entre le système nerveux central (SNC) (voir figure 2-3) et le système musculo-squelettique, ce qui est donc un problème de traitement d'information, de coordination, de mécanique, de physique et de cognition, c'est-à-dire un problème de *contrôle*. Un contrôle moteur efficace est crucial pour interagir avec le monde, non seulement pour déterminer les capacités d'action, mais aussi la régulation de l'équilibre et la stabilité (Rosenbaum, 1991; Wise & Shadmehr, 2002).

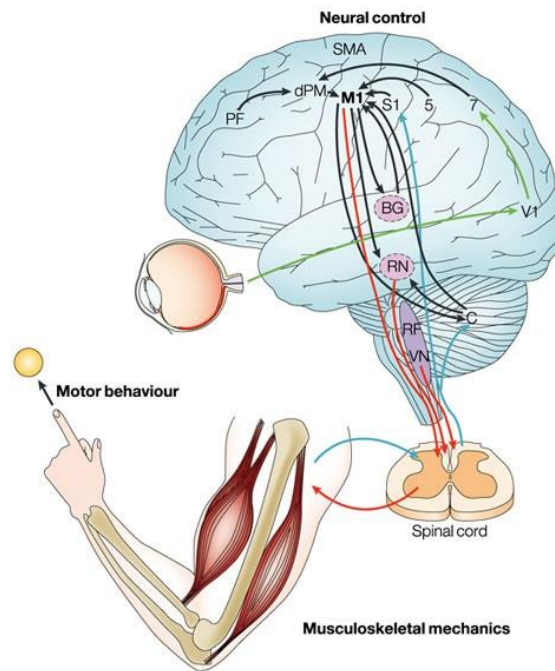


Figure 2-3: Interaction entre le SNC et le système musculo-squelettique afin d'exécuter une habileté motrice (Scott, 2004).

La planification et l'exécution d'un mouvement est un problème complexe, de sorte que l'étude du contrôle moteur a été abordée à partir de nombreuses disciplines, y compris la psychologie, les sciences cognitives, la biomécanique et les neurosciences. Bien que l'étude moderne du contrôle moteur soit un domaine de plus en plus interdisciplinaire, les questions de recherche ont souvent été définies comme étant neurophysiologiques en investiguant les propriétés physiques et biologiques du système sensorimoteur humain. Des domaines d'études liés au contrôle moteur sont par exemple le couplage action-perception, l'interaction posture-mouvement, la coordination motrice, les apprentissages moteurs ou encore la planification du mouvement volontaire.

2.1.2.2 La notion de degré de liberté

Le problème des degrés de liberté (DDL) est central dans le contrôle moteur. Il indique qu'il y a plusieurs façons pour les êtres humains et les animaux d'effectuer un mouvement en vue d'atteindre un même objectif. En d'autres termes, dans des circonstances normales, aucune correspondance simple et encore moins unique n'existe entre un problème moteur (ou tâche motrice) et sa solution. Selon le neurophysiologiste russe Nikolai Bernstein, il est clair que les difficultés de base pour la coordination des différents segments corporels proviennent précisément de l'extrême abondance des degrés de liberté, que le SNC ne gère sans doute pas indépendamment et individuellement (Bernstein, 1967).

Le corps humain a des DDL anatomiques redondants (il y a plusieurs trajectoires articulaires et activités musculaires pour effectuer un même signe en LS), des DDL cinématiques redondants (les mouvements peuvent avoir des trajectoires, des durées, des vitesses et des accélérations différentes pour réaliser un signe donné ; cela peut d'ailleurs varier d'un signe à l'autre et entre différents signeurs), et des DDL neurophysiologiques redondants (plusieurs motoneurons ont des synapses sur le même muscle)² (Bernstein, 1967; Scott, 2004). Comment le SNC choisit un sous-ensemble de ces DDLs quasi-infinis est une question cruciale dans la compréhension globale du contrôle moteur et son développement à travers l'évolution ou au cours de la vie d'un individu.

2.1.2.3 Les synergies ou primitives motrices

Afin de saisir, utiliser et manipuler des objets ou bien réaliser certains signes comme dans notre cas, la main est chargée de créer une multitude de postures dans la vie quotidienne (Jerde, Soechting, & Flanders, 2003; Tyrone & Mauk, 2012). La main dispose de plus de 25 DDL pour accomplir ces tâches (MacKenzie & Iberall, 1994). Comment le cerveau contrôle les nombreux DDL du corps reste l'un des problèmes fondamentaux du contrôle moteur (e.g., Bernstein, 1967; Latash, 2008; Turvey, 2007). Il est souvent supposé que le système nerveux central appelle quelques modèles de mouvements clés lors du contrôle de la main, cette hypothèse a été soutenue par des études neuronales (Saleh, Takahashi, & Hatsopoulos, 2012), musculaires (d'Avella, Portone, Fernandez, & Lacquaniti, 2006; d'Avella, Saltiel, & Bizzi, 2003) et cinématiques (Santello, Flanders, & Soechting, 2002; Vinjamuri et al., 2010; Vinjamuri, Mao, Scialabassi, & Sun, 2007) sur les mouvements d'atteinte et de saisie d'un objet (Santello, Baud-Bovy, & Jörntell, 2013).

Les concepts de synergie et de « contrôle de la synergie » ont suscité un intérêt considérable dans le contrôle moteur au cours des dernières années comme une solution possible à ce problème des degrés de liberté. Le concept a d'ailleurs été étendu aux mouvements des membres supérieurs et inférieurs, que ce soit au niveau cinématique (Berret, Bonnetblanc, Papaxanthis, & Pozzo, 2009) ou musculaire (d'Avella et al., 2003). Selon Turvey (Turvey, 2007), une synergie est « une collection de degrés de liberté relativement indépendants qui se comportent comme une seule unité fonctionnelle, ce qui signifie que les degrés de liberté internes prennent soin d'eux-mêmes, s'adaptant à leurs fluctuations mutuelles et aux fluctuations

² Les motoneurons constituent la voie de sortie du SNC ou la voie finale de tout acte moteur. La synapse désigne une zone de contact fonctionnelle qui s'établit entre deux neurones, ou entre un neurone et une autre cellule

du champ de force externe, et le font d'une manière qui préserve l'intégrité de la fonction de la collection ». En d'autres termes, la synergie est une sorte de regroupement de plusieurs éléments du système qui agit comme une seule unité à travers quelques motifs de coordination pour exécuter une tâche. L'intérêt des synergies, ou primitives motrices, serait de créer des mouvements de manière simplifiée en se basant sur des patterns qui peuvent être combinés afin d'en créer de nouveaux, un peu à la manière des lettres de l'alphabet qui permettent de créer des mots puis des phrases par association. La structure de la LS diffère cependant de la langue écrite dont les termes primitives et synergies sont à entendre dans le domaine du contrôle moteur. Un certain nombre de principes ou lois du mouvement humain ont été identifiés par ailleurs, et nous reviendrons dessus en section II.2. Pour étudier la structure et les caractéristiques des mouvements produits en LS (et en dehors), mesurer le mouvement humain est alors un point essentiel qui est abordé ci-dessous.

2.1.3 Qu'est-ce que la cognition motrice et le couplage perception/action?

Selon Sommerville et Decety, le concept de cognition motrice saisit l'idée que la cognition est incarnée dans l'action et que le système moteur participe à ce qui est généralement considéré comme des processus cognitifs, y compris ceux impliqués dans l'interaction sociale (Sommerville & Decety, 2006).

L'action une série de mouvements produit en réaction à un événement significatif dans les environnements physiques et sociaux ou pour satisfaire une intention vers un but moteur spécifique. Elle est l'unité fondamentale du paradigme de la cognition motrice.

La cognition motrice prend en compte la préparation et la production des actions, ainsi que les processus impliqués dans la reconnaissance, la prédiction, l'imitation et la compréhension du comportement d'autrui.

Selon Sperry, la perception est un moyen d'action et l'action est un moyen de perception. Autrement dit, Les processus de perception et d'action sont en effet fondamentalement imbriqués (Sperry, 1952). En effet, une fonction essentielle de cerveau est de transformer les modèles de la perception sensorielle en modèles de coordination motrice.

Plus récemment, Prinz a démontré que la perception et l'action partagent des codes de calcul communs et des architectures neuronales sous-jacentes (Prinz, 1997). Il repose sur l'hypothèse qu'une action est codée par le biais d'effets perceptibles. En d'autres termes, le modèle d'une action se fait à partir d'un modèle de perception sensorielle.

L'habilité à percevoir le mouvement biologique, c'est-à-dire les mouvements produits par des humains ou des animaux semble une capacité innée. Cette capacité a été vérifiée chez des nourrissons, où il a été montré que les nouveau-nés sont capables de différencier les cinématiques biologiques et non biologiques (Méary, Kitromilides, Mazens, Graff, & Gentaz, 2007). Cependant, pour permettre une identification précise des actions perçues, la perception et la reconnaissance visuelle des mouvements humains se construisent au cours de la vie et nécessitent des représentations sensori-motrices.

2.1.4 Qu'est-ce que la capture de mouvement ?

Cette section présente la capture du mouvement, sa mise en œuvre (section 2.1.4.1), avec les différentes méthodes de capture (section 2.1.4.2), ainsi que son utilisation (section 2.1.4.3)

La motion capture (mocap), ou capture de mouvement en français, est une technologie qui permet d'enregistrer des mouvements d'une personne ou d'un objet (figure 2-4). Les données mocap enregistrées sont transformées en un format numérique pour des analyses et des traitements ultérieurs ou pour en contrôler une contrepartie virtuelle sur ordinateur (caméra, modèle 3D, ou avatar). L'enregistrement fournit les coordonnées de déplacements comme une matrice numérique qui peut être utilisée comme source de données pour analyser les mouvements des parties du corps à partir d'un point de vue cinématique.

La mocap a été inventée à la fin du 19^e siècle où Étienne-Jules Marey et Eadweard Muybridge ont présenté une chronophotographie pour enquêter scientifiquement sur les mouvements d'animaux et d'humains. Les premiers essais ont été faits à travers l'étude des mouvements d'un cheval au galop avant d'être projetés sur écran, c'est ce qui a permis de créer le cinéma. A la base la mocap a été conçue dans les années 70 pour l'armée et la recherche scientifique. Cette nouvelle technologie a grandement contribué à étudier les mouvements humains, et elle a permis de bien progresser dans les domaines de la santé et de l'anatomie. Pour des systèmes complexes comme le corps humain avec un grand nombre de marqueurs (ce sont de petites boules recouvertes d'une matière blanche, réfléchissante), la mocap reste un exercice difficile et qui requiert une grande expertise, certaines entreprises étant d'ailleurs spécialisées en mocap au service de l'industrie du cinéma ou des jeux vidéo voire de la recherche académique comme en LS (cf. Mocaplab³ par exemple).

³ <http://www.mocaplab.com/fr/>



Figure 2-4: Un danseur portant une combinaison avec des marqueurs passifs réfléchissants, lors d'une capture optique de mouvement. (wikipédia).

2.1.4.1 Comment faire la mocap ?

Dans la plupart des cas, un sujet (l'homme, l'animal, ou un objet), est utilisé comme source de données qui seront utilisées dans des analyses ou transformées en animation. De cette façon, les mouvements du sujet virtuel seront très proches de la réalité.

Il y a plusieurs étapes de capture de mouvement, cinq pour être précis. La première et la deuxième étape sont le réglage de la disposition du studio (caméras) et des marqueurs, c'est-à-dire préparer le studio pour la capture de mouvement, installer les caméras aux bons endroits et positionner les marqueurs sur le sujet. La troisième et la quatrième étape sont la calibration de la zone de capture et la capture de mouvement. La dernière étape consiste à nettoyer les données enregistrées.

2.1.4.2 Les méthodes de capture

Il existe trois principales techniques pour effectuer la capture de mouvement : les technologies optiques, par centrale inertielle et mécaniques.

2.1.4.2.1 Capture optique

Dans cette catégorie de technologie, on distingue trois types de dispositifs, selon qu'ils sont basés sur des caméras infrarouges, des cellules photosensibles ou des caméras sans marqueurs.

a. Caméras infrarouges et marqueurs passifs réfléchissants

Pour cette technique, les sujets, en général, sont vêtus d'une combinaison noire, sur laquelle on dispose, à des points stratégiques (souvent les articulations) des marqueurs (voir figure 2-4). Les caméras envoient un rayonnement infrarouge qui est alors réfléchi par la surface des marqueurs, vers ces mêmes caméras. Les images filmées sont en noir et blanc, sur 2 bits (1 pour chaque couleur).

Un objet avec des marqueurs attachés à des positions connues est utilisé pour calibrer les caméras et obtenir leurs positions, et la distorsion de l'objectif de chaque caméra est mesurée. Si deux caméras calibrées voient un marqueur, l'emplacement de ce marqueur en trois dimensions peut être obtenue.

Le problème principal de cette méthode, c'est sa sensibilité aux obstacles entre les capteurs et les caméras, les rayons infrarouges ne pourront pas traverser l'obstacle et donc ne seront pas réfléchis par les capteurs. Cela peut se produire par exemple lorsque les mains d'un sujet se trouvent toutes deux sur le même axe vis-à-vis des caméras.

b. Cellules photosensibles et marqueurs actifs

Cette technique est similaire à celle d'avant, la différence est que les marqueurs sont actifs. Le principe de fonctionnement est donc celui « d'émetteur-récepteur » : les marqueurs actifs émettent directement un signal infrarouge captés par les cellules photosensibles. Celles-ci ne sont sensibles qu'à un type de longueur d'onde et identifient les marqueurs actifs en temps réel grâce à une séquence d'émission multi-complexe. Ces marqueurs sont donc identifiés automatiquement et immédiatement en temps réel comme étant des points dynamiques avec des coordonnées 3D dans l'espace virtuel.

c. Caméras sans marqueurs (la technologie Kinect)

Cette technique utilise une seule caméra Kinect, cette caméra permettant de filmer en 3D. Par le biais de la caméra, une trame de lumière infrarouge est projetée sur les objets, qui, en la réfléchissant, permettent de définir une capture de profondeur. Cette technologie avait à l'origine été conçue pour la Xbox en septembre 2008, et elle a très vite été détournée de son utilisation première.

2.1.4.2.2 Capture par centrale inertielle

Le principe de ce système est le même que pour la méthode de la capture optique avec marqueurs, mais les capteurs sont remplacés par des capteurs gyroscopiques, des accéléromètres et magnétomètres et il n'y a pas de caméras. Ils permettent de capter en temps réel l'orientation, la vitesse angulaire et l'accélération linéaire de la partie du corps où ils sont positionnés. Un émetteur relié à l'ordinateur transmet l'information. Le point positif de cette méthode est qu'il n'y a pas de problème d'obstacle, mais le point négatif est que cela reste moins précis que la capture optique même si ces technologies progressent très vite.

2.1.4.2.3 Capture mécanique

Cette méthode fonctionne grâce à un exosquelette construit autour de l'élément à capter (voir figure 2-5). L'exosquelette est une structure dont les articulations sont des capteurs angulaires. Ces articulations sont reliées par des fils à un ordinateur. Chaque capteur connaît sa position par rapport aux autres, ce qui permet à l'ordinateur de reconstituer les mouvements de l'ensemble de la structure. Cette méthode est très précise et calcule rapidement les données enregistrées par les capteurs. La contrainte de cette méthode est la nécessité d'être relié par des fils à l'ordinateur, ce qui réduit le rayon d'action, et impose des contraintes mécaniques.



Figure 2-5: Système de mocap mécanique (modèle GYPSY 7⁴).

⁴ <http://metamotion.com/gypsy/gypsy-motion-capture-system.htm>

2.1.4.3 Les utilisations de la Mocap

La mocap est utilisée dans de nombreux secteurs, principalement dans les jeux vidéo, afin de rendre notamment les mouvements sportifs de plus en plus réalistes comme FIFA, Call of Duty et Tomb Raider. Peu à peu la mocap fait son apparition dans le cinéma comme dans le Seigneur des Anneaux, où le personnage de Gollum est entièrement fait en image de synthèse. Elle est utilisée également dans le sport, où elle permet de détailler les mouvements d'un athlète afin de pouvoir les améliorer ou simplement les enseigner à d'autres. Dans le domaine scientifique, elle est utilisée à différentes fins par des chercheurs travaillant dans le cadre des sciences du mouvement. L'extrême exactitude du système et son utilisation en dynamique sont grandement mises à profit dans la recherche biomécanique, les sciences sportives, la rééducation et l'ergonomie.

2.2 Etat de l'art

Cette deuxième partie présente un état de l'art dans le domaine des LS (section 2.2.1), des lois du contrôle moteur (section 2.2.2), de mouvement dans les LS (section 2.2.3) et de la perception du mouvement de LS (section 2.2.4).

2.2.1 Langue des Signes

Après avoir présenté quelques corpus de LS parmi les plus représentatifs (section 2.2.1.1), les principes de l'annotation sont introduits (section 2.2.1.2) et un bilan est donné (section 2.2.1.3).

2.2.1.1 Les corpus de LS

Le terme corpus désigne une composition de plusieurs documents regroupés ensemble, ces documents peuvent être des textes, des vidéos, des images, des données numériques, etc. Les corpus sont utilisables dans plusieurs domaines, tels que les études linguistiques, littéraires, scientifiques, philosophiques, etc. En LS les corpus les plus utilisés sont des corpus conçus avec une ou plusieurs caméras. Ces dernières années, la LS commence à disposer de corpus 3D grâce à l'utilisation de la mocap. Dans cette partie, nous allons présenter les corpus vidéo les plus représentatifs en LS, ainsi que les corpus 3D qui ont été constitués récemment.

2.2.1.1.1 Corpus Vidéo

Les corpus vidéo sont les plus courants en LS, ils sont utilisés par les linguistes et les informaticiens. Ils sont enregistrés à l'aide d'une ou plusieurs caméras. Ici nous listons quelques-uns parmi les plus représentatifs.

Tout d'abord, il y a des corpus destinés à servir les lexiques, ils permettent de suivre et d'analyser les changements dans le vocabulaire des langues des signes d'un point de vue linguistique. Par exemple, "The American Sign Language Lexicon Video Dataset" (Neidle & Vogler, 2012) constitue un lexique pour la LS américaine (ASL), contenant plus de 3000 signes. Le site AUSLAN SignBank⁵ fournit des définitions pour plus de 4500 signes de la LS australienne.

Deuxièmement, il y a des corpus destinés à la recherche linguistique sur les signes isolés et les discours. Par exemple, le corpus de la LS néerlandaise « *Nederlandse Gebarentaal: NGT* » (Crasborn & Zwitterlood, 2008) contient 12 heures de signation de la partie haute du corps avec une vue de face, annoté à l'aide de gloses qui correspondent à des mots de la langue écrite

⁵ <http://www.auslan.org.au/dictionary/>

servant à labelliser un signe. Depuis 2008, l'annotation a été étendue à des traductions dans différentes langues écrites. Rutkowski a créé un corpus de la LS polonaise contenant environ 300h de séquences vidéo de 80 signeurs sourds effectuant des tâches linguistiques prédéfinies (Rutkowski, Łacheta, Mostowski, Filipczak, & Łozińska, 2013). Le corpus CopyCat (Zafrulla, Brashear, Hamilton, & Starner, 2010) couvre l'ASL signée par les enfants incluant 420 phrases formées à partir d'un vocabulaire de 19 signes. L'Université de Hambourg, en Allemagne, est quant à elle en train de créer un dictionnaire et un corpus de la LS allemande de très grande taille (environ 300 signeurs et 400 heures)⁶.

Troisièmement, il y a des corpus explicitement créés ou adaptés pour le traitement du langage naturel et/ou des tâches de vision par ordinateur. Contrairement aux ressources linguistiques, les corpus créés pour les tâches de traitement automatique contiennent des petits vocabulaires de quelques centaines de signes au lieu de milliers, et se concentrent sur un petit nombre de thématiques linguistiques. L'objectif est par exemple de fournir des statistiques pour élaborer des modèles statistiques robustes pour les systèmes de reconnaissance automatique. Dreuw donne un aperçu sur ces corpus (Dreuw, Forster, & Ney, 2012). Parmi eux on peut citer les corpus RWTH-BOSTON, créés à l'origine pour la recherche linguistique à l'Université de Boston et adaptés à des fins de reconnaissance de formes par l'Université RWTH Aachen. Ils comportent plusieurs signeurs et jusqu'à 7768 occurrences au sein d'un vocabulaire de 483 signes. Braffort, Matthes et Efthimiou présentent des corpus de dialogues de langues des signes pour la LS française, allemande, grec et britannique (A Braffort et al., 2010; Efthimiou et al., 2012; Matthes et al., 2010). Ils ont été créés dans le cadre du projet Dicta-Sign⁷. Les corpus comprennent des vidéos en LS tournés en haute définition en vue frontale et latérale. Semblable au Corpus NGT, les corpus Dicta-Sign contiennent une vue plongeante (vue de haut) sur les signeurs permettant l'étude des mouvements de la main dans l'espace de signation. Forster a créé RWTH-PHOENIX- Weather un corpus de la LS allemande (Forster et al., 2012; Forster, Schmidt, Koller, Bellgardt, & Ney, 2014). Le corpus contient des prévisions météorologiques interprétées simultanément dans la LS qui ont été enregistrées à la télévision publique allemande. Le corpus a été créé dans le cadre du projet SignSpeak⁸ avec 9 interprètes, contient 1081 signes et 7000 phrases. Le corpus SIGNUM (von Agris, Knorr, & Kraiss, 2008) a été

⁶ <http://www.sign-lang.uni-hamburg.de/dgs-korpus/>

⁷ <http://www.sign-lang.uni-hamburg.de/dicta-sign/portal/>

⁸ <http://www.signspeak.eu/>

explicitement créé à des fins de reconnaissance des formes. Il se compose de 25 signeurs et près de 14000 occurrences au sein d'un vocabulaire de 450 gloses.

Enfin, pour des projets en traduction automatique, Almohimeed a créé un corpus qui contient 230 phrases de la LS arabe et de l'arabe écrit avec un vocabulaire de 710 signes (Almohimeed, Wald, & Damper, 2010). Morrissey a créé un corpus bilingue anglais/irlandais en LS qui contient 350 phrases (Morrissey, Somers, & Smith, 2010). San-Segundo a créé un corpus de LS espagnole et la langue espagnole écrite dans le domaine des documents d'identité et de renouvellement du permis de conduire (San-Segundo et al., 2010). Le corpus comprend 1360 phrases en LS espagnole, et plusieurs variantes de traduction en espagnol ont été produites pour chaque phrase, ce qui a conduit à un total de 4080 paires de phrases.

Les corpus vidéo ne donnent pas accès aux données de profondeurs, d'où l'intérêt de constituer des corpus en capture de mouvement.

2.2.1.1.2 Corpus 3D

Les corpus captés à l'aide de mocap sont utilisés pour des analyses linguistiques et pour le traitement automatique, ainsi que pour l'analyse des mouvements. Même si ces corpus fournissent des données 3D et nécessitent moins de traitement que les corpus vidéo, il y a encore peu de corpus 3D. La raison principale pourrait être la disponibilité et l'accessibilité des systèmes mocap qui restent chers pour les laboratoires de recherche, ainsi que sa complexité d'utilisation pour enregistrer la LS. Dans cette partie, nous listons les corpus 3D de LS qui existent et le matériel utilisé pour leur capture.

En ce qui concerne l'ASL il existe trois corpus 3D. Le premier corpus CUNY qui a été utilisé dans plusieurs études (Huenerfauth & Lu, 2010; Lu & Huenerfauth, 2010, 2011, 2012, 2014). Il a été enregistré par le système Animazoo IGS-190 pour capturer le mouvement des bras et du torse, et le système Intersense IS-900 pour capturer le mouvement de la tête, en plus deux gants Immersion Cyber et un eye tracker pour capturer respectivement les mouvements des mains et des yeux. Le corpus a été enregistré avec 9 sujets sourds et sa durée totale est 215 minutes. Tandis que le deuxième corpus (Malaia et al., 2008) est relativement petit. Il a été enregistré avec le système Gypsy 3.0. Un sujet bilingue natif a été enregistré, qui produit un groupe de 50 signes. Tyrone et ses collaborateurs ont enregistré aussi des données 3D de l'ASL (Tyrone, Nam, Saltzman, Mathur, & Goldstein, 2010). Ces données ont été enregistrées à l'aide de six caméras du système de capture de mouvement optique Vicon, avec une fréquence de 100Hz.

Un petit corpus 3D de la LS finlandaise (FinSL) a été constitué (Jantunen, Burger, Weerdt, Seilola, & Wainio, 2012). Il se compose d'un monologue et d'un dialogue, il a été enregistré en utilisant huit caméras de capture de mouvement optique (Qualisys ProReflex MCU120), avec une fréquence de 120Hz. Ce corpus est très court, il dure environ 2,5 minutes.

Pour la LSF, il existe le corpus 3D (Duarte & Gibet, 2010) qui a été créé dans le cadre du projet SignCom⁹. Il a été enregistré à l'aide de douze caméras du système de capture de mouvement optique (Vicon MX), avec une fréquence d'enregistrement de 100Hz. Il n'est malheureusement pas disponible pour la communauté scientifique. Un autre corpus 3D a été créé dans le cadre du projet MarqSpat¹⁰, en plus de la LSF il contient aussi la ASL, et de la langue des signes québécoise (LSQ). Ce corpus non plus n'est pas disponible pour la communauté scientifique. Le dernier corpus 3D de LSF a été constitué dans le cadre du projet CIGALE¹¹ à l'aide du système Vicon et une fréquence d'enregistrement de 120 Hz (Catteau, Blondel, Vincent, Guyot, & Boutet, 2016). Ce dernier n'est pas encore disponible et les données n'ont pas encore été nettoyées.

Les expressions faciales sont importantes pour transmettre des informations grammaticales et prosodiques dans la LS, leur présence ou absence a un impact sur la compréhensibilité et la qualité perçue des animations de la LS. Huenerfauth et ses collaborateurs ont évalué expérimentalement les animations de la ASL avec et sans différents types d'expressions faciales et ont constaté que leur inclusion mène à des avantages mesurables pour la compréhensibilité et la qualité perçue des animations (Huenerfauth, Lu, & Rosenberg, 2011). D'autres études ont montré l'importance des expressions faciales dans la LS où ils se sont focalisés sur les mouvements des sourcils dans des questions interrogatives (Schnepp, Wolfe, McDonald, & Toro, 2013; Wolfe, Cook, McDonald, & Schnepp, 2011).

Ainsi, pour pouvoir mener des études variées sur la LSF à l'aide d'un corpus 3D, il nous faut créer ce corpus de manière à ce qu'il soit d'une taille suffisante et avec un contenu varié. Il faut aussi enregistrer les mouvements des composantes du visage. Enfin, l'annotation doit être

⁹ <http://www-valoria.univ-ubs.fr/signcom/en/>

¹⁰ <https://www.irit.fr/marqspat/index.html>

¹¹ <http://cigale.labex-arts-h2h.fr/>

effectuée à l'aide d'ANVIL¹², seul logiciel d'annotation qui permet de synchroniser les enregistrements vidéo et 3D.

2.2.1.2 Annotation

Un corpus est constitué de données primaires constituées par les enregistrement (vidéo ou mocap pour les LS) et de données secondaires constituées d'annotations. L'annotation consiste à rajouter des informations à des fins d'analyse. Par exemple, un type commun d'annotation est l'ajout d'étiquettes indiquant la description fonctionnelle, syntaxique, sémantique d'un élément du corpus ou d'une glose.

Le schéma d'annotation définit la structure de l'annotation. Il est à constituer avant l'annotation, car il rassemble toutes les valeurs et les phénomènes à annoter, ainsi que les relations entre les différents niveaux d'annotation. Le schéma d'annotation doit prendre en compte plusieurs paramètres, notamment : le type de données, le domaine de recherche dans lequel le corpus va être utilisé, le logiciel d'annotation et les possibilités qu'il donne, ainsi que la manière dont le corpus a été capturé.

Dans le schéma, plusieurs pistes d'annotation peuvent être définies, chacune spécialisée pour un type d'annotation (linguistique, phonétique, sémantique, lexicale, mouvement...etc.). Une piste peut être complémentaire à une autre. Les pistes contiennent les segments d'annotation. Ces segments sont définis par un temps de début, un temps de fin et des valeurs d'attributs. Les attributs sont les données à mettre dans les pistes, ils peuvent être des textes, des valeurs numériques, des valeurs booléennes, ou même un lien vers une autre annotation. Le livre de Braffort et ses collaborateurs donne plus de détails sur l'annotation de la LSF (Annelies. Braffort et al., 2016).

Plusieurs outils d'annotation sont disponibles gratuitement. ELAN¹³, ILEX¹⁴ et ANVIL, sont les plus utilisés pour les LS et proposent presque les mêmes fonctions. Rohlfing a comparé les logiciels d'annotation les plus utilisés (Rohlfing et al., 2006). L'avantage d'ELAN est qu'il permet de synchroniser deux vidéos ou plus en même temps. Dans les corpus vidéo de LS, généralement, il y a plus qu'une vidéo (plusieurs caméras de face de côté, etc..), ce fait rend ELAN l'outil le plus utilisé dans l'annotation des corpus de LS en plus de sa facilité

¹² <http://www.anvil-software.org/>

¹³ <https://tla.mpi.nl/tools/tla-tools/elan/>

¹⁴ <https://www.sign-lang.uni-hamburg.de/ilex/>

d'utilisation. ILex a pour particularité de permettre une annotation distribuée entre plusieurs annotateurs, car il est implanté sous la forme d'un serveur qui centralise les diverses annotations. Il est de ce fait plus complexe à installer et à utiliser. Quant à ANVIL, il permet de visualiser et de synchroniser les données de capture de mouvement visualisées à l'aide d'un squelette 3D avec la vidéo. Cette fonction le rend plus adapté dans le cadre de l'annotation d'un corpus de données 3D.

2.2.1.3 Bilan

Dans cette section nous avons listé quelques corpus de LS représentatifs. Nous avons ainsi vu qu'il existe beaucoup de corpus vidéo. Les corpus vidéo peuvent être classés en plusieurs catégories selon leur utilisation. Il existe des corpus destinés à servir les lexiques, des corpus destinés à la recherche linguistique, des corpus adaptés pour le traitement du langage naturel et/ou des tâches de vision par ordinateur.

Contrairement au corpus vidéo, peu de corpus 3D existent en raison du manque de disponibilité et d'accès à des systèmes mocap pour les laboratoires de recherche, ainsi que la complexité de l'utilisation de la mocap pour capter la LS. Le seul « grand » corpus de ce type est le corpus CUNY qui contient moins de 4h d'enregistrement. Les autres corpus 3D ne contiennent que quelques minutes d'enregistrement. Contrairement au corpus CUNY qui a été créé dans le but de créer un corpus de ASL assez représentatif, les autres corpus ont été enregistrés pour des études spécifiques plus limitées.

Le nombre et la disposition des marqueurs pour ces corpus est différent, chaque disposition a été établi selon le besoin de l'étude. Les seuls corpus qui prennent en compte le visage sont celui de Duarte et Gibet (Duarte & Gibet, 2010) dont l'objectif est d'utiliser les données pour animer un signeur virtuel, ainsi que celui de Jantunen (Jantunen et al., 2012).

Ainsi, les corpus existants ne sont pas disponibles ou ne répondent pas aux exigences de la recherche multidisciplinaire que l'on souhaite mener. Cela implique de créer un corpus 3D de LSF d'une taille suffisante et avec un contenu varié.

2.2.2 Contrôle moteur

Cette section présente trois lois du contrôle moteur, respectivement le principe d'isochronie (section 2.2.2.1), la loi des asymétries (section 2.2.2.2) et la loi de puissance 2/3 (section 2.2.2.3), puis un bilan (section 2.2.2.4). L'idée est d'étudier ces lois connues du contrôle moteur dans le contexte de la langue des signes.

2.2.2.1 Le principe d'isochronie

Le principe d'isochronie stipule que la vitesse du mouvement est proportionnellement liée à l'augmentation de la distance de mouvement de manière à maintenir la durée du mouvement à peu près constante sur une plage de distance de déplacement (P. Viviani & McCollum, 1983). La figure 2-6 montre bien l'effet de la distance sur la vitesse moyenne (Sartori, Camperio-Ciani, Bulgheroni, & Castiello, 2013), où nous observons dans (A) que la vitesse de pointe augmente avec la distance, et dans (B) nous observons que la vitesse de pointe augmente et que cela a pour effet de préserver un temps de déplacement globalement constant. Ce type de régularité temporelle dans le comportement du moteur humain a été trouvé assez tôt dans la littérature (Binet & Courtier, 1893; Derwort, 1938; Fitts, 1954; Freeman, 1914) et il a été vérifié dans une variété d'actions.

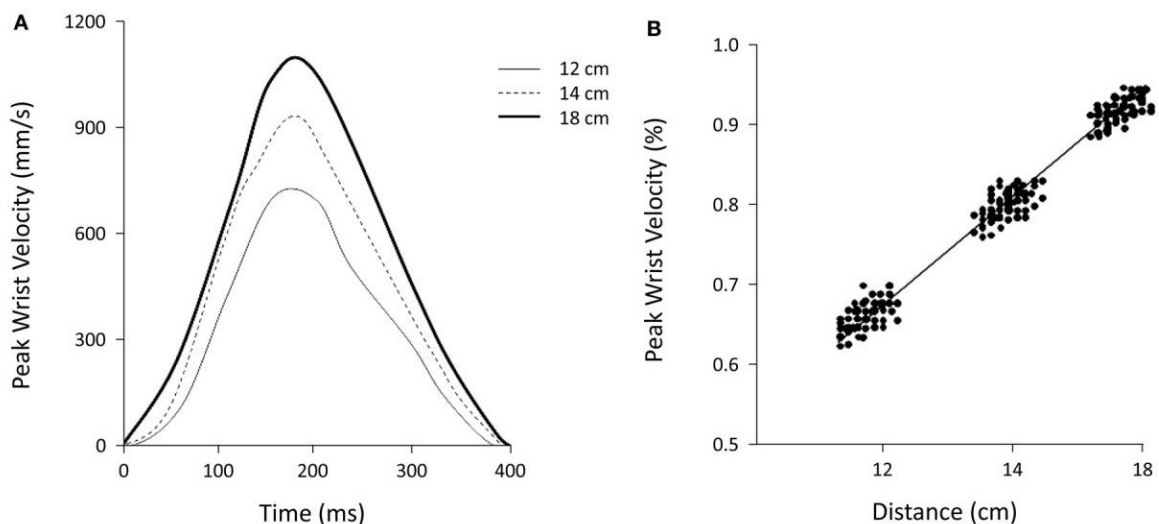


Figure 2-6 : (A) Vitesse moyenne du poignet pour pointer des objets situés à des distances différentes. (B) Relation entre la vitesse moyenne du poignet et la distance par rapport à la cible (Sartori et al., 2013).

Les études sur les mouvements d'écriture, par exemple, ont montré qu'écrire une lettre ou un mot avec des tailles différentes implique des changements proportionnels à la vitesse (Lacquaniti, Terzuolo, & Viviani, 1983; Michel, 1971). Cette relation entre la distance d'un mouvement et la vitesse semble être une caractéristique assez commune se rapportant non seulement à l'écriture, mais aussi à une variété d'actions telles que la dactylographie (P. Viviani & Terzuolo, 1982), le mouvement de dessin (P. Viviani & McCollum, 1983; P. Viviani & Schneider, 1991), le mouvement de la levée de poids (Gachoud, Mounoud, Hauert, & Viviani, 1983), l'activité de coup de pied chez les nourrissons (Thelen & Fisher, 1983), et les mouvements de la main et du bras (Freund & Büdingen, 1978; Jeannerod, 1984). Cette relation

a été aussi vérifiée dans les mouvements de saisie des objets pendant le comportement quotidien normal des macaques crabiers (singes) (Sartori et al., 2013).

Dans tous les cas, la constance relative de la durée de mouvement résulte directement du fait que la vitesse moyenne augmente avec la distance de la trajectoire. En particulier, en changeant l'échelle à laquelle on trace une trajectoire donnée, cela produit un changement semblable de la vitesse moyenne.

Cependant, il a été montré que l'augmentation de la vitesse avec l'amplitude d'un geste n'est en fait pas suffisante pour préserver une durée constante. Sur des gestes plus amples par exemple, il a été montré clairement que non seulement la vitesse mais aussi la durée augmentaient avec la distance parcourue (Berret & Jean, 2016; Young, Pratt, & Chau, 2009). Sur des distances allant de quelques centimètres à plusieurs dizaines de centimètres, la relation est globalement affine. Le principe d'isochronie a donc une portée limitée et il faudra surtout retenir en général qu'il existe des relations affines croissantes entre distance et durée ainsi qu'entre distance et vitesse moyenne pour les mouvements du membre supérieur.

2.2.2.2 Loi des asymétries verticales

L'accélération gravitationnelle est omniprésente dans nos actions de la vie de tous les jours et joue un rôle important dans plusieurs fonctions. Par exemple, l'action permanente sur les organes otolithes du système vestibulaire fournit au SNC des informations précieuses sur l'orientation spatiale, la perception visuelle, et le contrôle de l'équilibre (Lackner & DiZio, 2000; Pozzo, Papaxanthis, Stapley, & Berthoz, 1998). La gravité influe également sur l'élaboration de mouvement. Des études ont montré que le SNC profite de la force de gravité d'une manière optimale pendant les mouvements du bras ou du corps entier (Berret et al., 2008; Berret, Chiovetto, Nori, & Pozzo, 2011; Crevecoeur, Thonnard, & Lefèvre, 2009; Papaxanthis, Dubost, & Pozzo, 2003), différenciant ainsi le plan moteur d'un mouvement dirigé vers le haut de celui d'un mouvement dirigé vers le bas. En outre, il a été rapporté que le SNC utilise un modèle interne de la gravité pour compléter l'information sensorielle lors de l'estimation du temps de contact avec un objet approchant (Indovina et al., 2005; McIntyre, Zago, Berthoz, & Lacquaniti, 2001). Ainsi, il a été montré que le SNC met en œuvre des solutions optimales selon le contexte dynamique de l'action (Gaveau et al., 2014; Gaveau, Paizis, Berret, Pozzo, & Papaxanthis, 2011).

L'exécution des mouvements que l'on maîtrise nécessite la représentation interne de l'interaction du corps avec le monde extérieur. L'étude des mouvements verticaux du bras offre

un paradigme intéressant pour comprendre comment le système moteur se comporte avec la force de gravité et quels critères sont appliqués lors de l'élaboration d'un mouvement. Par exemple, si les trajectoires de la main (le chemin de la trajectoire ou le profil de vitesse) sont équivalentes pendant les mouvements ascendants et descendants (c.à.d., sous divers effets de la gravité), cela peut indiquer un plan moteur purement cinématique qui intègre précisément les couples de gravité pour préserver la cinématique du bras.

D'autre part, des changements significatifs dans la cinématique du bras selon la direction du mouvement peuvent indiquer l'existence d'un processus de planification dynamique qui profite des forces externes agissant sur le membre au détriment de l'invariance de la trajectoire de la main. En effet, des études ont rapporté que la cinématique différait entre les mouvements ascendants et descendants, argumentant ainsi en faveur d'un plan dynamique. Plus précisément, pour divers mouvements de durée et d'amplitude équivalentes (pointage, atteindre une cible, dessin etc.) la durée d'accélération est supérieure pendant les mouvements descendants comparée aux mouvements ascendants (Gentili, Cahouet, & Papaxanthis, 2007; Papaxanthis, Dubost, et al., 2003; Papaxanthis, Pozzo, & Schieppati, 2003; Papaxanthis, Pozzo, & Stapley, 1998; Papaxanthis, Pozzo, Vinter, & Grishin, 1998).

L'étude de Papaxanthis et ses collaborateurs sur le mouvement de pointage (Papaxanthis, Pozzo, & Stapley, 1998) a montré que le temps de mouvement était approximativement équivalent dans les deux directions (ascendant et descendant) et dans des conditions de charges et de vitesse différentes (figure 2-7).

La caractéristique la plus importante de la cinématique de cette étude était que les temps d'accélération et de décélération étaient inégaux, indiquant une asymétrie de la production de mouvement (voir figure 2-7). La figure 2-8 illustre qualitativement les effets de la direction du mouvement sur les profils de vitesse des doigts, en moyenne pour tous les sujets.

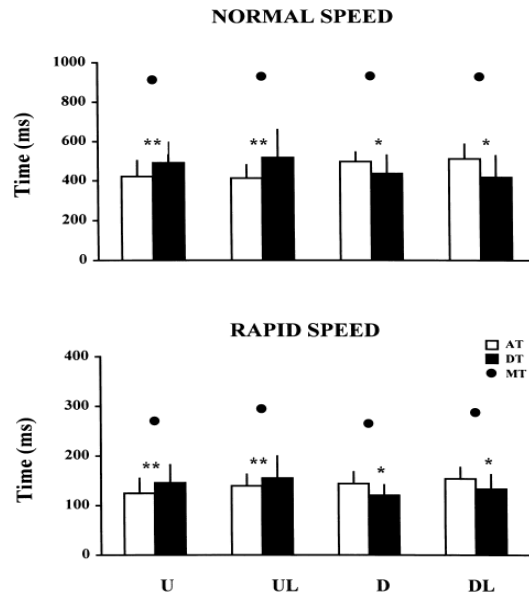


Figure 2-7: Histogramme montrant les effets de la direction du mouvement, la vitesse et la charge sur le temps d'accélération (AT), le temps de décélération (DT) et le temps de déplacement (MT). (Papaxanthis, Pozzo, & Stapley, 1998). U : mouvement ascendant, UL : mouvement ascendant avec une charge à la main, D : mouvement descendant et DL : mouvement descendant avec une charge à la main.

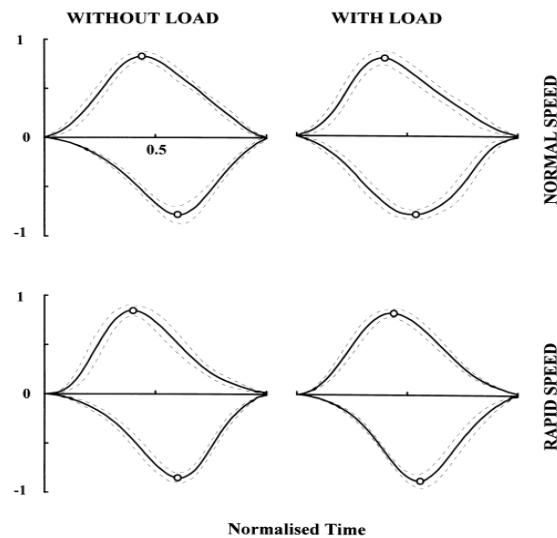


Figure 2-8: Profils de vitesse normalisée du mouvement du doigt dans des mouvements de pointage ascendants et descendants. Les mouvements ascendants sont représentés par des mouvements positifs et descendants par des profils de vitesse négatifs (Papaxanthis, Pozzo, & Stapley, 1998).

De plus, des expériences de vols paraboliques ont révélé que l'exposition à de nouveaux environnements gravitationnels (microgravité et hyper-gravité) modifier progressivement cette asymétrie directionnelle, ce qui suggère que la force de gravité est intégrée dans le processus de planification central et que les asymétries ne sont pas seulement dues à l'influence de groupes

musculaires différents (Crevecoeur et al., 2009; Gaveau, Berret, Angelaki, & Papaxanthis, 2016; Gaveau et al., 2011; Papaxanthis, Pozzo, & McIntyre, 2005).

2.2.2.3 Loi de la puissance 2/3

Il a été établi à la fin du XIXe siècle que, pour les mouvements des bras, la courbure et la vitesse sont des variables corrélées, la vitesse suit généralement une relation inverse à la courbure (Jack, 1895). Cette règle a été quantitativement formalisée comme la loi de puissance 2/3 (deux-tiers) ou 1/3 (un-tiers) suivant la formulation cartésienne ou angulaire de la loi. Cette règle dit que lorsqu'on dessine une figure, comme une ellipse ou un huit, la vitesse instantanée varie avec la courbure de la trajectoire selon une loi de puissance. Plus précisément, la vitesse de déplacement augmente dans les parties peu courbées de la trajectoire et inversement, elle diminue dans les parties courbées (Lacquaniti et al., 1983). Elle peut être formulée comme suit :

$$v(t) = \text{const} * k(t)^\beta \quad (1)$$

Ou, dans l'espace logarithmique :

$$\log(v(t)) = \text{const} + \beta \log(k(t)) \quad (2)$$

Expérimentalement β est environ égal à $-\frac{1}{3}$, et l'équation (1) devient :

$$v(t) = \alpha * k(t)^{-\frac{1}{3}} \quad (3)$$

où v est la vitesse euclidienne, k est la courbure et α est le facteur de gain de vitesse, qui est constant par morceaux sur des segments ou parties du mouvement voire constant sur tout le mouvement.

La figure 2-9 montre le tracé de la vitesse tangentielle $v(t)$ par rapport à la courbure $k(t)$ dans une échelle logarithmique. Selon l'équation 3, si le facteur de gain de vitesse α est constant par morceaux, les points de données se regrouperont le long d'un segment qui a une pente (l'exposant β) autour de $-1/3$.

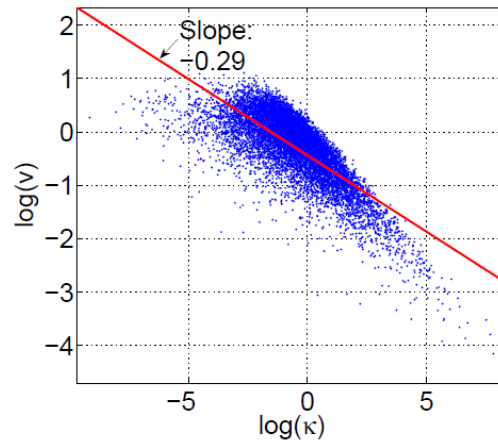


Figure 2-9: Courbe $\log(k)$ versus $\log(v)$. La valeur de l'exposant β mesurée à partir de cette courbe est $\beta = -0.29$, reproduisant la valeur théorique (Maoz, Portugaly, Flash, & Weiss, 2005).

Des recherches approfondies ont étudié les conditions et les origines de la loi de la puissance $2/3$. L'équation (3) a d'abord été développée pour les dessins en deux dimensions, mais a également été appliquée plus tard pour le dessin en trois dimensions dans des conditions de force isométrique (Massey, Lurito, Pellizzer, & Georgopoulos, 1992) et à diverses modalités de mouvement, y compris des mouvements de poursuite oculaire (De'Sperati & Viviani, 1997) et les mouvements de la parole (Perrier & Fuchs, 2008; Tasko & Westbury, 2004). En outre, l'exposant de la loi de la puissance $2/3$ varie chez les enfants et devient plus stable avec l'âge (P Viviani & Schneider, 1991). Ce principe a été appliqué aussi à la perception du mouvement : il a reçu l'appui des études sur la perception de l'écriture et des mouvements de dessin (Kandel, Orliaguet, & Viviani, 2000; Soechting, Lacquaniti, & Terzuolo, 1986) et du mouvement de stimuli visuels abstraits (Levit-Binnun, Schechtman, & Flash, 2006; P Viviani & Stucchi, 1992). Enfin, une étude d'imagerie par résonance magnétique fonctionnelle (IRMf) a mis en évidence une représentation centrale de la perception de cette loi cinématique (Dayan et al., 2007). Ces études montrent que la loi de la puissance $2/3$ n'est très probablement pas l'expression de contraintes biomécaniques mais reflète certainement l'implication du SNC (Tamar Flash, Meirovitch, & Barliya, 2013).

2.2.2.4 Bilan

Les lois classiques du contrôle moteur (LCCM) ont été vérifiées dans plusieurs contextes comme dans l'écriture, le dessin, le pointage, le mouvement de la levée de poids etc. Elles ont même été vérifiées dans des conditions de gravité spécifique (microgravité et hyper-gravité) (Crevecoeur et al., 2009; Papaxanthis et al., 2005) et même dans les mouvements des Macaques

crabier¹⁵ (Sartori et al., 2013). Nous savons aussi que certaines pathologies neuro-motrices peuvent affecter ces lois nominales du mouvement humain (Paolo Viviani, Burkhard, Chiuvé, Dell'Acqua, & Vindras, 2009). Malgré cette variété d'études sur les LCCM, aucune étude à notre connaissance n'a vérifié si ces lois restaient valables pour les mouvements portant un sens linguistique (sauf l'étude d'Endres qui a utilisé la structure de la loi de la puissance 2/3 pour segmenter la LS (Endres, Meirovitch, Flash, & Giese, 2013), mais comme un a priori). La démarche classique a d'ailleurs été de supposer que les lois connues en contrôle moteur s'appliquaient directement en LS, où Lebourque dans ces travaux de thèse s'est inspiré de ces théories en espérant de générer des gestes naturels (Lebourque, 1998). Or, il est tout à fait possible d'observer un éloignement des LCCM en LS puisqu'un mouvement portant un message linguistique doit certainement avoir des trajectoires ayant des caractéristiques spatiales et temporelles particulières et qu'une séquence de LS est nécessairement constituée de signes et de transitions entre les signes qui doivent être bien distinguables pour l'autre locuteur.

2.2.3 Le mouvement et la LS

Le mouvement est un paramètre inhérent à la LS, or ce paramètre n'a été que rarement étudié. Parmi ces rares études, il y a celles qui ont utilisé le mouvement pour la segmentation semi-automatique de la LS. (Lefebvre-Albaret & Patrice Dalle, 2009) ont fait une étude qui utilise des caractéristiques de bas niveau du mouvement afin de segmenter semi-automatiquement les signes. Ils ont étudié les mouvements dans le but d'identifier plusieurs types de symétries. Cette méthode s'est révélée efficace pour résoudre la segmentation de haut niveau. Par ailleurs, Gonzalez a présenté une autre étude sur la segmentation semi-automatique afin de permettre d'identifier le début et la fin des signes, elle a pris en considération les caractéristiques du mouvement ainsi que la forme des mains (Gonzalez, 2012).

Des modèles de mouvement ont été déterminés ayant pour objectif d'être utilisés lors de la synthèse de l'animation. Segouat et Braffort ont examiné la façon de produire des mouvements de coarticulation précis entre les signes (Segouat & Braffort, 2009). Ils ont analysé les mouvements humains d'un corpus vidéo de LSF pour déterminer des modèles mathématiques de mouvement mais très simplifiés car en 2D. Toro a modélisé l'incorporation de mouvements répétitifs indiquant des variations dans l'aspect temporel (répétition, régularité, etc.) pour les verbes avec plusieurs objets (Toro, 2004, 2005). Cette modélisation a été faite en se basant sur

¹⁵ Le macaque crabier (*Macaca fascicularis*) est un singe catarhinien de la famille des cercopithécidés, originaire d'Asie du Sud-Est.

des vidéos de verbes en ASL. Ainsi, il a analysé ces vidéos pour annoter l'emplacement des mains pour ces différents verbes. De plus, il a proposé des algorithmes pour la planification de trajectoires des mains pour chaque verbe.

McDonald et ses collaborateurs ont proposé une nouvelle approche dans le contexte de la génération automatique du changement de rôle par un signeur virtuel (McDonald et al., 2013; McDonald, Wolfe, Moncrief, & Baowidan, 2014; Wolfe, McDonald, Moncrief, Baowidan, & Stumbo, 2015). Cette approche utilise les informations linguistiques pour calculer les orientations globales de la posture du signeur virtuel qui assume un rôle et les mouvements des différents articulateurs du corps (les yeux, les épaules, la colonne vertébrale, etc.).

L'apparition et l'utilisation des corpus 3D en LS ces dernières années ont permis de prendre en considération le mouvement de la LS. Lefebvre-Albaret dans sa thèse a analysé les profils de vitesses des mouvements balistiques et des mouvements balistiques répétés issus des données 3D (Lefebvre-Albaret, 2010). Une autre étude utilisant des données 3D de LSF a été consacrée à la segmentation automatique du mouvement de la main sur la base de l'analyse en composantes principales (ACP) (Heloir et al., 2006). Malia a analysé le mouvement d'un corpus 3D de la ASL pour identifier le type de verbe, sur la base des paramètres de vitesse et d'accélération du mouvement (Malaia et al., 2008; Malaia & Wilbur, 2012), suggérant une modification contrôlée de la cinématique des mouvements de LS pour leur donner un sens linguistique, en l'occurrence le fait qu'une action soit complète (verbe téléique) ou pas (verbe atélique).

Plusieurs études ont montré que le mouvement est un paramètre important dans la prosodie en LS. Wilbur a indiqué que la réalisation du mouvement pour un signe donné peut varier en fonction de nombreux facteurs, tels que le rythme, l'accent, ou le contexte prosodique (R. B. Wilbur, 1999, 2009). Grosjean a constaté que l'augmentation du rythme de signation causait plus de modifications du mouvement que les autres paramètres phonologiques (Grosjean, 1979). De même, Wilbur et Schick ont constaté que le signe souligné peut être produit avec une plus grande amplitude de mouvement ou de vitesse (R. Wilbur & Schick, 1987).

Par ailleurs, Tyrone a étudié la prosodie des signes en regardant le timing du mouvement de la main dominante, comme un effet de la position du signe dans une phrase, en utilisant des données 3D (Tyrone et al., 2010). Ils ont examiné deux simples types de mouvements (mouvement vers le contact avec un emplacement sur le corps, et le mouvement d'éloignement du contact avec un emplacement sur le corps) à travers de multiples positions dans une phrase.

Bien que la LS soit une langue fondée sur le mouvement, la plupart des études en traitement automatique de la LS se sont intéressées à modéliser des propriétés linguistiques, et se sont concentrées sur les mains et les expressions faciales, mais peu se sont intéressées à comprendre la cinématique et la dynamique du mouvement des bras. Or dans l'optique d'une modélisation informatique et d'outils de production ou de traduction automatique de LS, il semble crucial de se poser ces questions fondamentales sur les paramètres cinématiques ou dynamiques qui sont modifiés en LS comparés au contrôle moteur lambda. Les rares études qui ont utilisé le mouvement dans des analyses linguistiques et notamment dans la prosodie sont basées sur des analyses de corpus vidéo, ce qui ne permet pas d'estimer des vitesses et des accélérations avec précision et fiabilité. Les résultats doivent donc être confirmés avec des analyses 3D.

2.2.4 La Langue de Signes et les informations pertinentes pour la compréhension

Cette section présente la perception de la LS. Après avoir présenté la prise d'information en LS, la privation et la compréhension (section 2.2.4.1), nous présentons la perception des signeurs virtuels (section 2.2.4.2).

2.2.4.1 Prise d'information en Langue des Signes, privation et compréhension.

Chez l'Homme, la capacité du système visuel à traiter les informations en provenance de l'environnement à chaque moment dans le temps serait limitée (Broadbent, 1958). Ainsi, la prise d'information d'un individu dans l'environnement peut nécessiter une focalisation de son attention sur un ou quelques points particuliers. Il est alors fait appel à des mécanismes d'attention visuelle sélective de manière à sélectionner, « filtrer » et « extraire » du flot d'informations visuelles, les stimuli pertinents du point de vue de l'individu. Hauthal et les autres auteurs ont mené une étude pour tester les capacités d'attention visuelle chez des individus sourds et non-sourds (Hauthal, Neumann, & Schweinberger, 2012). Les participants devaient évaluer si le symbole « cible » (i.e., une silhouette) était « masculin ou féminin », dans des situations mettant en jeu plus ou moins de distracteurs (e.g., un visage masculin ou féminin, un objet à connotation masculine ou féminine) et de symboles sans signification (i.e., « remplisseurs »). Leurs résultats semblent montrer que les capacités d'attention visuelle en situation de charge perceptuelle élevée sont plus importantes pour des individus sourds que pour des individus non-sourds. Pour ces auteurs, ce serait la privation d'un sens qui pourrait amener ces capacités. La production de la LS implique une grande partie du corps de l'individu : le visage, les yeux, les bras, les mains, les doigts, le buste deviennent des moyens de communiquer une information. Ainsi la nécessité de focaliser son attention visuelle sur des

éléments particuliers et ne pas être distrait par d'autres devient primordial pour les personnes pratiquant la LS.

Comprendre quels sont les éléments perçus et retenus lors de l'observation de la LS semble donc intéressant notamment pour l'apprentissage de cette langue mais aussi pour la création des signeurs virtuels compréhensibles par un large public.

L'utilisation des systèmes d'oculométrie (Eye tracking), mesurant l'orientation du regard d'un individu observant une scène, a permis d'acquérir des informations sur le placement du regard lors de l'observation d'un discours signé (Emmorey, Thompson, & Colvin, 2009; Muir & Richardson, 2005; Watanabe, Matsuda, Nishioka, & Namatame, 2011). Les résultats de ces études convergent vers l'idée que l'attention visuelle d'un observateur ayant une expérience en LS orienterait davantage son regard au niveau des yeux et de la bouche que ce soit lors de l'observation d'une personne qui signe ou de l'évaluation d'une émotion sur un visage statique. Les individus débutant la LS fixeraient également 80% du temps la tête du signeur avec une plus grande attention vers la bouche (Emmorey et al., 2009). Cependant, il faut savoir que certaines LS emploient plus que d'autres la labialisation, qui consiste à articuler avec la bouche les mots de la langue parlée correspondant au signe effectué. Il s'avèrerait également que les individus entendants orienteraient également leur regard vers ces mêmes zones du visage pour l'évaluation d'une émotion.

Le déplacement du regard vers les mains de la personne observée semble se faire uniquement lorsque le narrateur regarde lui-même ses mains pour les individus pratiquant la LS. Il apparaît donc des stratégies particulières dans la prise d'informations lors de l'observation de la LS par des personnes sourdes. Ainsi, un maintien de l'attention visuelle au niveau du visage semble être plus important chez des individus sourds par rapport à des individus entendants lors de l'observation de la LS (Heimler et al., 2015)

Des études sur la perception de points lumineux directement associées à une action motrice humaine ont montré que l'Homme présentait une capacité spécifique (car automatique et rapide) à détecter un mouvement biologique (Johansson, 1973; Neri, Morrone, & Burr, 1998; Poom & Olsson, 2002). Des études antérieures ont montré que l'Homme était même capable de reconnaître un grand nombre de caractéristiques sociales par le biais de ces quelques points, telles que : le sexe (Catharine D. Barclay, James E. Cutting, 1978), les intentions (Runeson & Frykholm, 1983), les émotions (Atkinson, Dittrich, Gemmell, & Young, 2004), etc.

La capacité d'un observateur sourd à reconnaître la LS à partir de points lumineux a également été démontrée dans différentes études (Campbell et al., 2011; Tartter & Fischer, 1982; Tartter & Knowlton, 1981). Ces études semblent montrer un discernement minimum de contrastes dans le mouvement, l'orientation, la localisation des mouvements des mains notamment. La précision de ces paramètres était moins bonne dans les conditions où seuls des points lumineux sont présentés (informations réduites au minimum) comparées à des situations réelles (quantité d'information maximale). Ainsi, il a été mis en évidence que les individus aussi bien sourds et signeurs qu'entendants et non signeurs, arrivaient à distinguer des émotions (peur, joie, neutre) uniquement à partir de mouvements des mains en LS. Ainsi, des individus entendants et non signeurs semblent être capables d'identifier certaines émotions véhiculées par des mouvements de LS. Comme nous l'avons vu précédemment, lors de la perception de mouvements biologiques, les invariants dynamiques pourraient fournir des informations sur l'état émotionnel de l'individu (Atkinson et al., 2004).

L'apparition des signeurs virtuels (ou « avatars signants ») dans divers domaines (ex. Education, Information, Jeux) se développe actuellement. En effet, un des intérêts applicatifs de l'utilisation d'un avatar pour la diffusion d'informations dans les lieux publics par exemple est de ne pas être confronté à la question de droit d'image et de l'anonymat de la personne qui signe, contrairement à la sollicitation de personnes réelles. De plus, dans le domaine de l'apprentissage, l'utilisation de vidéos mettant en scène des avatars signeurs permet notamment de manipuler différentes informations précises afin de faire varier le niveau de difficulté de compréhension du message. Cela nous amène à nous poser la question de savoir si les individus sourds sollicitent les mêmes processus perceptifs et les mêmes stratégies de prise d'informations lorsqu'ils sont confrontés à ce type de stimulus, en comparaison à des situations plus écologiques (observation d'un individu signeur). Certaines études ont porté sur cette question, comme nous le verrons dans la partie qui suit.

2.2.4.2 Perception des signeurs virtuels

Les signeurs virtuels sont des incarnations en trois dimensions (3D) de personnages virtuels délivrant un message en LS. Des programmes informatiques permettent de restituer les mouvements de LS en animation à partir de données textuelles. La figure 2-10 montre des signeurs virtuels créés au laboratoire LIMSI utilisés dans des différentes tâches comme la traduction automatique ou la diffusion des messages d'information.



Figure 2-10: Signeurs virtuels créés par le laboratoire LIMSI.

Leur utilisation est multiple et dans des domaines d'application variés : amélioration de l'accès aux contenus écrits pour les personnes sourdes, traduction sur internet, information en gare, jeux éducatifs, etc. Cependant les principaux problèmes rencontrés aujourd'hui semblent être liés à la fluidité des mouvements, la qualité des signes manuels proposés ainsi que la lisibilité des expressions du visage (Efthimiou et al., 2010; Kipp, Heloir, & Nguyen, 2011). Des stratégies pour pallier à ces problèmes émergent progressivement et notamment dans la manière d'animer les avatars à partir de la mocap (Alexanderson & Beskow, 2015). L'intérêt de cette dernière stratégie est qu'elle permet d'animer un avatar selon un mouvement biologique, réellement produit par un individu. Les mouvements sont donc plus fidèles à la réalité cinématique et cinétique de la LS.

Nous allons maintenant présenter d'une part des études portant sur l'impact de la qualité du mouvement des avatars sur la compréhension (section 2.2.4.2.1) et d'autre part celles sur l'impact de la forme (section 2.2.4.2.2).

2.2.4.2.1 Impact des informations liées au mouvement

Dans une étude utilisant l'oculomètre (Kacorri, Harper, & Huenerfauth, 2013), les auteurs ont analysé, chez des individus nés sourds, la perception d'une vidéo en ASL. L'étude s'est déroulée dans deux conditions : une dans laquelle une personne réelle signe et une seconde dans laquelle un signeur virtuel signe. Leurs résultats semblent montrer que la visualisation de la personne réelle, comparée à l'avatar, entraîne un temps de fixation du visage plus important et moins de déplacement du regard entre le visage et le corps chez les observateurs signeurs. De

plus, il semble exister une corrélation significative entre les patterns de mouvements des yeux de l'observateur sourd et les scores subjectifs (ex. difficulté grammaticale, compréhensibilité) qu'il a lui-même assigné aux animations et aux vidéos. Ainsi, plus la qualité de la vidéo semble élevée (la meilleure étant la vidéo de l'individu réel), plus la quantité de déplacement du regard est faible (« Time Normalized Total Trail Length »).

2.2.4.2.2 Impact des informations liées à la forme

Dans une expérimentation en ligne (figure 2-11), les auteurs ont essayé de comprendre quel pouvait être l'impact de l'apparence de l'avatar sur la compréhension et la perception des usagers (Adamo-Villani, Wilbur, Eccarius, & Abe-Harris, 2009). Ainsi deux modèles d'avatars ont été testés dans cette étude. Leurs résultats indiquent qu'un avatar « segmenté » (figure 2-11A) semble pouvoir être utilisé pour la reconnaissance de signes linguistiques simples, c'est-à-dire ne sollicitant que des mouvements des mains. Cependant lorsque les mains produisent des signes plus complexes, il semblerait que la présence de déformations organiques (e.g., pliure de la peau) puisse augmenter la compréhension, ce qui semble être le cas avec le modèle « homogène » (figure 2-11B).

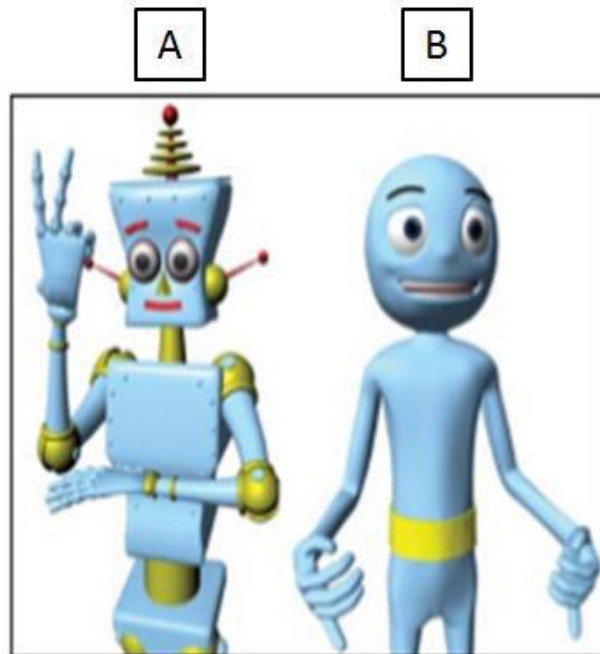


Figure 2-11: Représentation des deux modèles d'avatars présentés dans (Adamo-Villani et al., 2009). A : un modèle « segmenté » ; B : un modèle « homogène ».

La compréhension de la LS à différents niveaux de difficulté (e.g., formes des mains plus ou moins complexes) sur un avatar semble donc être possible même à partir d'avatars moins réalistes (i.e., avatar « segmenté ») et nécessitant moins de travail de développement technique

en amont. Il semble alors intéressant de se demander quelle quantité et qualité minimales d'informations sont nécessaires pour qu'un observateur comprenne encore le message transmis par un avatar signeur.

La manière dont les utilisateurs, à savoir essentiellement les personnes sourdes, perçoivent ces signeurs virtuels est intéressante dans la mesure où ces analyses serviraient à comprendre comment l'Homme interagit avec ces nouveaux types de stimuli mais aussi à améliorer la manière dont ceux-ci sont conçus. Ainsi, pousser l'individu aux limites de ce qu'il est capable de percevoir devrait permettre d'avoir une idée plus précise des éléments nécessaires à une meilleure compréhension de ces animations virtuelles.

2.2.5 Conclusion

Les LS sont des langues naturelles visio-gestuelles qui s'expriment par le mouvement de multiples composantes corporelles.

Si l'on souhaite développer des systèmes de traitement automatique tels que la reconnaissance ou la génération automatique, il est nécessaire de s'appuyer sur des connaissances à la fois sur la manière dont ces langues fonctionnent du point de vue linguistique, mais aussi sur la manière dont les mouvements humains de LS sont produits et perçus au sein d'une théorie de la cognition motrice. En effet, si l'on s'intéresse plus particulièrement à la génération d'énoncés en LSF, c'est-à-dire au contrôle de l'animation d'un signeur virtuel, il faut aussi s'appuyer sur des connaissances relevant de la perception du mouvement d'autrui en LSF puisqu'un signeur virtuel est avant tout créé pour être compris. Il faut donc comprendre non seulement les caractéristiques des mouvements de LSF mais aussi quelles caractéristiques sont cruciales pour la compréhension et devraient être reproduites en priorité par un modèle. Ainsi, cette thèse s'inscrit dans une démarche pluridisciplinaire et inclut des problématiques issues de ces trois domaines, les deux derniers pris en ensemble représentant la théorie de la cognition motrice :

- La linguistique et plus particulièrement la linguistique de corpus, par la création et l'annotation d'un corpus de LSF.
- Le contrôle du mouvement en LSF (analyse des propriétés cinématiques des mouvements de la main propre aux signes).
- La perception du mouvement en LSF (étude de la compréhensibilité des mouvements des signeurs virtuels).

Le chapitre suivant présente le corpus de LSF qui a été créé et annoté dans le cadre de cette thèse.

3 Création et annotation du corpus de la Langue des Signes Française

Nous avons présenté dans le chapitre précédent l'état des lieux pour les corpus de LS, qui a conclu sur le manque de données 3D pour la LSF. Or les corpus sont indispensables pour faire des études linguistiques et scientifiques de la LS. Les données 3D sont requises pour faire des études quantitatives fiables. Dans le domaine de la LSF, la création d'un corpus 3D ouvert à la communauté constitue donc une production scientifique à part entière. Nous avons donc constitué un corpus de données 3D de LSF (nommé MOCAP1). Dans ce chapitre, nous allons détailler la procédure de création (partie I) et d'annotation du corpus (partie II).

3.1 Création du Corpus MOCAP1

Le corpus MOCAP1 est un corpus de données 3D de la LSF qui est destiné à la recherche pluridisciplinaire, notamment pour les études linguistiques, informatiques et en sciences du mouvement. Le corpus a un contenu riche constitué de plusieurs tâches, il est enregistré avec un système mocap et une vidéo. Cette partie décrit le contenu du corpus (section 3.1.1), le matériel utilisé (section 3.1.2), la disposition des marqueurs (section 3.1.3) et le positionnement optimal des caméras (section 3.1.4). Elle présente aussi le processus de calibration (section 3.1.5), d'enregistrement (section 3.1.6), de post-traitement (section 3.1.7) ainsi que l'exportation des données (section 3.1.8).

3.1.1 Contenu du corpus

Huit signeurs¹⁶ sourds (4 hommes et 4 femmes) ont été invités à participer à ce corpus. Ils ont différents profils sociolinguistiques et différents styles dans leur manière de signer, afin d'avoir un corpus comportant des variabilités inter-individuelles.

Afin de fournir des données pouvant être utilisées dans diverses études de plusieurs disciplines scientifiques, nous avons défini cinq tâches différentes de manière à couvrir divers phénomènes à la fois sur le mouvement et les aspects linguistiques. Toutes les tâches sont de type monologue, car notre équipement ne permet pas d'enregistrer des dialogues. Ces tâches sont respectivement des tâches de description de photos (section 3.1.1.1), traduction (section

¹⁶ Locuteurs de LSF

3.1.1.2), explication (section 3.1.1.3), et narration à partir d'images (section 3.1.1.4.1) et de film (section 1.4.2).

3.1.1.1 Description de photos

Cette tâche a consisté à décrire le contenu de 25 photos. Les photos montrent de nombreux objets ayant des propriétés géométriques particulières (par exemple, des arrangements horizontaux ou verticaux, etc.) (figure3-1).

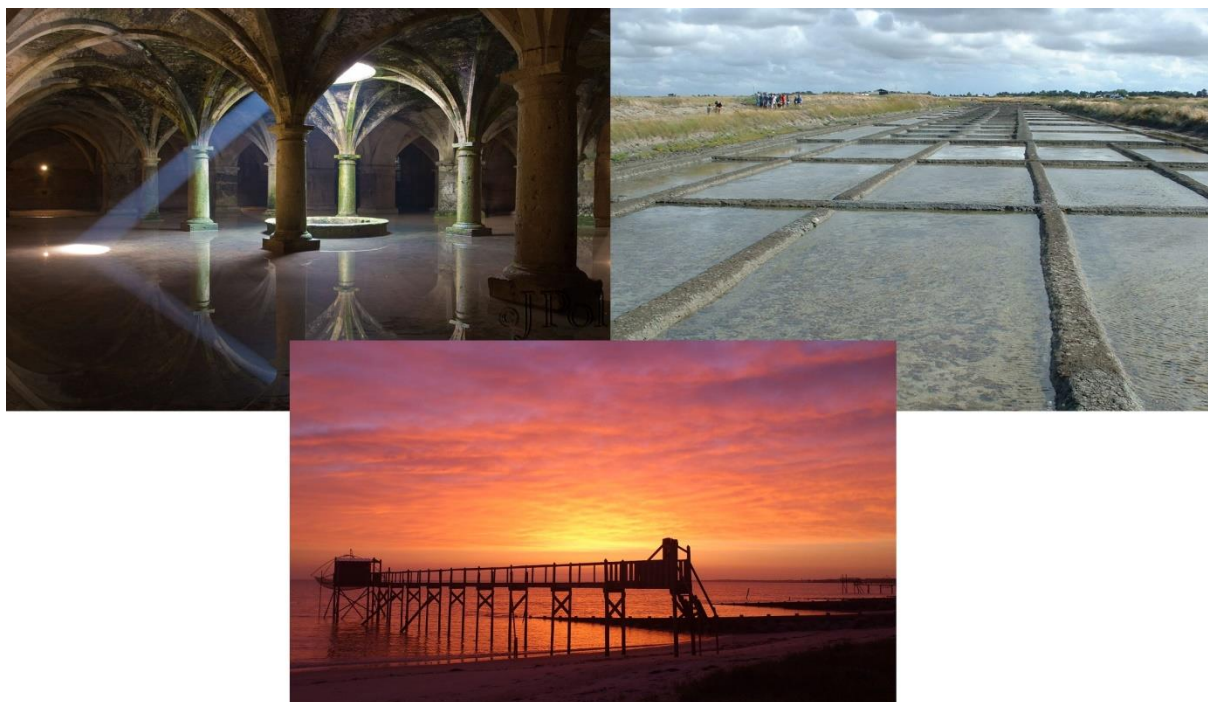


Figure 3-1: Exemples des photos pour la tâche 1.

Nous avons demandé aux signeurs de décrire les photos d'une manière spontanée. Chaque signeur regardait chaque photo pendant quelques instants puis la décrirait avant de passer à la suivante.

Toutes les photos utilisées dans cette tâche sont en annexe.

3.1.1.2 Traduction de brèves journalistiques AFP¹⁷

Au sein de notre équipe certaines études s'appuient sur l'étude de brèves journalistiques en LSF pour définir des règles linguistiques ayant pour objet la modélisation de la LSF (Filhol & Hadjadj, 2016; Filhol, Hadjadj, & Choisier, 2014). Le corpus utilisé jusqu'à maintenant était

¹⁷ Agence France-Presse, <https://www.afp.com/>

constitué de vidéos fournies par la société WebSourd¹⁸ qui était en charge de leur production au quotidien.

Cette tâche a consisté à fournir aux locuteurs 2 phrases en français écrit afin qu'ils en produisent la traduction en LSF. Cette tâche avait pour objectif d'enrichir le corpus WebSourd avec de la 3D afin d'analyser plus finement les structures spatiales en LSF. Les données de cette tâche ne seront pas analysées dans ces travaux de thèse mais font partie intégrante du corpus, tout comme les données issues des autres tâches décrites ci-dessous.

3.1.1.3 Explication

Dans cette tâche, le locuteur est invité à expliquer à un modérateur sourd placé derrière la caméra vidéo les procédures à l'aéroport depuis l'achat des billets d'avion jusqu'à l'arrivée à destination, comme si l'interlocuteur n'avait jamais voyagé par avion auparavant.

Les sujets regardent les consignes sous forme de vidéo en LSF et de diaporama avant de commencer l'enregistrement. Cette présentation contient des images présentées dans l'ordre chronologique montrant les différentes étapes : comme l'enregistrement, l'embarquement, la récupération des bagages, etc. (figure 3-2). Cette tâche est identique à l'une des tâches conçues dans le corpus Dicta-Sign (Matthes et al., 2010), afin, comme précédemment, d'enrichir ce corpus avec des données 3D.



Figure 3-2: Présentation de la tâche de l'aéroport

¹⁸ La société WebSourd a fermé en 2015.

3.1.1.4 Narration

Afin de pouvoir étudier des structures linguistiques au niveau du discours, deux tâches ont été conçues de manière à obtenir des productions de plus longue durée.

3.1.1.4.1 Narration basée sur des images

Cette tâche consiste à raconter deux événements bien connus en France, qui ont été souvent traités par les médias journalistiques. Ici, des durées plus longues de discours structurés sont attendues. Ces deux événements sont l'histoire du Titanic et l'affaire d'Omar Raddad. Le matériel d'élicitation de cette tâche était des images et des informations clés de ces événements, structurés sous forme de graphes afin d'éviter une éventuelle influence de la structure linéaire du français (figure 3-3). La documentation a été envoyée aux locuteurs une semaine avant l'enregistrement afin qu'ils puissent préparer la tâche.

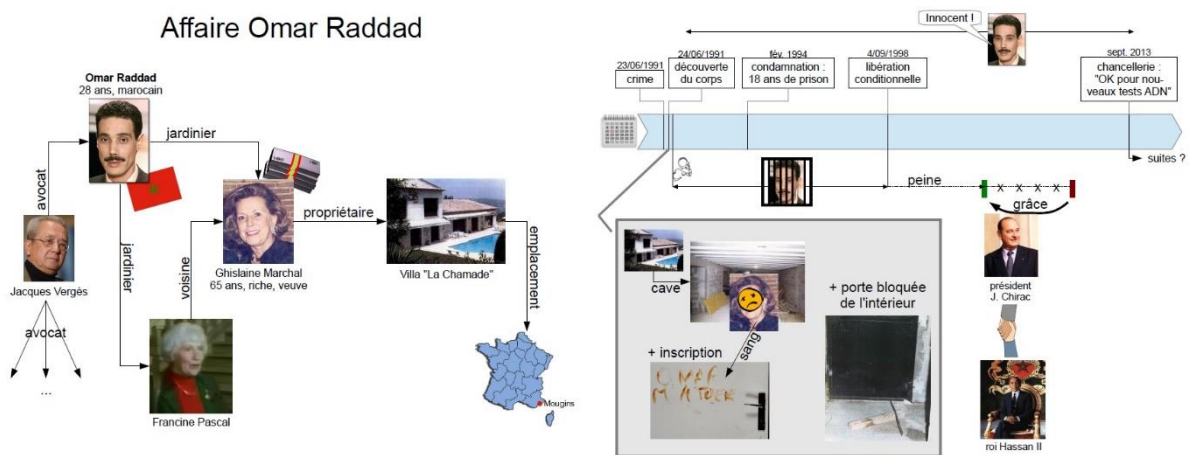


Figure 3-3 : Graphes relatant l'affaire d'Omar Raddad.

3.1.1.4.2 Narration basée sur un film

Cette tâche consiste à raconter le film « The Artist ». Le DVD a été envoyé aux locuteurs avant la session d'enregistrement afin qu'ils puissent le regarder. D'autres phénomènes linguistiques (par exemple, des prises de rôle) sont attendus de cette tâche.

Nous avons demandé aux sujets qu'ils fassent deux claques avec les mains avant de commencer la signation, et deux claques à la fin de la signation. Cette procédure nous permet de synchroniser les données 3D avec la vidéo dans la procédure d'annotation du corpus, et pour d'autres éventuelles analyses qui utiliseraient à la fois la vidéo et les données 3D.

La durée de ce corpus est d'environ 2 heures et 20 minutes, il contient 207 vidéos et 207 fichiers 3D.

3.1.2 Matériel utilisé

Dans cette section, tous les détails relatifs à la méthode de mocap utilisée sont donnés. D'une part, cela a pour objectif de permettre la diffusion et la réutilisation du savoir pour de futures études. D'autre part, le corpus constituant une ressource fondamentale et diffusable dans la communauté de LS, ces informations sont importantes et doivent accompagner les données de mocap afin qu'elles soient ré-utilisables par d'autres chercheurs.

Pour enregistrer le corpus, nous avons combiné un système de capture de mouvement optique Optitrack avec une caméra vidéo HD. Le système Optitrack S250e comporte 10 caméras, avec une résolution de 0.7 Mpixels et une fréquence maximale de 250 Hz. Le logiciel Motive¹⁹ fourni avec le système, propose plusieurs squelettes calibrés (disposition des marqueurs sur le corps) et des méthodes de post traitement des données comme l'interpolation des données perdues durant l'enregistrement, le lissage de la trajectoire des marqueurs, la suppression des bruits etc. Ces méthodes seront détaillées dans la section post traitement (section 7).

Pour que le corpus soit plus complet et respecte notre contrainte de pluridisciplinarité, une caméra vidéo HD a été utilisée en parallèle avec le système de mocap. Les vidéos ont été synchronisées avec les données 3D et ont été utilisées dans l'annotation du corpus (plus de détails seront donnés en partie II). En outre, les vidéos seront elles-mêmes utilisées dans plusieurs études en analyse linguistique et informatique.

Nous avons inclus dans nos tests l'utilisation des lunettes Tobii (Oculomètre). C'est un dispositif qui suit les mouvements des yeux et la direction du regard. Finalement, nous n'avons pas utilisé ce dispositif dans l'enregistrement du corpus, car il cache les marqueurs attachés aux sourcils et aux paupières dont l'enregistrement était plus important pour nous.

3.1.3 Les Marqueurs

Les marqueurs sont des objets sphériques fixés sur les objets dont on veut suivre les trajectoires. Dans cette partie, nous détaillons les types de marqueurs qui existent (section 3.1) et la disposition des marqueurs utilisés (section 3.2).

3.1.3.1 Type de marqueurs

Deux types de marqueurs existent selon le type du système (actif ou passif). Les marqueurs passifs sont revêtus d'un matériau réfléchissant la lumière qui est générée à proximité des

¹⁹ <http://optitrack.com/products/motive/>

objectifs des caméras. Contrairement aux marqueurs passifs, les marqueurs actifs sont eux-mêmes alimentés pour émettre leur propre lumière à destination des caméras.

Notre système Optitrack est un système passif et nous avons donc utilisé des marqueurs passifs. Les marqueurs sont habituellement fixés directement sur la peau, ou ils sont attachés à un costume conçu spécifiquement pour la capture de mouvement. Ce type de système peut capturer un grand nombre de marqueurs à des cadences généralement autour de 120 à 250 fps. Le seuil des caméras peut être réglé de façon à ce que la brillance des marqueurs réfléchissants soit captée et non pas la peau et le tissu.

Deux tailles de marqueurs ont été utilisées : des marqueurs de taille moyenne et des petits marqueurs. Les marqueurs de taille moyenne ont une taille de 12.7 mm. Ils ont été utilisés sur le buste, la tête, les bras et les mains. Les petits marqueurs ont une taille de 3 mm. Ils ont été utilisés pour le visage. Chaque sujet a porté une combinaison spéciale pour la mocap. La combinaison est élastique et complétée par un chapeau. Les marqueurs du buste, des bras, et des mains ont été placés sur la combinaison dans les endroits définis dans la section 3.2, les marqueurs de la tête ont été placés sur le chapeau. Les petits marqueurs de visage ont été placés directement sur la peau (figure 3-4).



Figure 3-4: Disposition des marqueurs

3.1.3.2 Disposition des marqueurs

Une étape importante de la création du corpus est de concevoir la meilleure configuration de l'emplacement des marqueurs. Pour cela, nous avons mené des évaluations de telle sorte que nous puissions enregistrer suffisamment de détails sur la performance humaine pour nos

différents besoins : le rejeu de l'animation avec signeur virtuel, l'analyse 3D de données, et l'annotation linguistique.

Une des questions les plus importantes dans l'enregistrement des données mocap est l'emplacement des marqueurs : où fixer les marqueurs sur le corps, et pourquoi ? Cette question est importante parce que l'emplacement des marqueurs affecte leur visibilité dans le système : les marqueurs couverts ne sont pas enregistrés. Chaque marqueur doit être vu par deux caméras au minimum. L'emplacement des marqueurs est également important du point de vue des étapes de post-traitement tels que la transformation des données de marqueurs tridimensionnelles en représentation en segments. En outre, le marqueur qui est placé de façon inappropriée pourrait rendre difficile pour le locuteur d'articuler correctement les signes. Enfin, l'emplacement du marqueur doit nous permettre de suivre, autant que possible tous les mouvements utiles d'un point de vue linguistique.

Au final, nous avons utilisé 40 marqueurs, la figure 3-5 montre la configuration des marqueurs que nous avons utilisés. Il y a 4 marqueurs sur le haut de la tête : 2 sur le front et 2 à l'arrière (os occipital). Le torse comporte 7 marqueurs : 4 sur la partie supérieure (sternum, clavicule, deux à l'arrière (C7 : apophyse épineuse de la 7e vertèbre cervicale, et T10 : processus épineux de la 10e vertèbre thoracique)), les 3 autres marqueurs sur le bassin. Chaque bras dispose de 5 marqueurs placés sur les principales positions communes (l'épaule, le coude intérieur et extérieur, l'ulnaire du poignet et le radial) et un autre marqueur sur les triceps. Il y a également 2 marqueurs sur chaque main. Un ensemble de 13 marqueurs est utilisé sur le visage : les sourcils, les paupières, les joues, le menton et 4 autour de la bouche (au-dessous, au-dessus, à gauche et à droite).

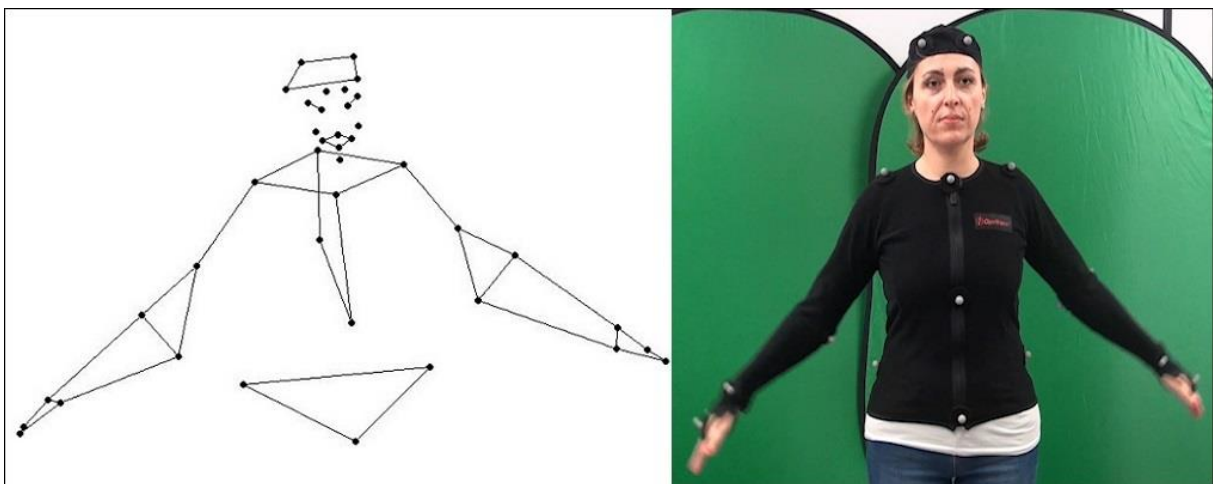


Figure 3-5: Disposition des marqueurs pendant l'enregistrement

Les positions des marqueurs dans notre configuration ont été choisies de manière à ce qu'ils soient au maximum visibles et identifiables par le système, et pour qu'ils captent les principaux mouvements des mains, des bras, du torse, de la tête et du visage. L'emplacement des 6 marqueurs de chaque bras a été choisi de manière à être en mesure de reconstituer l'orientation (angles conjoints) des 2 segments du bras (bras supérieur et l'avant-bras). La règle est qu'il doit y avoir au moins 3 marqueurs sur un corps rigide pour définir son orientation 3D. Les 2 marqueurs sur chaque main permettent d'avoir les mouvements et les rotations des mains. Les 4 marqueurs sur la partie supérieure du torse plus les 3 marqueurs sur le bassin permettant d'avoir la rotation de la partie supérieure du torse par rapport à la partie inférieure. De plus, ces 7 marqueurs permettent de capter les mouvements du torse. Les marqueurs sur le visage permettent de capter les mouvements des sourcils, la fermeture des yeux, la configuration des joues et les mouvements de la bouche.

Plusieurs tests avec des marqueurs sur les doigts en plus de la configuration précédente ont été réalisés, mais avec ce système de 10 caméras, il était impossible de capter ces marqueurs, car il y avait toujours un chevauchement de certains d'entre eux. De plus, ils étaient perdus pendant des temps trop longs ce qui ne permettait pas de les récupérer. Pour pouvoir capter les doigts, il faut un grand nombre de caméras concentrées sur les mains, car il y a beaucoup de contacts entre les doigts. Zhao et al se sont focalisés sur les mains 12 caméras mocap et une caméra kinect pour avoir des données de mouvements des mains de qualité (Zhao, Chai, & Xu, 2012). Pour cette raison nous avons décidé de ne mettre que 2 marqueurs sur le dos de chaque main pour capter au moins leurs mouvement et orientation.

En comparaison avec d'autres études mocap récentes, le nombre total de marqueurs dans notre configuration est assez complet : Jantunen et ses collaborateurs ont utilisé 20 marqueurs (7 sur chaque bras et main, 4 sur la tête, et 2 sur la partie supérieure du torse), une version avec une disposition plus grande a été testée où ils ont ajouté 5 marqueurs sur le visage et 6 sur la partie inférieure du torse (Jantunen et al., 2012). Tyrone et ses collaborateurs ont utilisé 30 marqueurs (7 sur chaque bras, 7 sur la tête, et 9 sur le torse) (Tyrone et al., 2010). Duarte et Gibet dont l'objectif est d'utiliser les données pour animer un signeur virtuel, ont utilisé 98 marqueurs (43 marqueurs sur visage, 43 marqueurs sur le corps, et 6 sur chaque main) (Duarte & Gibet, 2010).

3.1.4 Disposition des caméras

Le positionnement des caméras est une tâche importante dans la création du corpus. Des dispositions standards de 8 à 12 caméras sont donnés par les constructeurs des systèmes de

mocap pour une capture du mouvement du corps. La contrainte d'utiliser des marqueurs sur le corps et le visage avec seulement 10 caméras nous a obligés à élaborer une nouvelle disposition, qui permet que chaque marqueur soit capté par au moins deux caméras.

La figure 3-6 montre notre configuration optimale des caméras. Nous avons utilisé 4 caméras face au signeur et à la même hauteur que sa tête, 2 caméras de chaque côté, et 2 caméras derrière le signeur. Les caméras de face permettent une très bonne captation des marqueurs sur le visage et sur l'avant du torse. Les caméras sur les côtés et de face permettent de capturer les marqueurs posés sur les bras et sur le front, les 2 caméras derrière sont suffisantes pour capter les 2 marqueurs placés sur le dos (C7 et T10) et sur l'arrière de la tête (os occipital). La caméra vidéo numérique est placée en face. Cette configuration nous a permis une très bonne captation avec un minimum de lacunes et de chevauchement dans les données pendant l'enregistrement.



Figure 3-6: Disposition de caméras.

3.1.5 Calibration

Comme beaucoup d'autres systèmes de mesures, la calibration est indispensable pour les systèmes de capture de mouvement optique. Lors du calibrage des caméras, le système calcule la position et l'orientation de chaque caméra et les distorsions des images captées. En utilisant les données de calibration, le logiciel Motive construit le volume de capture en 3D. Plus

précisément, cela se fait en observant les images 2D à partir de plusieurs caméras synchronisées et en associant la position des marqueurs de calibration de chaque caméra par triangulation.

Un changement dans la configuration ou l'emplacement d'une caméra au cours ou après la calibration endommage la calibration, ce qui nécessite une re-calibration du système. En outre, même si les réglages ne sont pas modifiés, la précision de la calibration va naturellement se détériorer au fil du temps en raison de facteurs ambiants, tels que les changements de la température et d'autres conditions environnementales. Ainsi, pour des résultats précis, il est recommandé de calibrer le système avant chaque utilisation.

La calibration du système nécessite plusieurs étapes. Cette section détaille toutes ces étapes : réglage des caméras (section 3.1.5.1), masquage des objets réfléchissant (section 3.1.5.2), calibration spatiale (section 3.1.5.3), réglage du repère et de l'origine (section 3.1.5.4).

3.1.5.1 Réglages des caméras

La première étape est le réglage des caméras. Quatre paramètres sont à régler en fonction de la nature d'utilisation du système, du volume de l'espace utilisé, et de l'environnement. Ces paramètres sont la fréquence, l'exposition, le seuil de luminosité, et l'éclairage des LED (diode électroluminescente).

La fréquence (FPS) est le paramètre le plus important, les valeurs minimales et maximales de la fréquence dépendent du modèle des caméras utilisées. Dans notre cas la valeur minimale est 30Hz et la maximale est 250Hz. Une grande fréquence permet de capter plus d'images par seconde. Une petite fréquence permet d'avoir des valeurs d'exposition plus élevées (et des images plus lumineuses) et permet également de réduire la taille des données.

L'exposition (EXP) définit la durée d'exposition par image, la valeur minimale est 10 et la maximale est 3800. Une plus forte exposition permet d'avoir plus de lumière, créant une image plus lumineuse qui peut augmenter la visibilité pour les marqueurs petits et sombres. Toutefois, une exposition trop élevée peut introduire de faux marqueurs, et plus de bruit, cela donne un impact négatif sur la qualité des données du marqueur.

Le seuil de luminosité (THR) définit la luminosité minimale pour qu'un pixel soit vu par une caméra, tous les pixels en dessous du seuil sont ignorés, la valeur maximale est 255. L'augmentation du seuil peut aider à filtrer les interférences avec les éléments perturbateurs (par exemple, des réflexions et des sources lumineuses externes), un seuil bas peut permettre aux petits marqueurs qui sont loin des caméras d'être vus par le système.

L'éclairage (LED) définit le niveau de luminosité des LED des caméras infrarouges (IR), la valeur maximale est 15. Une valeur élevée permet à la caméra d'émettre plus d'IR ce qui peut permettre une meilleure détection de marqueurs à une plus longue portée. Cependant, une valeur trop élevée peut provoquer la réflexion sur des objets non marqueurs.

Les valeurs des paramètres que nous avons utilisés sont : fréquence à 250Hz, exposition à 510, seuil de luminosité à 227, et éclairage des LED à 12. Ces valeurs n'ont pas été choisies arbitrairement, le choix a été fait par rapport à des critères. Nous avons utilisé la plus grande fréquence disponible pour avoir plus de données dans les petits mouvements, ce qui nous a permis d'analyser plus de mouvements. La valeur de l'exposition utilisée est relativement petite, car notre volume de capture est relativement petit et la distance entre les caméras et les sujets (marqueurs) est petite. La valeur du seuil de luminosité utilisée est relativement grande, pour les mêmes raisons. Pour les quatre caméras de face, nous avons augmenté le seuil à 243, car pendant que le locuteur signe, la distance entre ces caméras et les marqueurs des bras est très petite. La valeur d'éclairage des LED utilisée est relativement élevée, car nous avons utilisé un seuil de luminosité élevé.

Ces paramètres sont influencés par l'environnement de la capture (l'emplacement du sujet, la disposition des caméras, la disposition des marqueurs, et le volume de capture). Tous ces paramètres sont importants, car un seul paramètre mal réglé a une influence sur la qualité des données enregistrées. Des tests ont été faits avec plusieurs valeurs afin de sélectionner ces valeurs optimales pour notre corpus.

3.1.5.2 Masquage des objets réfléchissants

Tout objet réfléchissant et les marqueurs inutiles doivent être retirés du volume de capture avant de calibrer le système. En fait, le processus de calibration ne démarre pas s'il y a beaucoup d'objets réfléchissants présents dans les vues des caméras autres que la baguette de calibrage. Cependant, dans certaines situations, les reflets indésirables ou les objets qui appartiennent au système ne pouvaient pas être retirés de la configuration tels que les caméras (dans certaines dispositions de caméras, une ou plusieurs caméras peuvent entrer dans le champ de vision des autres caméras). Dans ce cas, ces objets peuvent être ignorés par l'intermédiaire de l'outil de masquage.

Le masquage applique des masques rouges sur les objets réfléchissants, et tous les pixels dans les régions masquées sont entièrement filtrés. L'utilisation de l'outil de masquage pour éliminer toutes les réflexions indésirables est nécessaire avant de procéder à l'étape suivante.

Un outil permet de détecter et masquer tous les objets réfléchissants qui entrent dans la vue des caméras. Il permet de le faire automatiquement et aussi de masquer manuellement les objets réfléchissants indésirables avec des masques circulaires, rectangulaire ou de forme libre.

Les données dans les régions masquées ne sont pas collectées, donc l'utilisation excessive de masquage peut entraîner une perte de données. Par conséquent, tous les objets réfléchissants amovibles doivent être retirés ou recouverts avant l'utilisation de l'outil de masquage. La disposition de caméras doit être choisi de manière à ce que les caméras ne se voient pas entre elles autant que possible.

3.1.5.3 Calibration spatiale

Le processus de calibration spatiale est l'étape de base dans la calibration. Une baguette de calibration est agitée devant les caméras dans l'espace d'enregistrement à plusieurs reprises, ce qui permet à toutes les caméras de voir les marqueurs. Grâce à ce processus, chaque caméra capture des échantillons pour calculer leur position respective et leur orientation dans l'espace 3D. Il y a des baguettes de calibration adaptées aux différentes applications de capture.

Après avoir sélectionné le type de baguette utilisée, on lance le processus de calibration en dessinant le chiffre huit avec la baguette pour collecter des échantillons à des orientations différentes, et en couvrant autant que possible cet espace pour avoir des échantillons suffisants. Pour un meilleur résultat, il faut agiter la baguette de manière uniforme et complète dans tout le volume.

Pendant la calibration spatiale, il est recommandé de vérifier la vue 2D des caméras dans le volet d'aperçu pour évaluer la couverture de chaque caméra. S'il y a des lacunes importantes, il faut concentrer la calibration spatiale sur ces endroits pour augmenter la couverture. Lorsque des quantités suffisantes d'échantillons de calibration sont collectées par chaque caméra. Motive calcule la calibration pour le volume de capture. Généralement, 2000 à 10000 échantillons sont suffisants selon la grandeur du volume à calibrer. Dans notre cas, nous avons collecté entre 3000 et 4000 échantillons, car notre volume était relativement petit.

Bien qu'il faille collecter des échantillons dans tout le volume, il est très utile de collecter plus d'échantillons autour de la région où les marqueurs se déplacent durant l'enregistrement. En procédant ainsi, les résultats de calibration ont une meilleure précision dans cette région.

Le temps nécessaire pour le calcul de calibration varie en fonction du nombre de caméras incluses dans l'installation, ainsi que le nombre d'échantillons collectés. Dans notre cas, c'est

de l'ordre 5 minutes. La note générale du résultat est la note la plus basse des caméras. La qualité globale montre, la qualité du solveur actuel (figure 3-7).

Cam	Samp	Quality	Focal	MeanEr	PixelEr
Ave		Exceptional		0.27	0.067
Wand		Exceptional		0.11	0.114
1	1576	Exceptional	6.55	0.29	0.071
2	1956	Exceptional	6.57	0.27	0.069
3	2000	Exceptional	6.57	0.26	0.072
4	1576	Exceptional	6.56	0.28	0.064
5	1576	Exceptional	6.55	0.26	0.057
6	1820	Exceptional	6.58	0.26	0.070
7	1900	Exceptional	6.57	0.27	0.067
8	1580	Exceptional	6.57	0.28	0.062

Figure 3-7: Volet de calibration pendant le calcul.

A la fin du processus de calibration, un rapport fournit une évaluation de la qualité de la calibration (figure 3-8).

Overall Reprojection	Mean 3D Error: 0.486 mm	Mean 2D Error: 0.095 pixels (Exceptional)
Worst Camera	Mean 3D Error: 0.553 mm	Mean 2D Error: 0.126 pixels (Exceptional)
Triangulation	Recommended: 2.8 mm	Residual Mean Error: 0.5 mm
Overall Wand Error	Mean Error: 0.102 mm (Exceptional)	
Ray length	Suggested Max: 16.7 m	

Figure 3-8: Rapport de calibration

La calibration est enregistrée dans un fichier autonome qui contient toutes les informations requises pour restaurer un volume de la calibration, y compris les caméras utilisées et leurs

paramètres de distorsion. Ce fichier est utile lorsque la configuration du système a été maintenue constante, et peut-être chargée automatiquement au démarrage pour fournir une calibration prête.

3.1.5.4 Réglage du repère et de l'origine

L'étape finale du processus de calibration est la fixation du repère (plan de masse) et de l'origine. Cela est fait en plaçant un triangle de calibration dans l'endroit où l'on souhaite que l'origine soit localisée. La partie la plus longue du triangle de calibration indique l'axe Z, et la partie courte indique l'axe X, donc automatiquement l'axe Y est dirigé perpendiculairement vers le haut.

3.1.6 Enregistrement des données (corpus)

La capture (enregistrement) des données peut commencer lorsque le volume de capture a été calibré et tous les marqueurs ont été placés sur le sujet. Une fois que la capture a commencé toutes les caméras sont actives. En d'autres termes, Motive reconstruit en continu les réflexions détectées dans le volume, et les données de capture sont enregistrées en temps réel.

Les fichiers d'enregistrement des données contiennent toutes les informations nécessaires pour recréer l'intégralité de la capture à partir du fichier, y compris la calibration des caméras, les données des caméras 2D, les données 3D reconstruites, les modifications. Dans le cas où on utiliserait les dispositions de marqueurs de Motive (des squelettes avec des dispositions de marqueurs sont proposés par Motive), les données d'angle et les modèles de suivi (squelettes) sont aussi enregistrés dans ce fichier.

Ces fichiers peuvent être utilisés ultérieurement pour éditer les données, faire un post-traitement, et exporter les données.

3.1.7 Post-traitement

L'utilisation d'un grand nombre de marqueurs (40 pour ce corpus) présente un inconvénient, qui est la grande quantité de lacunes dans les données, ainsi que le chevauchement entre les marqueurs qui sont proches au repos ou en mouvement. Cet inconvénient peut être réduit par l'utilisation d'un plus grand nombre de caméras mocap (ici nous avons utilisé 10 caméras). Au moins 2 caméras doivent voir un marqueur à chaque trame pour pouvoir récupérer instantanément son emplacement en trois dimensions. Si nous avons un système avec plus de caméras (18 ou 20) nous n'aurions pas (ou beaucoup moins) de lacunes dans les données et de

chevauchement entre les marqueurs. Pour régler ce problème, il faut faire un post-traitement sur les données.

Une fois que les données brutes sont enregistrées, il y a plusieurs étapes essentielles qui doivent être effectuées avant que les données puissent être exportées et exploitées. En raison de l'occlusion possible entre les diverses parties du corps, et parce que les marqueurs ne sont pas identifiés et peuvent apparaître identiques, un post-traitement est nécessaire pour nettoyer et exporter les données. Ces étapes sont la labélisation (identification) des marqueurs (section 3.1.7.1), la suppression du chevauchement possible entre les différents marqueurs (section 3.1.7.2), l'interpolation des données perdues (lacunes) durant l'enregistrement (extrapolation ou Gap-filling en anglais) (section 3.1.7.3), la suppression des bruits (marqueurs fantômes) (section 3.1.7.4), et finalement le filtrage des données (section 3.1.7.5).

3.1.7.1 Labélisation

La labélisation (étiquetage) est le processus d'identifier et de donner des noms aux marqueurs utilisés. Le logiciel Motive donne plusieurs dispositions de marqueurs qui peuvent être utilisés durant l'enregistrement. L'utilisation de ces dispositions permet d'obtenir directement les marqueurs étiquetés à la fin de l'enregistrement. Malheureusement, ces dispositions ne prennent pas en considération le visage. Comme notre but est de créer un corpus qui inclus les mouvements du visage, donc, nous n'avons pas utilisé ces dispositions.

La première étape consiste à créer des ensembles de marqueurs. Nous avons choisi de faire des ensembles qui contiennent de petites listes de marqueurs pour simplifier la procédure d'étiquetage. Nous avons créé 5 ensembles, chacun représente une partie du corps : la tête contient 4 marqueurs, le visage contient 13 marqueurs, le buste contient 7 marqueurs, le bras droit avec la main droite, forme un ensemble de 8 marqueurs, et la même chose pour le bras et la main gauche.

La deuxième étape consiste à donner à chaque marqueur son nom (identifiant). Chaque fois que le marqueur est perdu durant l'enregistrement et que le système ne le reconnaît pas lorsqu'il réapparaît une nouvelle fois, le marqueur doit être ré-identifié. À la fin de cette procédure tous les marqueurs sont étiquetés (identifiés).

3.1.7.2 Suppression du chevauchement

Cette étape consiste à vérifier et à supprimer les chevauchements qui existent entre les marqueurs. Les chevauchements apparaissent lorsque deux marqueurs sont trop proches au

cours de l'enregistrement, le système confond les deux marqueurs et inverse leur identité. Il y a deux cas dans le chevauchement :

- Dans le premier cas, l'identité des marqueurs reste inversée après le chevauchement.
- Dans le second cas, le système inverse l'identité des marqueurs uniquement pendant le chevauchement, et lorsque les marqueurs s'éloignent les uns des autres, leurs identités deviennent correctes.

La vérification est faite marqueur par marqueur tout au long de la séquence d'enregistrement. Quand il y a un chevauchement entre deux marqueurs, on supprime les données des deux marqueurs pendant le chevauchement. Puis, dans le premier cas où l'identité des marqueurs reste inversée, nous enlevons l'étiquetage qui est après le chevauchement de ces deux marqueurs en sélectionnant la partie où l'étiquetage est inversé, et nous les ré-étiquetons avec les bonnes identifications. Dans le second cas, les marqueurs ont la bonne identification, donc il n'y a rien à faire de plus.

3.1.7.3 Extrapolation

Cette étape consiste à interpoler les données perdues (lacunes) durant l'enregistrement, et qui ont été supprimées durant la suppression des chevauchements. Dans le contexte de la LSF, il y a beaucoup de contacts entre les mains, les mains avec les bras, les mains avec le visage, et même les mains avec le buste. Beaucoup de marqueurs peuvent être cachés pendant les contacts, ce qui entraîne des lacunes dans les données qui doivent alors être interpolées.

L'interpolation peut être faite dans le logiciel Matlab avec des Toolboxes de mocap tels que celle proposée par Burger et Toiviainen, après l'exportation des données (Burger & Toiviainen, 2013). Mais ces Toolboxes donnent des résultats tout à fait arbitraires lorsque les lacunes sont trop longues. Nous avons préféré utiliser une option du logiciel. Cette option propose plusieurs algorithmes d'interpolation pour combler les lacunes (interpolation constante, interpolation linéaire, interpolation cubique, ou interpolation fondée sur des modèles (marqueurs)).

L'interpolation fondée sur des modèles utilise le fait que plusieurs marqueurs sont posés sur le même segment, et sont fixes les uns par rapport aux autres, c'est la méthode que nous avons utilisée, car elle est la meilleure méthode surtout pour les grandes lacunes. Elle interpole les données en suivant les trajectoires des autres marqueurs choisis. Dans le cas où tous les marqueurs du même segment étaient perdus au même moment, nous avons utilisé une des autres interpolations disponibles pour interpoler les données du marqueur qui a la plus petite lacune.

Puis les autres marqueurs ont été interpolés par l'interpolation par rapport au premier marqueur interpolé. L'interpolation cubique a été utilisée pour les lacunes dont le mouvement du marqueur est non linéaire. L'interpolation linéaire a été utilisée pour les petites lacunes et lorsque le mouvement du marqueur est linéaire. L'interpolation constante a été utilisée pour les lacunes qui sont dans des périodes où le marqueur ne bouge pas. La figure 3-9 montre un exemple de lacune interpolée avec les différentes méthodes d'interpolation.

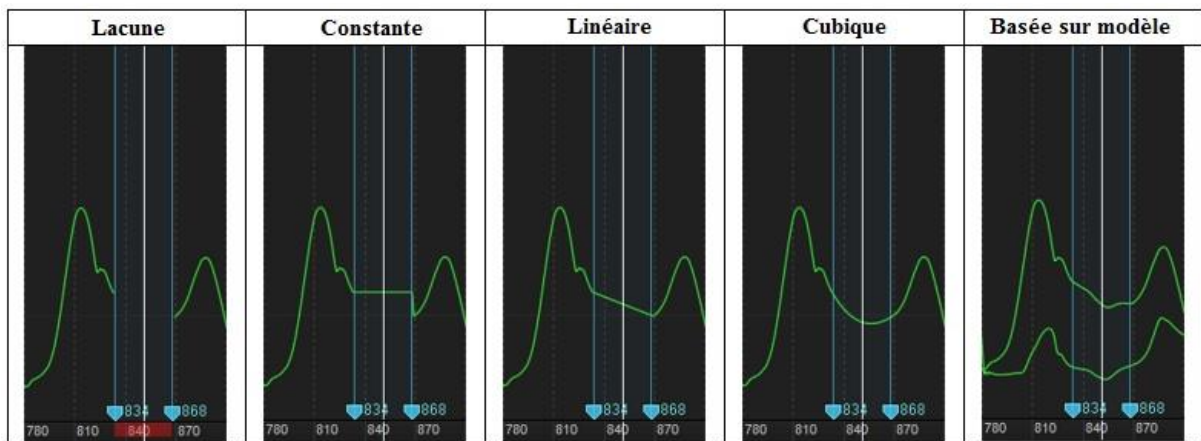


Figure 3-9: Exemple d'interpolation avec les différentes méthodes.

3.1.7.4 Suppression des faux marqueurs

Cette tâche consiste à supprimer les faux marqueurs, encore appelés phantoms. Ces phantoms apparaissent dans le logiciel Motive comme des vrais marqueurs, mais la différence est qu'ils sont d'une durée très courte.

Pour supprimer ces bruits, il suffit de diviser la liste des marqueurs en deux : les marqueurs identifiés et les marqueurs non identifiés. Les marqueurs qui apparaissent dans la liste des marqueurs non définis sont tous des faux marqueurs. Il suffit de sélectionner tous ces marqueurs et les supprimer. Cette méthode nous a permis de supprimer tous les faux marqueurs facilement et très rapidement.

3.1.7.5 Le filtrage des données

Il est connu que lorsque le mouvement est plus lent, il y a plus du bruit (vibration) dans les données, et comme dans un discours en LS la vitesse des mouvements est variable, cela a produit des vibrations dans les données (trajectoire des marqueurs). Ces vibrations peuvent être supprimées, après l'exportation, avec les logiciels de traitement de données comme Matlab. Mais pour gagner du temps, les vibrations peuvent être supprimées avant d'exporter les

données. Le logiciel Motive propose une fonction de filtrage pour supprimer ces vibrations et lisser les trajectoires des marqueurs.

La fonction de filtrage (lissage) applique un filtre passe-bas (Butterworth) sur les données, cela modifie la trajectoire du marqueur. Tous les mouvements vibrants sont filtrés avec cette fonction. Tout d'abord, nous avons choisi la fréquence de coupure du filtre pour définir à quel point nos données seront lissées. Ensuite, nous avons choisi le marqueur et la partie de la trajectoire à lisser. Quand la fréquence de coupure est haute, seuls des signaux de hautes fréquences sont filtrés. Quand la fréquence de coupure est basse, les signaux de basse fréquence sont aussi filtrés. Autrement dit, l'utilisation d'une fréquence de coupure basse lissera la plupart de la trajectoire, tandis que l'utilisation d'une haute fréquence de coupure lissera seulement les vibrations (figure 3-10).



Figure 3-10 : Exemple de filtrage avec une haute fréquence de coupure (10Hz) et une basse fréquence de coupure (5Hz).

Généralement la fréquence de coupure de filtre utilisée est entre 7 Hz à 12 Hz, mais lorsque les mouvements sont rapides, on peut augmenter la valeur pour éviter d'adoucir les transitions de mouvement. Dans notre cas, la fréquence de coupure utilisée est 10 Hz, elle est suffisante pour supprimer les vibrations sans modifier les données. De plus le filtrage n'a été utilisé que dans les parties où il y a des vibrations dans les données.

Les données de la première tâche (description des images), qui ont été utilisées dans les analyses des lois classiques du contrôle moteur, n'ont pas été filtrées dans le logiciel Motive, mais nous

les avons filtrées après l'export dans le logiciel Matlab. Nous avons appliqué un filtre passe-bas (Butterworth) avec une fréquence de coupure de 10Hz sur toutes les données.

Une fois le processus de filtrage terminé, les données peuvent être exportées et utilisées.

3.1.8 Export des données

Motive exportent les données 3D reconstruites dans des fichiers de différents formats et les fichiers exportés peuvent être importés dans d'autres logiciels pour utiliser les données de capture. Les formats supportés comprennent CSV, C3D, FBX, BVH et TRC.

Les options d'export sont : la fréquence, l'échelle de mesure, et la plage des trames de données à exporter. Ces options sont disponibles pour chaque format d'export. En fonction des options activées, l'export des données peut inclure les données des marqueurs reconstruites en 3D, ou les données du squelette. Chaque format peut exporter un ou plusieurs types de données. Le tableau 3-1 montre quels sont les types de données disponibles dans les différents formats d'exportation :

Type de données	C3D	CSV	BVH	FBX	TRC
Les données des marqueurs reconstruites en 3D	◆	◆		◆	◆
Les 6 degrés de liberté de données du corps rigide		◆	◆	◆	
Les données du squelette		◆	◆	◆	

Tableau 3-1: Les formats et le type de donnée qui peuvent être exportés.

C3D est un format de fichier binaire qui est largement utilisé en particulier dans les applications de biomécanique et d'étude de mouvement, il contient les données 3D des marqueurs. CSV est un format de fichier texte qui peut être importé par un logiciel de tableur ou un script. Selon les options d'export de données activées, les fichiers CSV exportés peuvent contenir des données de marqueurs, des données de corps rigides, et des données squelette. TRC est un format de fichier qui contient les données brutes de mocap, qui comprennent des données 3D des marqueurs étiquetés et non étiquetés. BVH est un format de fichier qui ne comprend pas les données des marqueurs. Au lieu de cela, un squelette est exporté en utilisant les relations de segments hiérarchiques. Un seul squelette est exporté par chaque fichier BVH. Le fichier BVH contient la définition du squelette qui est nécessaire pour caractériser le squelette dans d'autres

logiciels. Il existe deux types de fichiers FBX : Binaire et ASCII. Le format ASCII contient les données 3D des marqueurs ainsi que les 6 degrés de liberté de données selon les configurations de réglage d'exportation. Les fichiers ASCII peuvent également être ouverts et modifiés à l'aide des éditeurs de texte. Les fichiers Binaires sont plus compacts que les fichiers ASCII. Les fichiers Binaires ne comprennent pas les données des marqueurs, mais ils contiennent les squelettes, ainsi que les positions et orientations à l'origine du corps rigide.

La version de Motive que nous avons utilisée est une version limitée, elle ne permet d'exporter que les formats C3D et CSV. Les données ont été exportées dans ces deux formats avec deux fréquences 250 Hz et 25 Hz. La version en 25 Hz a été exportée pour être utilisée dans un logiciel d'annotation. La version 250 Hz est la version originale, elle a été utilisée dans les analyses.

3.2 Annotation du Corpus

Une fois les productions de LSF enregistrées, et avant de pouvoir analyser et étudier le mouvement, il faut repérer dans le flux de données quelles sont celles que l'on peut utiliser. Pour cela, on utilise un logiciel qui permet de visualiser les données et de les annoter.

Dans cette partie, nous présentons le logiciel d'annotation utilisé (section 3.2.1), l'utilisation des données de mocap (section 3.2.2), la structure de l'annotation (section 3.2.3), le processus d'annotation en lui-même (section 3.2.4), puis l'export de l'annotation (section 3.2.5).

3.2.1 Le logiciel ANVIL

Parmi les différents logiciels d'annotation, nous avons retenu le logiciel ANVIL, qui permet de visualiser et synchroniser les données de mocap et la vidéo à partir de laquelle les annotations sont effectuées, ce qui permet par la suite d'extraire des flux de données de mocap les segments qui nous intéressent pour les étudier.

ANVIL est un logiciel d'annotation de vidéo, développé par Michael Kipp. Il offre une annotation multi pistes à partir d'un schéma d'annotation défini par l'utilisateur (Kipp, 2001, 2008, 2012, 2014).

ANVIL permet d'importer les vidéos de formats AVI et MOV, les données de mocap sous forme de courbes issues de fichiers textes ou de représentations 3D issus de fichiers BVH (voir figure 3-11). Comme nos données 3D sont en format C3D, nous devons passer par une étape de création des fichiers pour pouvoir les importer dans ANVIL.

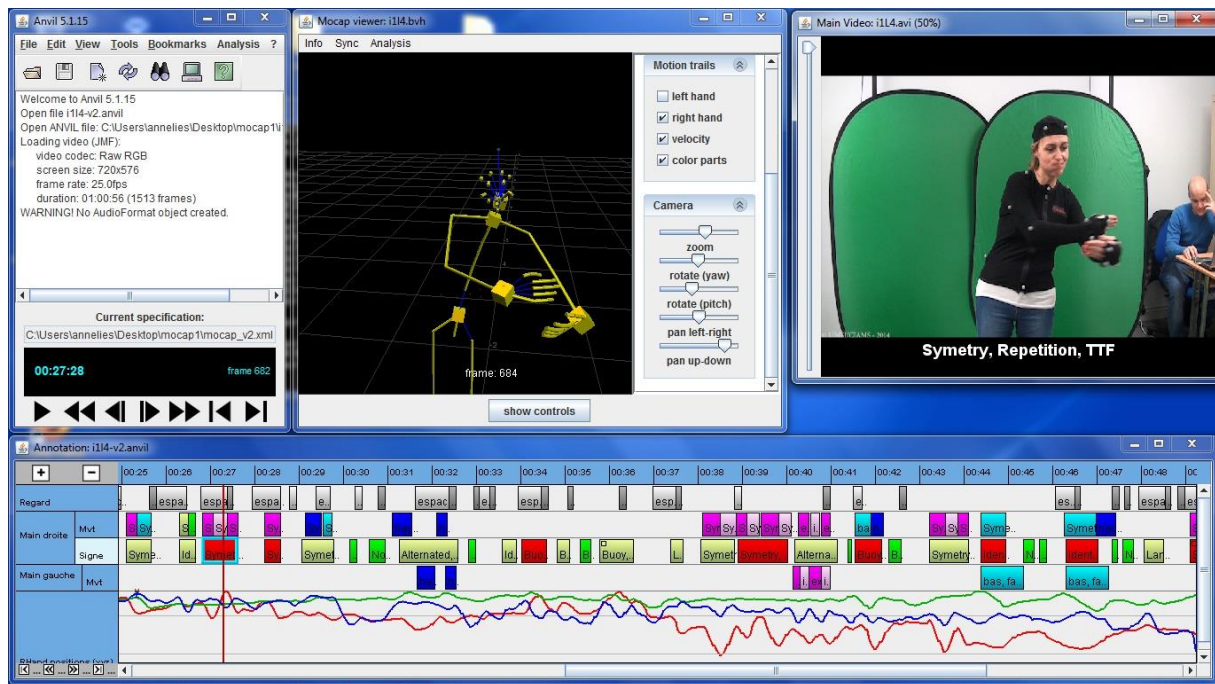


Figure 3-11: Capture d'écran du logiciel ANVIL montre l'annotation en utilisant la vidéo, les données de mocap en tant que squelette 3D et trajectoire des mains.

3.2.2 Création des fichiers BVH

Les fichiers BVH, comme nous l'avons expliqué dans la section d'avant, contiennent la spécification du squelette 3D. Pour créer le squelette à partir des données 3D des marqueurs, nous avons utilisé le logiciel 3DS-MAX²⁰ qui permet d'importer et d'éditer les fichiers C3D.

Pour chaque sujet, un squelette est créé, car chaque sujet a sa propre anthropométrie (les mesures de chaque membre du corps humain). La création du squelette est basée sur l'emplacement des marqueurs. Ces derniers sont les points-clés pour avoir les distances des membres supérieurs du corps, car ils sont attachés sur les jonctions des membres. Par exemple, la distance entre le centre des deux marqueurs attachés sur le poignet et le centre des deux marqueurs attachés sur le coude est la mesure de l'avant-bras. De cette façon les squelettes des 8 sujets ont été créés. Le visage n'a pas été modélisé, et les marqueurs de visage sont représentés par des points reliés au segment de la tête.

Une fois terminé, il suffit d'exporter l'animation dans des fichiers BVH.

²⁰ <http://www.autodesk.fr/products/3ds-max/overview>

3.2.3 Structure de l'annotation

La structure de l'annotation est basée sur un schéma d'annotation défini par les utilisateurs. Pour notre étude, nous avons réalisé un schéma permettant d'annoter les mouvements que nous souhaitons analyser et des informations de nature linguistique afin d'étudier l'éventuel « effet langue des signes » sur les lois du contrôle moteurs. Nous y avons adjoint des annotations sur le regard pour d'autres études menées dans l'équipe sur ce corpus.

Le schéma d'annotation contient 4 pistes principales, ainsi que des pistes secondaires pour afficher les courbes de la trajectoire des mains. Les quatre pistes principales sont : une piste pour annoter le regard, une piste pour annoter les signes manuels, et deux pistes pour annoter le mouvement des mains.

3.2.3.1 Piste d'annotation du regard

Cette piste est consacrée à l'annotation de la direction du regard. 2 valeurs sont définies :

- Espace de signation : Le sujet regard l'espace de signation.
- Fermé : le sujet ferme ses yeux.

3.2.3.2 Piste d'annotation des signes manuels

La piste d'annotation linguistique est destinée à l'annotation des signes. Elle contient 4 attributs : la catégorie linguistique du signe, la structure temporelle du signe, la relation entre les deux mains, et un attribut pour la glose qui est un mot français dont le sens correspond à celui du signe.

L'attribut de catégorie linguistique prend une des valeurs suivantes : TTF, Signe Lexical, TS, Pointage, et Autre.

L'attribut de structure temporelle peut prendre une des valeurs suivantes : Répétition, Composition, et Complexe.

L'attribut de relation entre les deux mains peut prendre une des valeurs suivantes : bi-manuel, Identique, Symétrique, Alterné, Landmark, Buoy, Pointage.

Le détail de chaque valeur est donné en annexe.

Par exemple, le signe illustré en figure 3-12 a comme attributs la glose « armoire », la structure temporelle « composé » car le mouvement est composé de deux trajectoires et la relation entre

les deux mains est « symétrique » car le mouvement de chaque main est globalement symétrique.

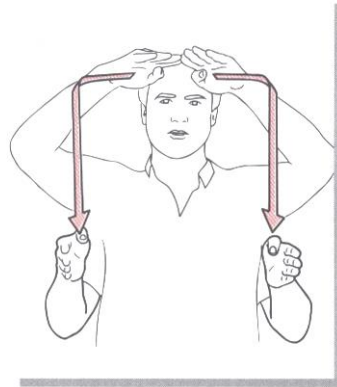


Figure 3-12: Le mot armoire en LSF

3.2.3.3 Piste d'annotation du mouvement de la main dominante

Cette piste est destinée à l'annotation de la direction des mouvements. Pour annoter le mouvement, nous avons créé deux attributs : la direction de la trajectoire du mouvement et la transition intra-signe. Cette piste annote les mouvements de la main dominante (droite pour les droitiers).

Le premier attribut prend une valeur parmi les suivantes : haut, bas, extérieur, intérieur, avant, et arrière. Dans le cadre de l'étude des lois de contrôle moteur, on s'est intéressé uniquement aux mouvements linéaires. La direction principale du geste est retenue et peut être vérifiée a posteriori grâce aux données 3D.

Le deuxième attribut « transition intra-signe » permet d'identifier les mouvements de transition qui se trouvent dans un signe. Ce type de mouvement se trouve dans les signes qui contiennent une structure temporelle comportant une répétition. Cet ensemble contient deux valeurs : Vrai, lorsque c'est une transition intra-signe, et Faux, lorsque ce n'est le cas.

Le détail de chaque valeur est donné en annexe.

Une autre valeur de type booléen « symétrique » permet d'identifier les mouvements qui se font par les deux mains d'une façon symétrique. Cette valeur permet de ne pas annoter la plupart des mouvements de la main dominée (gauche), dans les cas où les mouvements des mains sont symétriques, ce qui représente un gain de temps lors de l'annotation.

Par exemple, le signe « armoire » illustré en figure 3-12 a comme attributs de direction « extérieur » lorsque le mouvement est vers l'extérieur, puis « bas » lorsque le mouvement est vers le bas, ainsi que la valeur « symétrique » car le mouvement est fait par les deux mains de façon symétrique. Cette dernière valeur nous permet de ne pas annoter les mouvements de la main dominée de ce signe.

3.2.3.4 Piste d'annotation du mouvement de la main dominée

Cette piste sert à annoter les mouvements de la main dominée qui ne sont pas symétriques avec la main dominante. Elle contient les mêmes attributs « direction du mouvement » et « transition intra-signes » que pour la piste précédente.

3.2.3.5 Pistes des courbes de la trajectoire des mains

Ces pistes servent à afficher les courbes de trajectoires des deux mains et selon les 3 axes X, Y, et Z. Ces pistes sont affichées selon le besoin. Elles aident à synchroniser la vidéo avec les données mocap et à annoter les mouvements des mains.

Le code du schéma d'annotation est donné dans l'annexe.

3.2.4 La procédure d'annotation

Une fois le schéma d'annotation conçu, le processus d'annotation peut être effectué. Il comporte deux étapes : l'importation (section 3.2.4.1) et l'annotation manuelle (section 3.2.4.2).

3.2.4.1 Importation et synchronisation des données

La première phase consiste à importer les données dans ANVIL : les données de mocap sous forme de squelette dans le format BVH et la vidéo au format AVI. Les courbes des trajectoires (x, y, et z) peuvent être importées, mais aussi affichées à partir du fichier BVH.

La deuxième phase est la synchronisation des données mocap avec la vidéo. ANVIL synchronise par défaut au début des données mais il peut avoir un décalage entre la vidéo et les données mocap. Il faut tout d'abord supprimer l'association que fait ANVIL automatiquement à partir de l'interface d'affichage du squelette. Ensuite, on fait défiler la vidéo et les données mocap séparément jusqu'à ce qu'elles soient synchronisées, en se basant sur les claps produits par le sujet avant de commencer la signation comme mentionné dans la section I.6. Les courbes de trajectoires des mains aident beaucoup dans la synchronisation, car lors du clap, le mouvement s'arrête pendant un petit moment. Une fois que la synchronisation est effectuée, il

faut l'enregistrer. Finalement, une vérification de la synchronisation est faite par rapport aux deux claques que produit le sujet à la fin de la signation.

3.2.4.2 L'annotation manuelle

L'étape d'annotation peut alors débiter. Le choix du début et de la fin d'un segment d'annotation dépend de l'élément annoté. Pour le regard, le segment correspond aux images dans lesquelles les yeux sont fermés (valeur « fermé ») ou ne regardent pas vers la caméra (valeur « espace de signation »)²¹. Pour les signes, le début correspond au moment auquel tous les paramètres manuels des signes (configuration, orientation, emplacement, mouvement) sont stabilisés et la fin au moment où l'un de ces paramètres change (A Braffort et al., 2010). Pour le mouvement, le début et la fin du segment sont choisis par rapport au mouvement des mains du squelette dans la fenêtre dédiée à la visualisation des données de mocap. Le fait de pouvoir tourner le squelette à 360° permet de voir le début du mouvement et la fin, ainsi qu'une comparaison avec la vidéo.

Les phases nécessitant une connaissance de la LSF ont été réalisées par une personne experte en LSF.

Le processus d'annotation des pistes de mouvements s'est déroulé en trois phases.

- Tout d'abord, les deux pistes de mouvements ont été annotées avec les attributs de « direction » et de « symétrie ». Un segment a été créé pour tout mouvement dont la trajectoire avait une direction correspondant à une des six valeurs listées dans la section précédente.
- La deuxième phase a consisté à annoter la piste des signes manuels.
- La troisième phase a consisté à compléter l'annotation des pistes de mouvements avec l'attribut « transition intra-signe ». Cela concerne uniquement les mouvements repérés au sein de signes ayant une structure temporelle qui inclus des parties porteuses de sens et des transitions entre ces parties (par exemple : voir figure 3-13).

²¹ L'étude du regard sortant du cadre de cette thèse, nous en parlons plus dans la suite de ce mémoire.

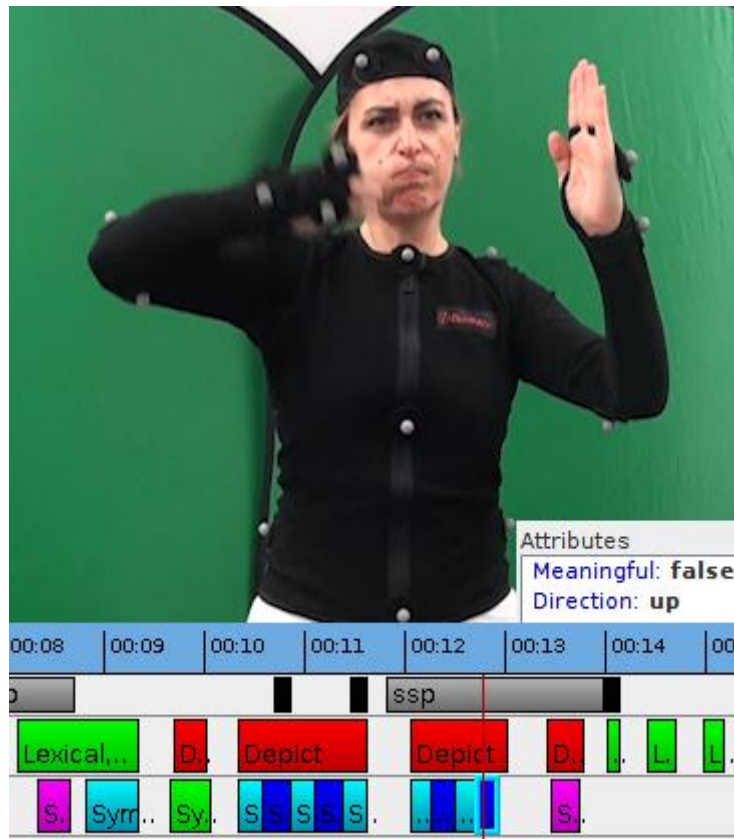


Figure 3-13: Locutrice décrit une scène d'une maison entourée de poteaux, elle décrit les poteaux successivement de haut en bas et entre les deux, il y a une transition vers le haut pour décrire le poteau suivant. Le mouvement de la main dominante (droite) vers le haut (Direction: up), ce mouvement est une transition intra-signe qui ne porte pas d'informations linguistiques (Meaningful: false).

3.2.5 L'export de l'annotation

Une fois l'annotation terminée, elle est exportée dans des fichiers textes de type CSV. L'export s'est fait dans des fichiers distincts pour simplifier l'utilisation de l'annotation dans l'analyse des mouvements.

Le fichier d'annotation des signes manuels contient : l'identifiant de chaque segment d'annotation, le nom de la piste, le temps du début, le temps de fin, la durée du segment, la catégorie du signe, le type de la structure, le type de relation entre les mains si elle existe, et la glose. Le fichier de l'annotation des mouvements contient : l'identifiant de chaque segment, le nom de la piste) le temps du début, le temps de fin, la durée du segment, la direction du mouvement, la valeur de l'attribut « Symétrique », et la valeur de l'attribut « Transition intra-signe ». Ainsi, on dispose de trois fichiers d'annotation pour chaque vidéo : pour les signes, et les mouvements des mains dominante et dominée.

3.3 Conclusion

Nous avons détaillé dans ce chapitre la méthode utilisée pour créer le corpus MOCAP, ainsi que son annotation. MOCAP1 est le premier corpus de LSF disponible pour la communauté scientifique et qui est utilisé dans plusieurs disciplines. Le contenu du corpus est riche, il contient différentes tâches pour qu'il satisfasse les besoins de la communauté.

Nous avons vu que les dispositions des caméras et des marqueurs sont des étapes importantes de la création du corpus. Ainsi, avant de commencer l'enregistrement, la calibration du volume d'enregistrement est nécessaire. En plus le réglage des paramètres des caméras est une étape essentielle pour avoir des données de bonne qualité. Ces paramètres dépendent de l'environnement : l'emplacement du sujet, la disposition des caméras, la disposition des marqueurs et le volume de capture.

Par ailleurs, après l'enregistrement des données, un post-traitement est nécessaire pour nettoyer et exporter les données, en raison de l'occultation possible entre les diverses parties du corps, les marqueurs non identifiés, le chevauchement possible entre les marqueurs, les faux marqueurs, et le bruit. Cette étape est cruciale pour disposer de données exploitables.

Pour rendre les données utilisables par les linguistes et aussi pour analyser les mouvements, les fichiers de mocap et les vidéos ont été importés dans le programme d'annotation ANVIL. L'annotation est composée de trois pistes, dédiées au regard, aux signes manuels et aux mouvements des mains

Les mouvements annotés sont rectilignes dont la trajectoire est orientée selon les trois axes principaux, l'axe transversal, l'axe longitudinal et l'axe antéro-postérieur. Ce sont ces mouvements que l'on va étudier pour vérifier la validité des lois classiques du contrôle moteur. Ces études sont présentées dans le chapitre suivant.

4 Méthode et analyse des données

Nous avons présenté dans le chapitre précédent la création et l'annotation du corpus. Comme nous l'avons dit dans le chapitre II sur l'état de l'art, il y a très peu d'études du mouvement en LS en général et en LSF en particulier. Or le mouvement est la base de la communication en LS. Avant de modéliser les mouvements de la LSF, il apparaît opportun de mieux les connaître. En particulier, il est souhaitable dans un premier temps de comprendre et vérifier la validité des LCCM dans les mouvements de la LSF. Dans ce chapitre, nous allons détailler la procédure du traitement des données de notre corpus utilisé pour investiguer la robustesse ou l'altération des LCCM et nous discuterons les résultats pour chacun des principes de contrôle moteur testé.

La première partie présente la méthode de traitement des données. Dans cette partie, nous détaillons les étapes suivies pour importer les données du corpus et d'annotation, l'organisation des données et l'extraction des données de déplacement, les opérations et les filtres appliqués sur les données selon les principes testés. La deuxième partie présente les résultats obtenus et propose des explications.

Dans ce chapitre, nous nous focalisons sur trois principes importants mis en évidence en contrôle moteur classique et le but est de comparer signes (portant un sens linguistique) et transitions (étant purement motrices) en LSF. Les trois LCCM étudiées sont : le principe d'isochronie, la loi des asymétries et la loi de puissance 2/3.

4.1 Méthode de traitement automatique des données

Avant d'étudier les mouvements de LSF, les données provenant de la mocap doivent être synchronisées, segmentées et filtrées afin de permettre le calcul des paramètres cinématiques des mouvements. L'analyse et l'annotation des vidéos standards est un élément crucial de ce processus. La principale difficulté rencontrée lorsque l'on compare des mouvements de LSF à des données de contrôle moteur classique (ex : pointage, geste d'atteinte de cible, dessin de formes de type ellipse etc.) est que les mouvements de LSF constituent un flux continu de déplacements des segments, sans interruption claire et nette dictée par la tâche ou l'expérimentateur. Il y a donc des étapes de pré-traitement complémentaires pour extraire des données comparables à celles obtenues en contrôle moteur. Le traitement et les calculs que nous présentons ici ont été appliqués aux données du déplacement des marqueurs placés sur les poignets des deux bras.

Cette partie est composée de deux sections, l'organisation des données (section 4.1.1) et les opérations sur le mouvement (section 4.1.2).

4.1.1 L'organisation des données

L'organisation des données est l'étape qui facilite le calcul des paramètres et l'application des filtres. Elle est composée de deux étapes : la synchronisation des annotations avec les données de déplacement en 3D et le filtrage de ces données.

4.1.1.1 La synchronisation de l'annotation

La synchronisation de l'annotation avec les données de mocap est une tâche simple mais nécessaire. Elle nécessite d'ajouter la durée du décalage entre la vidéo et les données mocap au temps de début et de fin de chaque annotation. La synchronisation a été faite par rapport aux deux claques des mains des sujets avant de commencer la signation. L'annotation est utilisée pour distinguer les types de mouvements, principalement pour dissocier les signes des transitions dans cette étude.

Comme l'annotation manuelle était basée sur la vidéo (25 images/s), elle ne peut pas être précise à 100%, ainsi, chaque image correspond à 10 frames de données de mocap (250 Hz). Donc, certains mouvements annotés peuvent ne pas être complets, c.à.d. une ou des parties de ces mouvements peuvent être manquantes lorsqu'on les regarde dans les données 3D de mocap. Ce problème a été résolu en élargissant l'intervalle du mouvement [début, fin] avec 40 frames (20 au début et 20 à la fin), ce qui est l'équivalent de 0.16 seconde. Ainsi, c'est l'équivalent d'ajouter 2 images au début et fin de chaque annotation. L'intervalle élargi va permettre d'agrandir la fenêtre du mouvement afin de s'assurer qu'il ne manque pas une partie du mouvement dans les analyses.

Cet élargissement de la fenêtre de mouvement peut ajouter une partie de mouvement qui n'appartient pas au mouvement annoté, mais ce problème est réglé par l'application de filtres présentés dans la section 2.1 et grâce aux positions x,y, z de la main.

4.1.1.2 Filtre passe-bas

La deuxième étape consiste à appliquer un filtre sur les données de la trajectoire. Comme cela est habituel en contrôle moteur, nous avons appliqué un filtre Butterworth d'ordre 5 et avec une fréquence de coupure égale à 10Hz comme dans (Gaveau et al., 2016, 2011; Gaveau & Papaxanthis, 2011). Comme nous avons dit dans le chapitre précédent ce filtre lisse les signaux de positions en enlevant le bruit de mesure. Cette étape a été réalisée seulement sur la première

tâche du corpus qui consiste à décrire des photos. Pour les autres tâches du corpus, le filtre a été appliqué dans le logiciel Motive comme expliqué dans le chapitre précédent.

4.1.2 Les opérations sur le mouvement

Afin d'étudier les LCCM dans la LSF et de comparer des choses comparables, plusieurs opérations sur les données 3D du mouvement doivent être effectuées. En outre, les lois que nous avons étudiées ont été appliquées sur deux types de mouvements. D'un côté, la loi des asymétries verticales et horizontales et le principe d'isochronie ont été vérifiés sur des mouvements verticaux et horizontaux présentant un début et une fin approximativement à l'arrêt. D'un autre côté, la loi de puissance $2/3$ a été vérifiée sur tous les mouvements quelle que soit leur direction. Cette distinction est importante car en contrôle moteur ce sont des mouvements de pointage vers des cibles (i.e. « point-to-point reaching » en anglais) qui sont typiquement étudiés. Ces derniers partent et finissent toujours à vitesse nulle. Au contraire, en LSF, l'ensemble de l'enregistrement est constitué d'une multitude de mouvements omnidirectionnels sans véritables arrêts imposés. Cependant, les signeurs marquent eux-mêmes des périodes d'arrêt où la vitesse cartésienne de la main tombe à des valeurs proches de zéro, ce qui permet de comparer les mouvements en LSF avec des mouvements analogues habituellement testés en sciences du mouvement humain.

En résumé, les mouvements ont été séparés en deux groupes selon la loi de contrôle moteur testée. Le premier groupe était destiné à l'étude du principe d'isochronie et de la loi des asymétries verticales et horizontales, et le deuxième groupe était destiné à l'étude de la loi de puissance $2/3$. Ainsi, les opérations appliquées sur les mouvements ne sont pas toutes les mêmes. Dans chacun des cas, nous avons fait attention à ce que les principes de contrôle moteur connus pour des mouvements spécifiques soient comparés et comparables à des mouvements similaires apparaissant en LSF. Dans cette section, nous présentons les opérations et les traitements appliqués sur chacun des deux groupes pour rendre cela possible.

4.1.2.1 Les opérations sur les mouvements destinés à l'étude de la loi des asymétries et du principe d'isochronie

Dans cette section, nous présentons les opérations appliquées sur les mouvements destinés à l'étude du principe d'isochronie et de la loi des asymétries verticales et horizontales.

4.1.2.1.1 La segmentation de la trajectoire

La segmentation de la trajectoire des bras est une étape nécessaire pour analyser le mouvement de la LSF, elle permet d'extraire les mouvements pertinents à étudier. Comme la loi des asymétries verticales s'applique sur les mouvements verticaux (ascendant et descendant) qui doivent être comparés à des mouvements horizontaux (vers l'intérieur et vers l'extérieur) et que le principe d'isochronie peut aussi être vérifié sur ce type de mouvements, la segmentation de la trajectoire peut tout d'abord être appliquée en se fondant sur l'annotation directionnelle du mouvement. Pour rappel, l'annotation du mouvement est présentée dans la deuxième partie du chapitre précédent.

En appliquant l'annotation du mouvement sur la trajectoire, nous segmentons d'abord les mouvements horizontaux et verticaux de la trajectoire avec une approche visuelle. Le début de chaque segment correspond au début de l'annotation du mouvement, et la fin du segment correspond à la fin de l'annotation du mouvement. Il est important de noter que, comme nous avons aussi les données 3D du mouvement de la main, la verticalité ou l'horizontalité des mouvements sera ensuite être quantifiée précisément (cf. plus bas), et que cette étape visuelle n'est qu'intermédiaire.

4.1.2.1.2 La coupure à 10%

La coupure à 10% de la vitesse maximale est une étape où les profils de vitesse des mouvements ont été coupés à 10% de la vitesse maximale du mouvement. En d'autres termes, cela veut dire que la partie du profil de vitesse où la vitesse est inférieure à un seuil relatif est supprimée (figure 4-1). Ce seuil est :

$$Seuil = 0.1 * v_{max}$$

Comme nous avons ajouté 40 frames à chaque mouvement afin d'assurer que les mouvements soient complets (section 2.1), cela peut ajouter une partie de mouvement qui n'appartient pas au mouvement original (comme nous avons dit que l'annotation ne peut pas être précise à 100%). Ce filtre permet de corriger tout biais et de ne s'intéresser qu'à une partie du mouvement possédant une accélération et une décélération, en supprimant une petite partie du début et de fin du mouvement. Cela est par ailleurs identique au traitement généralement fait en contrôle moteur (Berret et al., 2009; Gaveau & Papaxanthis, 2011; Gentili et al., 2007; Papaxanthis, Dubost, et al., 2003; Papaxanthis, Pozzo, et al., 2003; Papaxanthis, Pozzo, & Stapley, 1998; Papaxanthis, Pozzo, Vinter, et al., 1998), ce qui permet de se mettre dans les meilleures conditions de comparaison.

L'algorithme utilisé était le suivant : trouver le pic de la vitesse maximale dans le profil de vitesse, puis se déplace sur les deux côtés en comparant les valeurs de vitesse avec la vitesse max, une fois que la valeur est inférieure ou égale à 10% de la vitesse max, couper le profil à cet endroit comme montré dans la figure 4-1.

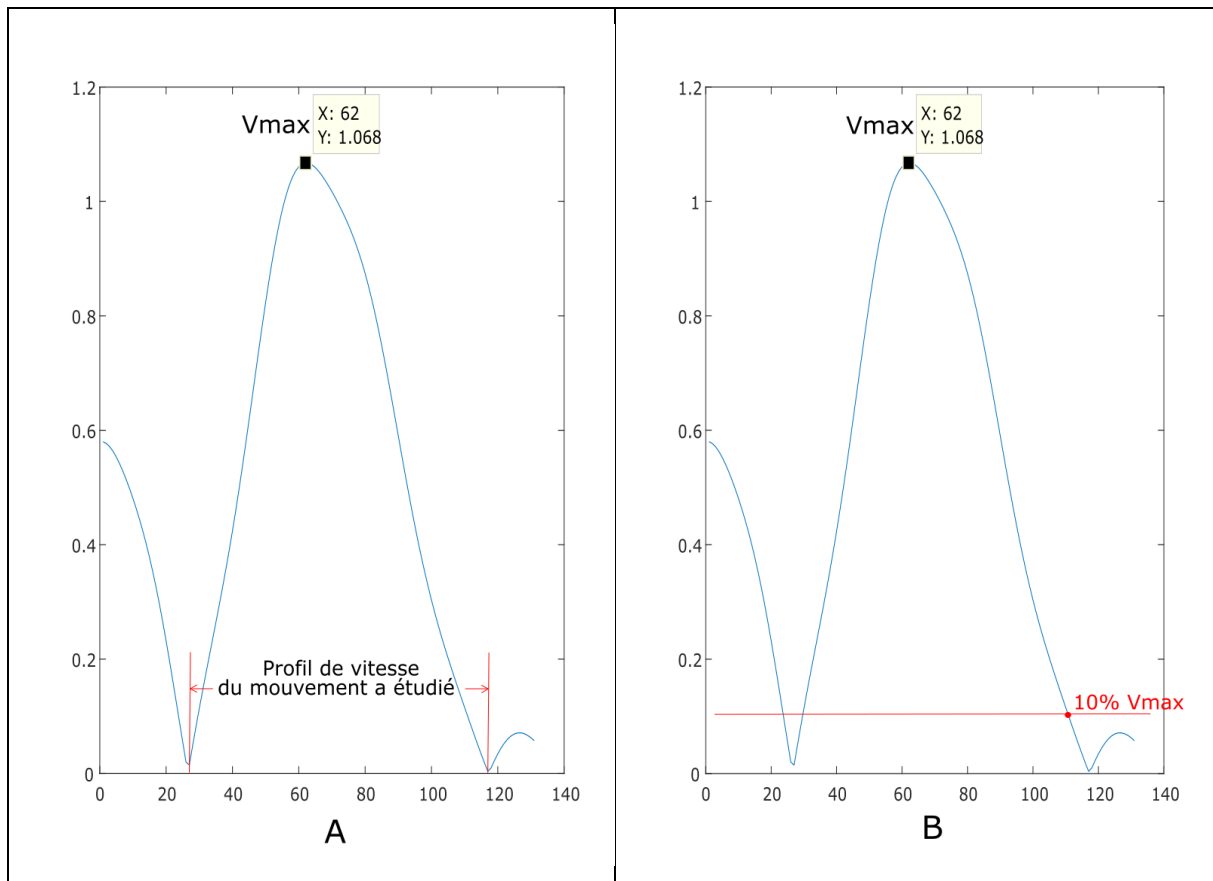


Figure 4-1: Méthode pour couper le profil de vitesse à 10% de la vitesse maximale.

4.1.2.1.3 Les paramètres cinématiques calculés

Afin d'étudier les LCCM en LSF, nous avons calculé des paramètres cinématiques standards du mouvement. Les paramètres principaux calculés sont : la distance parcourue par la main (Dist), la durée du mouvement (DM), la vitesse instantanée du mouvement (V_{inst}), la vitesse moyenne de chaque mouvement (V_{moy}), la vitesse maximale (vitesse de pointe) (V_{max}), l'indice de courbure global qui est défini comme la déviation maximale de la trajectoire de la main par rapport à la ligne droite normalisée par la distance linéaire totale (Curv), la durée d'accélération (DA), la durée de décélération (DD), l'accélération maximale (A_{max}), et la décélération maximale (D_{max}) (voir figure 4-2).

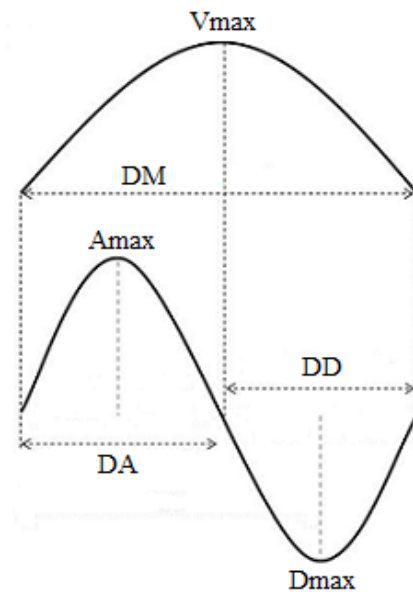


Figure 4-2: Définition des paramètres cinématiques étudiés (Gaveau & Papaxanthis, 2011).

Ainsi, nous avons calculé : le facteur DA/DM qui est la durée d'accélération par rapport à la durée du mouvement (aussi appelé « temps relatif au pic de vitesse » dans d'autres études), le facteur V_{max}/V_{moy} qui est le ratio de la vitesse maximale par rapport à la vitesse moyenne (et donne une information sur le caractère plus ou moins aplati du profil de vitesse), et le ratio d'horizontalité ou de verticalité que nous expliquons ci-dessous dans la section 2.1.4.4.

4.1.2.1.4 Les filtres appliqués

L'étude des LCCM doit être faite sur des mouvements qui présentent des caractéristiques communes entre signes et transitions (et aussi vis-à-vis des caractéristiques des mouvements de pointage étudiés en contrôle moteur). Les paramètres déterminants étaient la distance parcourue, la vitesse moyenne, et la durée du mouvement. Ces contraintes nous ont obligés à mettre un système de sélection automatique des mouvements à étudier, fondé sur des critères (filtres) bien précis.

Ces filtres ont pour but d'éliminer tous les mouvements qui ne respectent pas les contraintes que nous avons fixées. Dans cette partie, nous détaillons les contraintes et les filtres que nous avons choisis d'appliquer sur les mouvements horizontaux et verticaux.

a. Distance minimale et maximale

Afin que les distances (longueurs curvilignes) des mouvements à étudier soient raisonnables, nous avons appliqué un filtre qui élimine les mouvements trop petits et trop grands. Le critère a aussi été choisi en se basant sur la volonté d'avoir des distances moyennes des mouvements

de type « signe » égales à celles des mouvements de type « transition ». Le seuil minimal et le seuil maximal étaient 5 cm et 40 cm respectivement. Tous les mouvements qui ont une distance en dehors de ces intervalles ont donc été éliminés.

Nous avons appliqué ces filtres car les mouvements verticaux et horizontaux dans les signes sont généralement petits et ne dépassent pas les 40 cm contrairement aux transitions qui ont parfois des distances plus grandes, ce qui pourrait alors créer un biais dans les comparaisons signe versus transition. En effet, on sait en contrôle moteur que les grands mouvements ont tendance à être plus rapides tout en ayant une durée plus longue par rapport aux petits mouvements (cf. principe d'isochronie).

b. La vitesse moyenne minimale et maximale

Le deuxième critère élimine les mouvements dont la vitesse est trop grande ou trop petite. A la différence des protocoles de contrôle moteur où la vitesse peut être imposée ou être relativement stable, les mouvements en LSF ont des vitesses beaucoup moins régulières. Ce critère a été choisi pour que les mouvements à étudier aient des vitesses moyennes homogènes. Il élimine donc les mouvements dont la vitesse est hors norme (et donc atypique pour des mouvements hors LSF). Nous avons choisi une formule simple afin de calculer les deux seuils. La formule est la suivante:

$$s_{max} = v_{max} - \sigma_v$$

$$s_{min} = v_{min} + \sigma_v$$

où s_{min} : le seuil minimal, s_{max} : le seuil maximal, v_{max} : la vitesse maximale, v_{min} : la vitesse minimale, et σ_v : l'écart-type de la vitesse moyenne.

Cette formule a donné des seuils différents pour chaque sujet, ce qui est logique, car les sujets signe avec des vitesses assez différentes.

c. Différence de vitesse entre début et fin à 10%

Comme les mouvements de la LSF sont continus, les profils de vitesse des mouvements annotés ne commencent et ne se terminent pas toujours à l'arrêt. Cela veut dire que les mouvements ne s'arrêtent pas réellement dans de nombreux cas, malgré des changements notables de direction. Par exemple, comme nous l'avons déjà mentionné, les signes sont suivis par des transitions et inversement les transitions sont suivies par des signes. Nous voulons comparer les résultats de

notre étude avec ce qui est connu dans la littérature, où les profils de vitesse étudiée commencent et se terminent à la vitesse quasi nulle.

À cet effet, nous avons créé ce filtre qui a pour but d'éliminer les mouvements dont la différence entre les valeurs du début et de la fin dont le profil de vitesse est supérieur à 10% de la vitesse maximale du mouvement (figure 4-3).

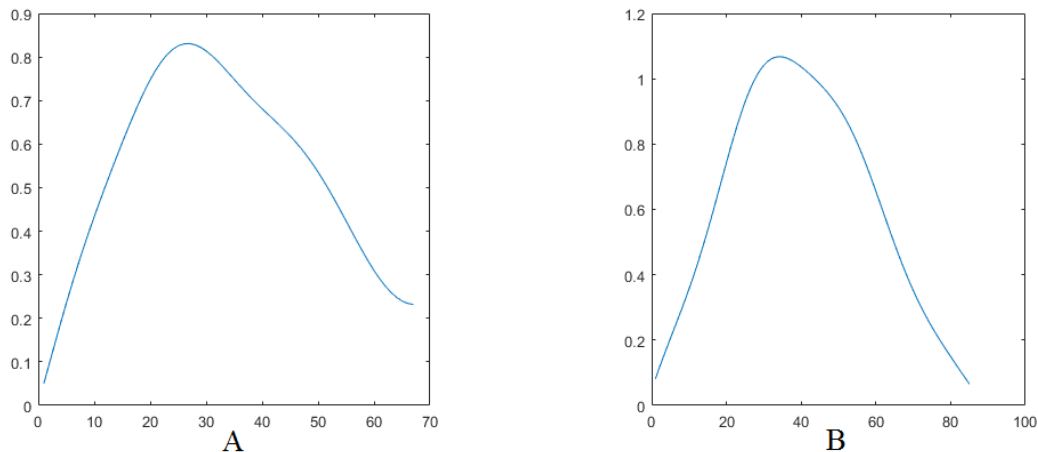


Figure 4-3: A : profil de vitesse éliminé, B : profil de vitesse accepté

d. Ratio d'horizontalité ou de verticalité

Comme nous avons dit dans la section précédente, l'annotation a été réalisée manuellement en se basant sur les vidéos, le squelette 3D, et les courbes des trajectoires. Malgré cette variété des ressources, l'annotation des mouvements a été un peu délicate voir arbitraire dans certains cas, car dans certains mouvements en 3D il est difficile de décider si le mouvement est vertical ou horizontal. Heureusement, les données de mocap permettent de vérifier l'horizontalité ou la verticalité d'un mouvement de la main très précisément.

Le ratio de déplacement vertical ou horizontal est le pourcentage de distance parcourue sur le plan horizontal ou vertical (selon la direction du mouvement) par rapport à la distance globale du mouvement dans l'espace. Ce ratio était calculé avec la formule suivante :

$$R_{h/v} = \frac{Dist_{h/v}}{Dist} * 100\%$$

où $R_{h/v}$ est le ratio d'horizontalité ou de verticalité, $Dist_{h/v}$ est la distance parcourue sur le plan horizontal ou vertical, et $Dist$ est la distance totale curviligne parcourue dans l'espace par le marqueur de la main étudiée.

Afin d'assurer l'horizontalité ou la verticalité d'un mouvement, nous avons mis en place un filtre qui élimine les mouvements dont le ratio est inférieur à 80% ($R_{h/v} < 80\%$).

L'application de ces filtres nous a permis d'éliminer tous les mouvements incompatibles avec les données de contrôle moteur habituelles qui s'effectuent dans des plans bien définis (ex : plan sagittal ou transverse), et de garder des mouvements de signes et de transition qui ont des distances moyennes, des vitesses moyennes, et des durées moyennes similaires. Cela nous permettra de comparer les résultats trouvés dans les signes avec ceux des transitions ainsi qu'avec les études faites dans la littérature du contrôle moteur.

À la fin de l'application de ces filtres, une inspection visuelle de tous les profils de vitesse a confirmé qu'ils avaient tous un seul pic et une forme en cloche.

4.1.2.1.5 Normalisation des profils de vitesse

Pour illustrer qualitativement des similitudes ou des dissemblances dans la cinématique des mouvements, les profils de vitesse ont été normalisés en durée et en amplitude pour chaque essai et chaque participant comme dans (Gaveau & Papaxanthis, 2011). La normalisation garantit que les profils de vitesse étaient indépendants de la distance parcourue, de la vitesse de déplacement et de la durée du mouvement et permet de voir des différences sur la forme de ces profils.

4.1.2.1.6 La correspondance entre les mouvements et leurs types

Trouver la correspondance entre les mouvements et leurs types (signe ou transition) est une étape cruciale pour analyser séparément les mouvements des signes (porteurs d'information linguistique) et des transitions (exempt de tout sens linguistique), autrement dit, trouver pour chaque mouvement sa correspondance (signe ou transition). Comme l'annotation du mouvement était séparée de l'annotation linguistique, nous n'avons pas traité pour le moment l'information qui lie le mouvement avec son type (signe ou transition).

Afin de s'assurer qu'un mouvement 3D appartient à un signe ou à une transition, nous avons décidé d'agrandir l'intervalle des signes [fin, début] avec 40 frames (20 au début et 20 à la fin) l'équivalent de 0.16 seconde, en utilisant la piste d'annotation linguistique.

La recherche des correspondances a été réalisée en deux étapes. La première étape a consisté à comparer le début et la fin de chaque mouvement avec le début et la fin de chaque signe. Si l'intervalle [début, fin] d'un mouvement 3D est inclus dans l'intervalle [début, fin] d'un signe

(annoté comme tel) alors ce mouvement appartient bien à ce signe, sinon ce mouvement est une transition²².

La deuxième étape a consisté à trouver les mouvements qui sont des transitions intra-signes (c.à.d. une transition dans un signe) et les considérer comme des transitions. En effet, dans un signe qui contient une ou plusieurs répétitions, il y a une ou plusieurs transitions. L'annotation du mouvement contient cette information qui indique si le mouvement est une transition intra-signes ou non. Nous avons vérifié cette information dans tous les mouvements qui ont été classés comme des signes dans la première étape, et ceux qui sont des transitions intra-signes ont été déplacés dans les mouvements de transition.

4.1.2.2 Les opérations sur les mouvements destinés à l'étude de la loi de puissance 2/3

Dans cette partie, nous présentons les opérations et tout le traitement et les calculs appliqués sur les mouvements destinés à l'étude de la loi de puissance 2/3.

4.1.2.2.1 La segmentation de la trajectoire

La segmentation de la trajectoire pour étudier la loi de puissance 2/3 est basée sur la segmentation linguistique. Afin d'étudier cette loi, nous avons besoin des mouvements des signes et des transitions quelles que soient leurs directions. Cette segmentation était ainsi basée sur l'annotation linguistique afin de savoir quels sont les mouvements des signes et les mouvements des transitions.

Les mouvements des signes correspondent à l'intervalle [début du signe, fin du signe] de chaque signe. En ce qui concerne les mouvements des transitions, c'est l'inverse : le début de chaque segment correspond à la fin du signe, et la fin du segment correspond au début du signe suivant. Autrement dit, l'intervalle des mouvements des transitions correspond à [fin du signe, début du signe suivant].

4.1.2.2.2 Les paramètres cinématiques calculés

Les paramètres cinématiques de ces mouvements ont été calculés avec la même méthode et les mêmes fonctions que les premiers mouvements. Afin d'étudier la loi de puissance 2/3, la courbure instantanée de chaque mouvement de la trajectoire doit être calculée.

²² Les mouvements qui commencent comme signe et se terminent comme transition ou inversement, sont éliminés de l'étude.

a. La courbure instantanée du mouvement

Comme la loi de 2/3 de puissance est une relation entre la courbure de la trajectoire et la vitesse du mouvement, nous avons calculé la courbure de chaque mouvement à étudier avec la formule suivante :

$$k = \frac{\sqrt{(z''y' - y''z')^2 + (x''z' - z''x')^2 + (y''x' - x''y')^2}}{(x'^2 + y'^2 + z'^2)^{\frac{3}{2}}}$$

où x'' , y'' , et z'' sont les accélérations instantanées et x' , y' , et z' sont les vitesses instantanées selon les axes x, y, et z.

4.1.2.2.3 Les filtres appliqués

Nous avons appliqué les mêmes filtres que sur les premiers mouvements sauf pour ce qui concerne la différence entre début et fin à 10% et le ratio de verticalité/horizontalité puisqu'ils n'étaient pas adaptés à cette loi du contrôle moteur qui implique des mouvements cycliques ou « lancés » (la vitesse ne devant pas être nulle pour tester la loi de puissance 2/3). Par contre, nous avons ajouté un autre filtre sur la vitesse minimale, où nous avons éliminé tous les mouvements dont les profils de vitesse tombaient en dessous de 0.05 m/s. Le but de ce filtre est d'enlever les mouvements qui contiennent un ou plusieurs arrêts autres que l'arrêt final car cela pose des problèmes dans l'évaluation du rayon de courbure (cf. division par la vitesse cartésienne au cube).

Le seuil minimal et le seuil maximal de la distance des mouvements étaient ici 10 cm et 80 cm respectivement. Tous les mouvements qui ont une distance en dehors de cet intervalle sont éliminés. Ici il est moins essentiel de préserver des distances totales comparables, car la loi de puissance 2/3 est une propriété « locale » entre courbure et vitesse instantanée. La coupure du profil de vitesse à 10% de la vitesse maximale a été choisie et appliquée de la même manière que dans les premiers mouvements.

Une fois les mouvements filtrés, et les paramètres cinématiques sont calculés, l'étude de ces lois et la comparaison signe/transition peut commencer. Dans la section suivante, nous présentons les résultats de ces études et nous les discuterons.

4.2 Résultats et discussions

Cette section décrit les résultats et les discussions de l'étude des trois lois abordées. Premièrement, nous présentons le principe d'isochronie qui donne des informations

« macroscopiques » sur les relations entre distance, durée et vitesse des mouvements (section 4.2.1). Ce principe a été étudié selon le type et la direction du mouvement. Deuxièmement, la loi des asymétries a été étudiée selon le type et la direction du mouvement (section 4.2.2). Cette analyse est plus fine et s'intéresse à la forme des profils de vitesse dans les différentes conditions de mouvement. Troisièmement, la loi de puissance $2/3$ a été étudiée selon le type de mouvement, pour des mouvements 3D dans toutes les directions pour voir les relations entre forme de la trajectoire (courbure) et vitesse tangentielle de la main (section 4.2.3).

L'objectif de ces études est de vérifier la robustesse de ces lois, énormément documentées en contrôle moteur, durant les mouvements de la LSF. La question que nous avons posée est : la transmission de l'information linguistique dans le contexte de LSF affecte-t-elle la cinématique des mouvements de type signe ? Une comparaison directe avec les mouvements de transition est donc envisagée dans chacun des cas.

4.2.1 Principe d'isochronie et relation distance-durée

Le principe isochronie est essentiellement une relation croissante entre la distance et la vitesse moyenne. Ce principe stipule que la vitesse moyenne (ou vitesse maximale) des mouvements d'un point à un autre augmente avec la distance entre les cibles et que, par conséquent, la durée du mouvement ne dépend que faiblement de l'augmentation de l'amplitude du déplacement. Les études ont montré qu'un tel phénomène tendait à se retrouver dans presque tous les types de mouvements, périodiques et apériodiques, réguliers et improvisés (voir Chapitre 2). Cependant, l'augmentation de la vitesse n'est généralement pas suffisante pour réellement garder une durée constante quelle que soit l'amplitude du mouvement, et une relation distance-durée croissante est aussi observée en règle générale. Nous étudierons donc les relations distance-vitesse et distance-durée dans la suite.

Avant d'étudier le principe, nous présentons dans le tableau 4-1, les valeurs moyennes des paramètres des mouvements du poignet dans les deux types de mouvements et les quatre directions. Ces paramètres sont : la distance, la durée du mouvement, la vitesse moyenne, la vitesse de pointe, la durée d'accélération et de décélération, et l'indice de courbure global.

Le tableau 4-1 montre que la distance et la durée des mouvements des transitions étaient un peu plus grandes que celles des signes. Le tableau montre aussi que les durées d'accélération et de décélération n'étaient pas égales, autrement dit les profils de vitesse n'étaient pas symétriques.

Direction du mouvement	Type de mouvement	Dist	DM	Vmoy	Vmax	DA	DD	Curv
Haut	Signe	0.177	0.269	0.696	1.622	0.143	0.125	0.102
	Transition	0.220	0.320	0.734	1.593	0.153	0.168	0.106
Bas	Signe	0.166	0.257	0.668	1.667	0.146	0.109	0.089
	Transition	0.223	0.323	0.737	1.594	0.174	0.149	0.096
Extérieur	Signe	0.192	0.279	0.713	1.633	0.161	0.117	0.091
	Transition	0.210	0.294	0.769	1.506	0.160	0.134	0.107
Intérieur	Signe	0.202	0.255	0.824	1.511	0.150	0.104	0.118
	Transition	0.212	0.279	0.807	1.509	0.146	0.134	0.077
Moyenne	Signe	0.184 ±0.016	0.265 ±0.011	0.725 ±0.068	1.608 ±0.068	0.150 ±0.008	0.114 ±0.009	0.100 ±0.013
	Transition	0.216 ±0.006	0.304 ±0.021	0.762 ±0.034	1.551 ±0.050	0.158 ±0.012	0.146 ±0.016	0.097 ±0.014

Tableau 4-1: Valeurs moyennes des paramètres de la cinématique du poignet dans les deux types de mouvements et les quatre directions.

Ce tableau donne une vue globale sur les paramètres des mouvements à étudier. Dans la suite, nous affinons les résultats pour étudier le principe d'isochronie.

Les mesures ont été soumises à une analyse de variance (ANOVA) à mesures répétées avec deux facteurs car elles ont montré une distribution normale (test de Kolmogorov-Smirnov). Les facteurs étaient du type de mouvement (signe et transition) et le plan du mouvement (horizontal et vertical). Une valeur de $P < 0.05$ indique que le résultat est significatif.

4.2.1.1 Résultats

La validité de ce principe a été testée séparément pour les mouvements des signes et des transitions. Au niveau qualitatif, cela a été fait en calculant la vitesse moyenne, la distance et la durée de chaque mouvement. Ainsi, nous avons tracé la variation de la vitesse moyenne et la durée en fonction de la distance (Vmoy vs Dist et DM vs Dist), en traçant en même temps la courbe de tendance pour chacun des deux paramètres (régression linéaire).

Le principe d'isochronie serait strictement valide si la courbe de tendance de DM vs Dist est parallèle à l'axe horizontal. Cependant une relation linéaire croissante de la relation Vmoy vs Dist permettrait de conclure à une validité partielle du principe d'isochronie.

Les figures 4-4 et 4-5 présentent les tracés de la vitesse moyenne et la durée en fonction de la distance du mouvement dans les signes (en bleu) et les transitions (en rouge) d'un sujet type. Nous avons aussi tracé les courbes de tendance en affichant leurs équations et les coefficients de détermination R^2 . Les résultats qui correspondent à chaque sujet sont présentés dans les tableaux 4-2 et 4-3. Dans ces tableaux, nous présentons les R^2 , les pentes des courbes de tendance ainsi que leurs ordonnées à l'origine pour la vitesse moyenne et pour la durée.

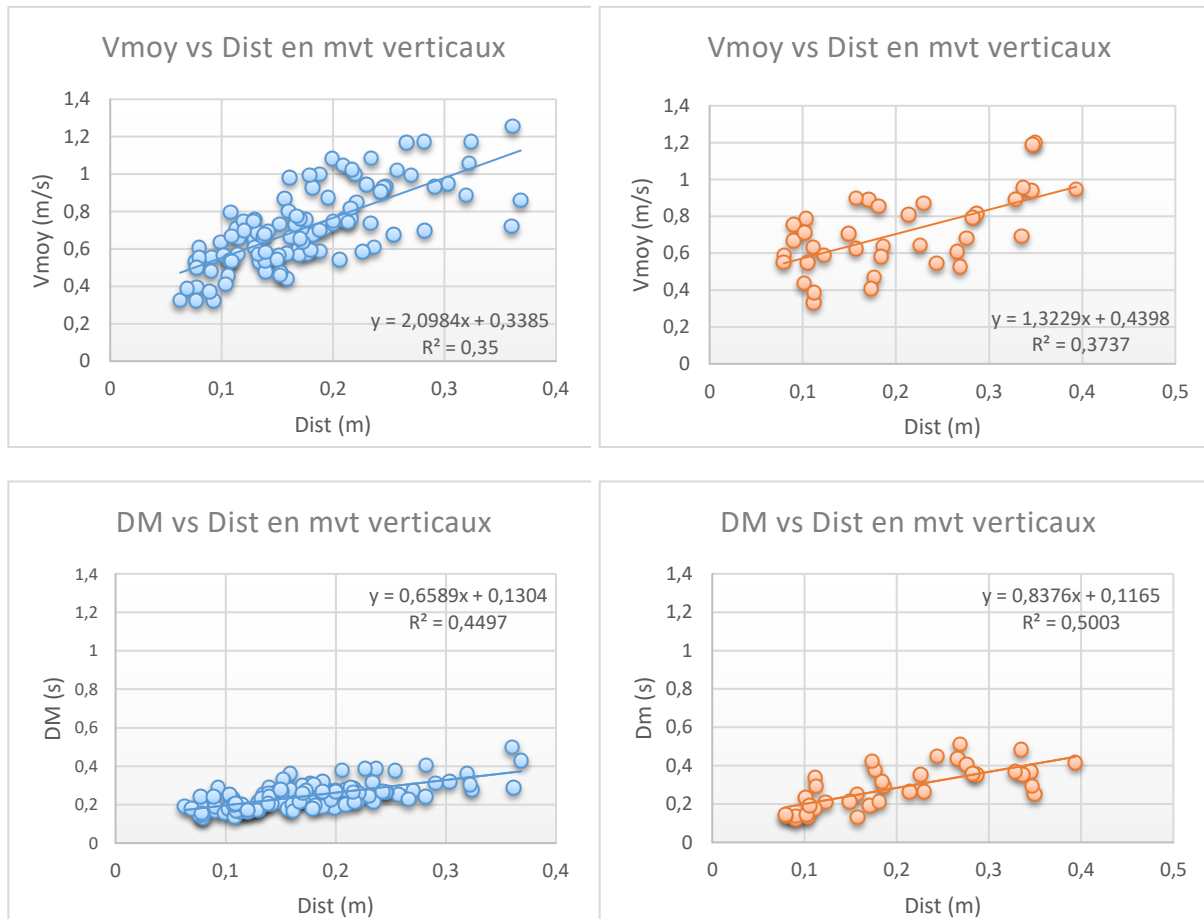


Figure 4-4: le tracé de la vitesse moyenne (V_{moy}) et la durée du mouvement en fonction de la distance, en bleu : les signes, en rouge : les transitions, pour tous les mouvements verticaux d'un sujet. En abscisses, la distance en mètre. En ordonnées, la durée en seconde et la vitesse en m/s.

4.2.1.1.1 V_{moy} vs Dist

Les rapports entre vitesse moyenne et distance sont donnés pour chaque sujet dans le tableau 4-2. Les valeurs moyennes de R^2 dans les signes étaient ($M = 0.50$, $SD = 0.06$ et $M = 0.42$, $SD = 0.14$ dans les mouvements verticaux et horizontaux respectivement). Les valeurs étaient beaucoup plus faibles pour les mouvements de transition ($M = 0.32$, $SD = 0.15$ et $M = 0.23$, $SD = 0.13$ dans les mouvements verticaux et horizontaux respectivement). L'ANOVA a montré qu'il n'y avait pas d'effet significatif de l'interaction type*plan ($F_{1,7} = 0.008$, $P = 0.23$) et de la

direction ($F_{1,7} = 2.001$, $P = 0.23$) sur le R^2 . Par contre, l'analyse statistique a montré un effet significatif du type de mouvement (signe vs transition) sur le R^2 ($F_{1,7} = 42.379$, $P < 0.001$).

Pour, les valeurs moyennes de la pente des courbes de tendance dans les signes, les valeurs étaient ($M = 2.13$, $SD = 0.40$ et $M = 2.05$, $SD = 0.50$ dans les mouvements verticaux et horizontaux respectivement). De même les valeurs de la pente étaient aussi beaucoup plus faibles pour les mouvements de transition ($M = 1.57$, $SD = 0.49$ et $M = 1.46$, $SD = 0.57$ dans les mouvements verticaux et horizontaux respectivement). L'analyse ANOVA a montré aussi qu'il n'y avait pas d'effet significatif de l'interaction type*plan ($F_{1,7} = 0.026$, $P = 0.23$), et de la direction ($F_{1,7} = 0.242$, $P = 0.23$) sur la pente. Par contre, elle a encore une fois montré un effet significatif du type de mouvement sur la pente ($F_{1,7} = 57.251$, $P < 0.001$).

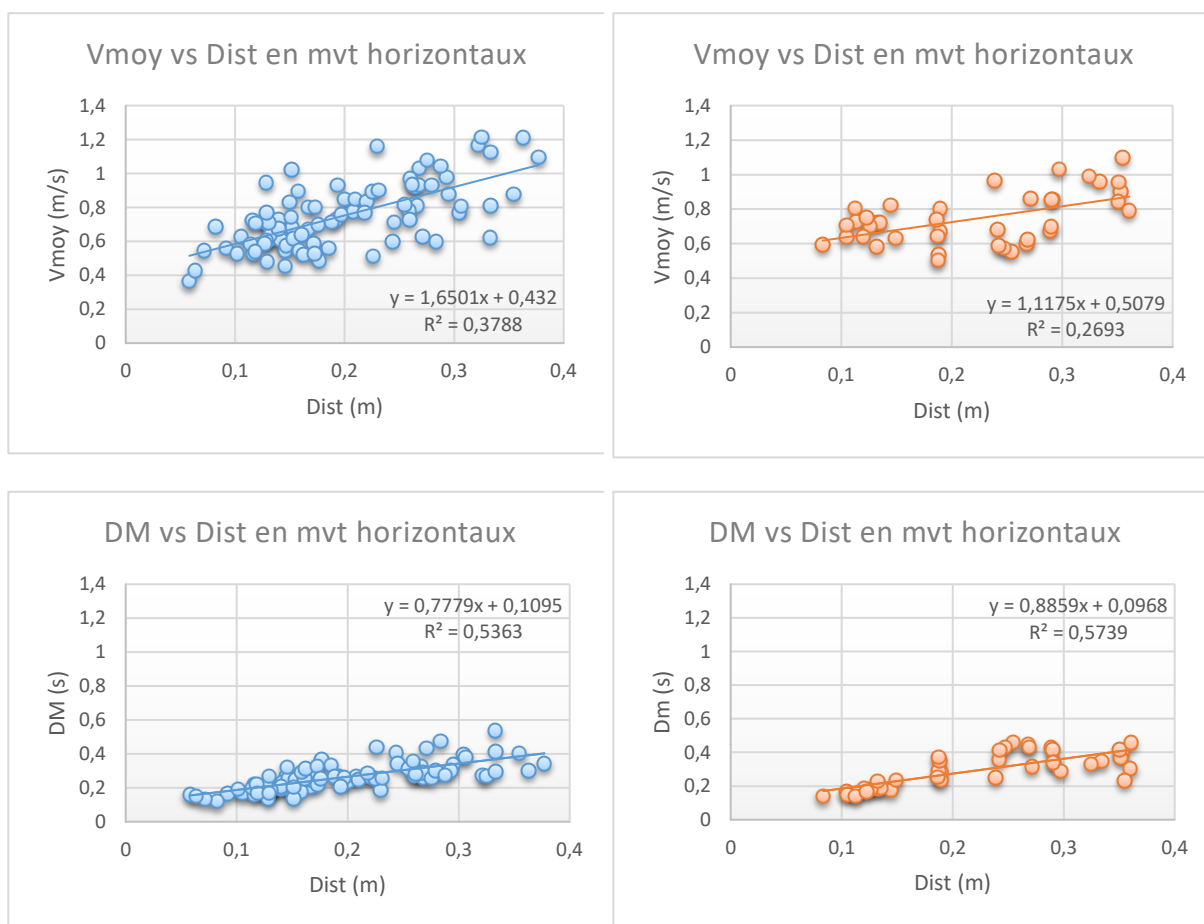


Figure 4-5: Tracé de la vitesse moyenne (Vmoy) et la durée du mouvement en fonction de la distance, en bleu : les signes, en rouge : les transitions, pour tous les mouvements horizontaux d'un sujet. En abscisses, la distance en mètre. En ordonnées, la durée en seconde et la vitesse en m/s.

Enfin, les valeurs moyennes de l'ordonnée des courbes de tendance dans les signes étaient ($M = 0.32$, $SD = 0.06$ et $M = 0.35$, $SD = 0.07$ dans les mouvements verticaux et horizontaux

respectivement). Contrairement aux valeurs de R^2 et de la pente, les valeurs de l'ordonnée étaient plus grandes dans les transitions ($M = 0.39$, $SD = 0.11$ et $M = 0.48$, $SD = 0.12$) dans les mouvements verticaux et horizontaux respectivement). L'ANOVA a montré également qu'il n'y avait pas d'effet significatif de l'interaction type*plan ($F_{1,7} = 3.047$, $P = 0.23$), et de la direction ($F_{1,7} = 3.068$, $P = 0.23$) sur l'ordonnée. Par contre, elle a également montré un effet significatif du type de mouvement sur l'ordonnée ($F_{1,7} = 20.572$, $P < 0.01$).

Vmoy vs Dist												
Sujets	Mouvements verticaux						Mouvements horizontaux					
	Les mouvements de Signes			Les mouvements de transitions			Les mouvements de Signes			Les mouvements de transitions		
	R^2	Pente	Ordonnée	R^2	Pente	Ordonnée	R^2	Pente	Ordonnée	R^2	Pente	Ordonnée
S1	0.53	1.55	0.27	0.29	1.52	0.33	0.34	1.07	0.28	0.12	0.94	0.45
S2	0.51	1.81	0.20	0.24	1.39	0.27	0.34	1.60	0.29	0.22	1.31	0.40
S3	0.50	2.21	0.37	0.51	2.09	0.37	0.16	1.20	0.51	0.49	2.04	0.41
S4	0.50	2.31	0.33	0.57	2.44	0.27	0.13	1.20	0.50	0.25	2.27	0.40
S5	0.39	2.88	0.35	0.29	2.31	0.45	0.31	2.53	0.51	0.04	0.80	0.76
S6	0.48	2.10	0.35	0.38	1.65	0.43	0.37	1.32	0.44	0.27	1.12	0.51
S7	0.51	2.28	0.35	0.58	2.85	0.27	0.30	1.87	0.30	0.26	2.05	0.40
S8	0.59	1.92	0.37	0.50	2.17	0.37	0.62	1.79	0.32	0.19	1.15	0.52
Moy	0.50 ± 0.06	2.13 ± 0.40	0.32 ± 0.06	0.42 ± 0.14	2.05 ± 0.50	0.35 ± 0.07	0.32 ± 0.15	1.57 ± 0.49	0.39 ± 0.11	0.23 ± 0.13	1.46 ± 0.57	0.48 ± 0.12

Tableau 4-2: Valeurs de R^2 , la pente et l'ordonnée des courbes de tendance des mouvements de signes des transitions selon la direction pour V_{moy} en fonction de la distance.

4.2.1.1.2 DM vs Dist

Les rapports entre la durée du mouvement et la distance sont donnés pour chaque sujet dans le tableau 4-3. Premièrement, les valeurs moyennes de R^2 dans les signes et les transitions étaient assez proche. Dans les signes les valeurs étaient ($M = 0.44$, $SD = 0.13$ et $M = 0.41$, $SD = 0.08$ dans les mouvements verticaux et horizontaux respectivement), dans les transitions, elles étaient ($M = 0.37$, $SD = 0.13$ et $M = 0.42$, $SD = 0.11$ dans les mouvements verticaux et horizontaux respectivement). Ce résultat a été confirmé par l'analyse ANOVA, qui a montré un

effet non significatif de l'interaction type*plan ($F_{1,7} = 3.154, P = 0.23$), du type de mouvement ($F_{1,7} = 2.068, P = 0.23$), et de la direction ($F_{1,7} = 0.001, P = 0.23$) sur le R^2 .

Deuxièmement, les valeurs moyennes de la pente des courbes de tendance dans les signes étaient ($M = 0.72, SD = 0.16$ et $M = 0.72, SD = 0.24$) dans les mouvements verticaux et horizontaux respectivement). Les valeurs étaient plus grandes dans les transitions ($M = 0.77, SD = 0.20$) et ($M = 0.84, SD = 0.17$) dans les mouvements verticaux et horizontaux respectivement. L'analyse ANOVA n'a pas confirmé ce résultat, où elle a montré qu'il n'y a pas d'effet significatif ni l'interaction type*plan ($F_{1,7} = 0.691, P = 0.23$), ni du type de mouvement ($F_{1,7} = 3.067, P = 0.23$), ni de la direction ($F_{1,7} = 0.390, P = 0.23$) sur la pente.

DM vs Dist												
Sujets	Mouvements verticaux						Mouvements horizontaux					
	Les mouvements de Signes			Les mouvements de transitions			Les mouvements de Signes			Les mouvements de transitions		
	R ²	Pente	Ordonnée	R ²	Pente	Ordonnée	R ²	Pente	Ordonnée	R ²	Pente	Ordonnée
S1	0.54	1.02	0.14	0.41	0.93	0.14	0.48	1.21	0.17	0.47	1.13	0.10
S2	0.35	0.86	0.19	0.45	1.17	0.15	0.25	0.74	0.20	0.40	0.91	0.13
S3	0.43	0.69	0.11	0.47	0.63	0.12	0.42	0.81	0.11	0.51	0.56	0.13
S4	0.34	0.64	0.13	0.27	0.44	0.18	0.24	0.77	0.14	0.27	0.92	0.07
S5	0.27	0.53	0.12	0.35	0.67	0.09	0.34	0.54	0.10	0.38	0.78	0.08
S6	0.45	0.62	0.13	0.54	0.78	0.11	0.50	0.84	0.12	0.57	0.89	0.10
S7	0.45	0.58	0.13	0.38	0.48	0.14	0.21	0.62	0.20	0.27	0.67	0.13
S8	0.70	0.74	0.11	0.43	0.62	0.12	0.56	0.66	0.16	0.46	0.87	0.10
Moy	0.44 ±0.13	0.72 ±0.16	0.13 ±0.03	0.41 ±0.08	0.72 ±0.24	0.13 ±0.03	0.37 ±0.13	0.77 ±0.20	0.15 ±0.04	0.42 ±0.11	0.84 ±0.17	0.10 ±0.02

Tableau 4-3: Les valeurs de R², la pente et l'ordonnée des courbes de tendance des mouvements de signes des transitions selon la direction pour la durée en fonction de la distance.

Finalement, les valeurs moyennes de l'ordonnée des courbes de tendance dans les signes et les transitions étaient assez proches, où dans les signes les valeurs étaient ($M = 0.13, SD = 0.03$ et $M = 0.13, SD = 0.03$ dans les mouvements verticaux et horizontaux respectivement), et dans les transitions les valeurs étaient ($M = 0.15, SD = 0.04$ et $M = 0.10, SD = 0.02$ dans les mouvements verticaux et horizontaux respectivement). L'ANOVA a montré un effet non

significatif de l'interaction type*plan ($F_{1,7} = 5.830$, $P =$), et du type de mouvement ($F_{1,7} = 0.205$, $P =$). Par contre, elle a montré un effet significatif de la direction sur l'ordonnée ($F_{1,7} = 9.7776$, $P < 0.05$).

4.2.1.2 Discussion

Dans la présente étude, nous avons tenté de vérifier la robustesse du principe d'isochronie dans les mouvements des signes et des transitions.

Un résultat important de cette étude concerne les valeurs de la pente des courbes de régression linéaire et le coefficient de détermination R^2 . En regardant les valeurs des pentes, nous trouvons que dans l'analyse V_{moy} vs Dist, les valeurs étaient significativement plus grandes dans les signes que dans les transitions. A l'inverse, dans l'analyse DM vs Dist, il n'y avait pas de différence significative entre les pentes des signes et des transitions.

Les valeurs du R^2 de V_{moy} vs Dist étaient aussi significativement plus grandes dans les signes que dans les transitions. Par contre, il n'y avait pas de différence significative entre R^2 des signes et des transitions de DM vs Dist.

Un résultat intéressant aussi est qu'il n'y avait pas d'effet de la direction du mouvement sur les deux valeurs (la pente et le R^2) et dans les deux analyses (V_{moy} vs Dist, et DM vs Dist).

Les valeurs de la pente des courbes de tendance du facteur V_{moy} vs Dist étaient supérieures à 1.5 dans les signes et les transitions. Cependant, dans le facteur DM vs Dist, elle était autour de 0.76. Ce qui veut dire que la vitesse moyenne augmente avec la distance du mouvement, par conséquent, la durée du mouvement dépendant plus faiblement de l'augmentation de la distance. Ce résultat permet de dire que le principe d'isochronie reste dans l'ensemble valide dans les signes et dans les transitions.

Afin de comparer les valeurs V_{moy} vs Dist avec des mouvements de pointage dans le plan horizontal, nous avons réalisé une étude sur les mouvements de pointage avec 18 participants. La valeur moyenne de la pente était 0.9. Il est intéressant de noter que la valeur de pente dans les signes et les transitions était plus importante et plus grande comparée à cette étude de mouvements de pointage (0.9) et à celle d'atteinte de cible des singes où elle était autour de 1 (Sartori et al., 2013), cela permet de dire que le principe d'isochronie était plus prononcé dans les mouvements de LSF que dans les mouvements de pointage ou d'atteinte de cible.

En comparant les signes et les transitions, la pente des régressions linéaires de l'analyse V_{moy} vs $Dist$, était significativement plus grande dans les signes. Cela veut dire que le principe d'isochronie est plus marqué dans les signes que dans les transitions. En d'autres termes, dans les signes, la vitesse des mouvements avait tendance à changer beaucoup plus fortement en fonction de l'amplitude par rapport au cas des transitions. Ce qui est important à retenir est que les signes varient plus leur vitesse en fonction de l'amplitude, alors que les transitions impliquent moins de variations de vitesses. Ce résultat peut être lié au fait que des signes amples sont souvent associés à des notions linguistiques qui impliquent aussi une vitesse clairement plus rapide (idée d'exagération pour dire « gros », « grand » etc.).

Concernant la valeur moyenne des ordonnées des pentes des courbes de tendance du facteur V_{moy} vs $Dist$, elle était significativement plus petite dans les signes que dans les transitions. Ce qui veut dire que la vitesse moyenne de référence était plus grande dans les transitions que dans les signes, même si cette dernière restait moins variable en fonction de la distance dans les transitions par rapport au signes.

Dans cette étude, le coefficient de détermination R^2 dans l'analyse V_{moy} vs $Dist$, était significativement plus petit dans les transitions que dans les signes. Ceci veut dire que les vitesses des transitions fluctuaient globalement plus en dehors de la droite de régression comparées aux signes. En réalité, dans cette étude, nous avons comparé signe vs transition, sans distinguer les différents types de signes (signe lexical, signe illustratif, pointage, etc.) et de transition (inter-signes et intra-signes). En effet, les différents types de signes et de transitions peuvent avoir des cinématiques de mouvements différentes. Une étude ultérieure sur ces types peut expliquer ce résultat.

Pour conclure cette discussion, nous avons montré que le principe d'isochronie restait valide dans les signes et les transitions, mais qu'il était plus fortement vérifié pour les mouvements de type « signe » quelle que soit le plan du mouvement. La pente de la régression linéaire ainsi que le R^2 étaient plus élevés durant les signes

Finalement, ce résultat sur le principe d'isochronie où la pente des courbes de tendance dans les signes est plus important que dans les transitions renforce notre hypothèse qui dit que la transmission d'informations linguistiques influence la cinématique des mouvements des signes. Ces différences pourraient apparaître pour marquer les différences entre signe et transition afin de rendre le discours compréhensible pour l'interlocuteur et ainsi faire la différence entre le mouvement porteurs d'information linguistique et les autres.

4.2.2 La loi des asymétries verticales et horizontales

Afin d'étudier la loi des asymétries, nous avons analysé chaque direction de mouvements séparément. Pour les mouvements verticaux, nous avons séparé les mouvements vers le haut de ceux vers le bas, et pour les mouvements horizontaux nous avons séparé les mouvements vers l'extérieur de ceux vers l'intérieur.

Le problème que nous avons rencontré est que pour certains sujets et dans les mouvements de transition il reste très peu de mouvements à étudier car les mouvements de transition sont plus grands, (voir section 4.1.2.1.4.1), ce qui a pour effet de rendre l'analyse peu fiable. Afin d'avoir suffisamment de mouvements à étudier, nous avons élargi l'intervalle de distance de mouvement jusqu'à 80 cm.

L'analyse des données a été basée uniquement sur la cinématique du poignet. Les paramètres cinématiques en trois dimensions ont été calculés pour ce marqueur. Le tableau 4-4 montre les nouveaux paramètres des mouvements analysés dans cette étude.

Direction du mouvement	Type de mouvement	Dist	DM	Vmoy	Vmax	DA	DD	Curv
Haut	Signe	0.182	0.272	0.701	1.131	0.145	0.127	0.099
	Transition	0.363	0.399	0.917	1.510	0.189	0.210	0.135
Bas	Signe	0.174	0.257	0.690	1.147	0.148	0.110	0.089
	Transition	0.372	0.460	0.805	1.318	0.246	0.214	0.118
Extérieur	Signe	0.206	0.287	0.727	1.174	0.167	0.120	0.093
	Transition	0.572	0.518	1.054	1.606	0.264	0.254	0.163
Intérieur	Signe	0.226	0.273	0.848	1.267	0.158	0.115	0.138
	Transition	0.413	0.390	0.948	1.406	0.215	0.175	0.130
Moyenne	Signe	0.197 ±0.024	0.272 ±0.012	0.742 ±0.073	1.180 ±0.061	0.155 ±0.010	0.118 ±0.007	0.105 ±0.023
	Transition	0.430 ±0.097	0.442 ±0.060	0.931 ±0.102	1.460 ±0.125	0.229 ±0.033	0.213 ±0.032	0.137 ±0.019

Tableau 4-4: Valeurs moyennes des paramètres de la cinématique du poignet, pour l'étude de la loi des asymétries.

L'effet de la direction du mouvement sur les profils de vitesse a été testé en utilisant le facteur DA/DM entre le temps d'accélération et le temps de déplacement total comme dans (Berret et al., 2009; Gentili et al., 2007; Papaxanthis, Pozzo, & Stapley, 1998). Une valeur supérieure à

0,5 indique une durée relative d'accélération supérieure à la durée relative de décélération et inversement.

L'effet de la vitesse sur les profils de vitesse a été vérifié par le facteur V_{max}/V_{moy} comme dans (Gentili et al., 2007; Papaxanthis, Pozzo, & Stapley, 1998). Si le paramètre V_{max}/V_{moy} est petit cela veut dire que le profil de vitesse a une forme « écrasée », un peu plate, avec des phases d'accélération/décélération moins prononcées.

Les invariances dans les profils de vitesse ont été évaluées qualitativement en les normalisant en temps et en amplitude. Les mouvements des signes et les mouvements des transitions ont été étudiés séparément afin de pouvoir comparer leurs cinématiques.

Les mesures ont été soumises à une analyse de variance (ANOVA) à mesures répétées avec deux facteurs, car elles ont montré une distribution normale (test de Kolmogorov-Smirnov). Les facteurs étaient le type de mouvements (signe et transition) et la direction (haut, bas, extérieur, intérieur). En outre, les asymétries des profils de vitesse haut vs bas et vers l'intérieur vs vers l'extérieur ont été soumises à des tests post-hoc de Bonferroni. Une valeur de $p < 0.05$ indique que la différence de direction sur le paramètre étudié est significative.

4.2.2.1 Résultats

Les valeurs moyennes des facteurs DA/DM et V_{max}/V_{moy} de tous les sujets et dans les deux types de mouvements (signe et transition) et les différentes directions du mouvement sont montrées dans la figure 4-6.

En se focalisant sur le facteur DA/DM , une caractéristique importante est que les temps d'accélération et de décélération étaient inégaux dans les mouvements verticaux et horizontaux, indiquant une asymétrie de la production de mouvement. Si l'analyse ANOVA sur ce facteur a montré un effet non significatif de l'interaction type*direction ($F_{3, 21} = 2.359$, $P = 0.23$), elle a par contre montré un effet significatif du type de mouvement ($F_{1, 7} = 33.901$, $P < 0.01$) et de la direction ($F_{3, 21} = 15.957$, $P < 0.001$).

Concernant le facteur V_{max}/V_{moy} , l'analyse ANOVA a montré un effet significatif de l'interaction type*direction ($F_{3, 21} = 4.721$, $P < 0.05$) et de la direction ($F_{3, 21} = 24.473$, $P < 0.001$). Par contre, elle a montré un effet non significatif du type de mouvement ($F_{1, 7} = 0.537$, $P = 0.23$).

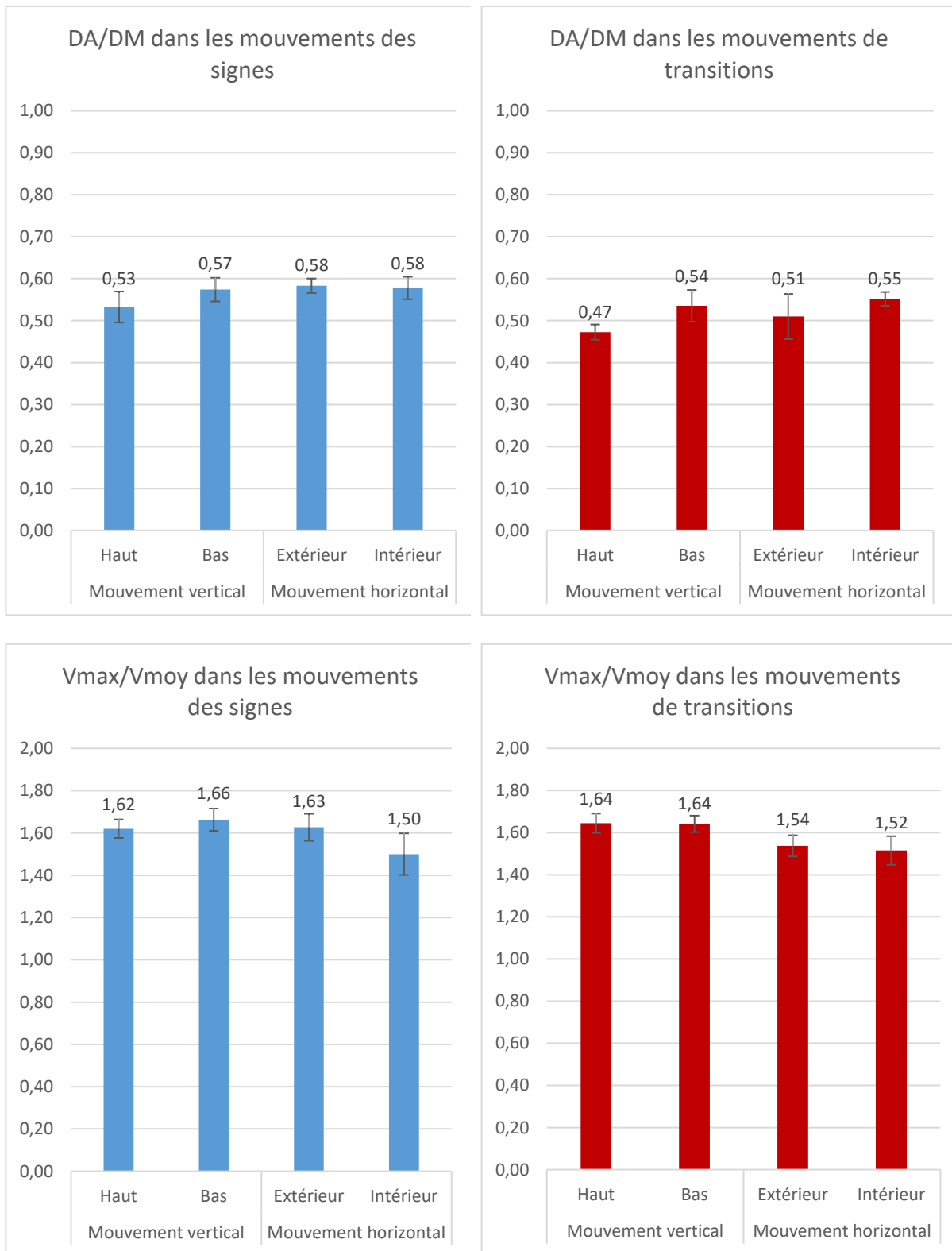


Figure 4-6: Histogrammes montrant les effets de la direction de mouvement sur les facteurs DA/DM et Vmax/Vmoy, en bleu : les mouvements de signes, en rouge : les mouvements de transitions.

Dans la suite, nous présentons les comparaisons post-hoc de Bonferroni sur le type et la direction du mouvement.

4.2.2.1.1 L'effet de la direction du mouvement

Dans cette section, nous vérifions les asymétries des profils de vitesse des mouvements haut vs bas et vers l'intérieur vs vers l'extérieur avec des tests post-hoc de Bonferroni, dans les deux types de mouvements.

a. Mouvements classés « signes »

Le facteur DA/DM varie selon la direction haut vs bas dans les mouvements verticaux. Il est plus élevé dans les mouvements descendants que dans les mouvements ascendants. Les comparaisons post-hoc (de Bonferroni) ont indiqué que le score moyen pour les mouvements ascendants ($M = 0.53$, $SD = 0.04$) était significativement différent ($P < 0.05$) des mouvements descendants ($M = 0.57$, $SD = 0.03$). Par contre, ce facteur reste quasiment identique dans les mouvements horizontaux quelle que soit la direction. Les tests post-hoc ont indiqué que le facteur dans les mouvements vers l'extérieur ($M = 0.58$, $SD = 0.02$) n'a pas été significativement différent ($P = 0.23$) des mouvements vers l'intérieur ($M = 0.58$, $SD = 0.03$).

Concernant le facteur V_{max}/V_{moy} , les comparaisons post-hoc ont indiqué que la valeur moyenne pour les mouvements ascendants ($M = 1.62$, $SD = 0.04$) n'a pas été significativement différente ($P = 0.23$) des mouvements descendants ($M = 1.66$, $SD = 0.05$). Les tests ont indiqué que le facteur dans les mouvements vers l'extérieur ($M = 1.63$, $SD = 0.06$) était significativement différent ($P < 0.05$) des mouvements vers l'intérieur ($M = 1.50$, $SD = 0.10$). Ce qui est intéressant aussi est que la valeur moyenne dans les mouvements vers l'intérieur était significativement différente aussi des mouvements ascendants et descendants.

b. Mouvements classés « transitions »

Le facteur DA/DM varie selon la direction dans les mouvements verticaux comme dans les mouvements des signes, il est plus élevé dans les mouvements descendants que dans les mouvements ascendants. L'analyse post-hoc a indiqué que la valeur moyenne pour les mouvements ascendants ($M = 0.47$, $SD = 0.02$) était significativement différente ($P < 0.05$) des mouvements descendants ($M = 0.54$, $SD = 0.04$). Contrairement aux mouvements des signes horizontaux, le facteur DA/DM varie aussi dans les mouvements de transition horizontaux selon la direction. L'analyse post-hoc n'a pas validé cette différence en indiquant que le facteur dans les mouvements vers l'extérieur ($M = 0.51$, $SD = 0.05$) n'était pas significativement différent ($P = 0.23$) des mouvements vers intérieur ($M = 0.55$, $SD = 0.02$).

Concernant le facteur V_{max}/V_{moy} , ces valeurs étaient identiques dans les mouvements verticaux. Les comparaisons post-hoc ont indiqué que la valeur moyenne pour les mouvements ascendants ($M = 1.64$, $SD = 0.05$) n'a pas été significativement différente ($P = 0.23$) des mouvements descendants ($M = 1.64$, $SD = 0.04$). Cependant, il y'avait une légère variation dans le facteur des mouvements horizontaux, mais les tests post-hoc ont indiqué que la valeur moyenne du facteur dans les mouvements vers l'extérieur ($M = 1.54$, $SD = 0.05$) n'a pas été significativement différente ($P = 0.23$) des mouvements vers l'intérieur ($M = 1.52$, $SD = 0.07$). Ce qui est intéressant aussi dans ce résultat est que le facteur dans les mouvements verticaux était significativement différent des mouvements horizontaux quel que soit la direction.

4.2.2.1.2 L'effet du type de mouvement

Dans cette section, nous vérifions les asymétries des profils de vitesse entre les deux types de mouvements (signe et transition) avec des tests post-hoc de Bonferroni, dans les quatre directions.

Le facteur DA/DM varie selon le type de mouvement (signe vs transition). Il est plus élevé dans les signes que dans les transitions. Les comparaisons post-hoc ont indiqué que le score moyen pour les mouvements ascendants, descendants et vers l'extérieur des signes était significativement différent ($P < 0.001$, $P < 0.05$, et $P < 0.01$ respectivement) de ceux des transitions. Par contre, le score moyen pour les mouvements vers l'intérieur des signes n'était pas significativement différent ($P = 0.23$) de ceux des transitions.

Concernant le facteur V_{max}/V_{moy} , l'ANOVA a montré qu'il n'y avait pas d'effet significatif du type de mouvement sur le facteur ($P = 0.23$ pour les quatre directions).

La figure 4-7 illustre qualitativement les effets de la direction du mouvement sur les profils de vitesse du poignet. Une fois cette vitesse normalisée (en amplitude et en durée), cette figure montre la différence qui existe entre la durée d'accélération et la durée de décélération. Ainsi, les profils de vitesse pour les mouvements verticaux des signes et des transitions diffèrent entre eux selon la direction. Cependant, ils ne diffèrent pas entre eux selon la direction dans les mouvements horizontaux des signes et des transitions. Il faut néanmoins souligner que les asymétries (différences) dans les signes et les transitions ne sont pas identiques.

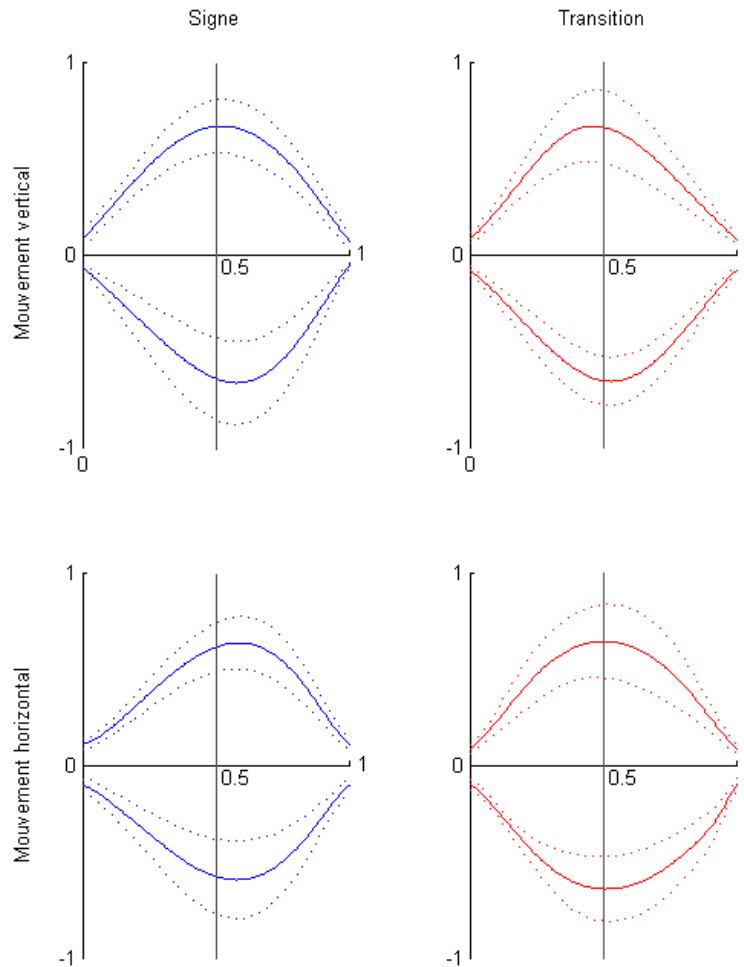


Figure 4-7: Moyennes des profils de vitesse du poignet normalisé pour tous les sujets. Les mouvements ascendants et vers l'extérieur sont représentés par des mouvements positifs, les mouvements descendants et vers l'intérieur par des profils de vitesse négatif. En abscisses, la durée du mouvement normalisée à 1. En ordonnées, la vitesse normalisée.

4.2.2.2 Discussion

Dans la présente étude, nous avons tenté de mieux comprendre la cinématique des mouvements des bras verticaux et horizontaux dans le contexte de la LSF. Un résultat important est que les profils de vitesse des mouvements horizontaux et verticaux des signes et des transitions sont asymétriques (figure 4-7). Ce résultat correspond aux résultats connus sur la loi des asymétries documentée abondamment dans la littérature du contrôle moteur (Gaveau & Papaxanthis, 2011; Gentili et al., 2007; Papaxanthis, Dubost, et al., 2003; Papaxanthis, Pozzo, et al., 2003; Papaxanthis, Pozzo, & Stapley, 1998; Papaxanthis, Pozzo, Vinter, et al., 1998).

En effet, en ce qui concerne nos résultats expérimentaux sur les signes, notamment les valeurs comparables de DA/DM, il est difficile de supposer que la transmission d'informations

linguistiques n'a pas été bien intégrée dans le programme moteur. L'intégration de la gravité dans les modèles internes a été prouvée par des études qui ont exploré l'adaptation motrice en conditions de microgravité (Augurelle, Penta, White, & Thonnard, 2003; Courtine, Papaxanthis, & Pozzo, 2002; Courtine & Pozzo, 2004; Papaxanthis et al., 2005; Papaxanthis, Pozzo, Popov, & McIntyre, 1998; White, McIntyre, Augurelle, & Thonnard, 2005), et l'influence de la gravité sur le système vestibulaire (Angelaki, 2004; Bresciani, Gauthier, Vercher, & Blouin, 2005; Zupan, Merfeld, & Darlot, 2002). En revanche, si la gravité est le seul facteur intégré dans la génération des mouvements des signes, les asymétries A/D auraient eu une allure similaire à ceux des études (Gentili et al., 2007; Papaxanthis, Dubost, et al., 2003; Papaxanthis, Pozzo, et al., 2003; Papaxanthis, Pozzo, & Stapley, 1998; Papaxanthis, Pozzo, Vinter, et al., 1998).

En regardant les valeurs du facteur DA/DM dans les mouvements des signes, ces valeurs sont plus grandes que celles qui sont connues dans les études précédentes de la littérature (mouvements de pointage, d'atteinte de cible, ...etc.). Plus précisément, les valeurs du facteur pour les mouvements ascendants, descendants, vers l'extérieur et vers l'intérieur respectivement étaient : 0.53, 0.57, 0.58, et 0.58. Par contre, dans la littérature, les valeurs étaient autour de 0.46, 0.53, 0.49, et 0.49 respectivement (Gentili et al., 2007; Papaxanthis, Dubost, et al., 2003; Papaxanthis, Pozzo, et al., 2003; Papaxanthis, Pozzo, & Stapley, 1998; Papaxanthis, Pozzo, Vinter, et al., 1998). La question qui se pose ici est : cette différence est-elle liée à la transmission des informations linguistiques dans les signes ?

En ce qui concerne les valeurs du facteur dans les mouvements de transition, ces valeurs sont comparables à celles des études précédentes sur les mouvements de pointage, d'atteinte de cible, etc., où nous avons trouvé des valeurs typiques (0.54, 0.46, 0.51 pour les mouvements ascendants, descendants, vers l'extérieur respectivement). Ce résultat est logique, car les transitions sont des mouvements biologiques comme les mouvements d'atteints de cible qui ne transmettent pas d'informations linguistiques.

En ce qui concerne le facteur V_{max}/V_{moy} , dans la littérature les valeurs de ce facteur étaient différentes de celles de notre étude. Comme exemple, dans les mouvements verticaux, dans une étude de Papaxanthis, dans des mouvements dont la distance était inférieure à 35 cm, la valeur du facteur était autour de 1.95, et pour les mouvements dont la distance était entre 35 et 50 cm, la valeur du facteur était autour de 1.90. Ainsi, dans l'étude de Gentili, la valeur était autour de

1.7. Concernant les mouvements horizontaux, les valeurs du facteur étaient entre 1.8 et 1.9 (T Flash & Hogan, 1985; Gentili et al., 2007).

Par rapport à ces études, les valeurs du facteur dans notre étude quels que soient le type et la direction du mouvement étaient faibles (moins de 1.65). Autrement dit, les profils de vitesse, dans notre étude sont plus plats (ont une forme écrasée) et les phases d'accélération/décélération sont moins prononcées que celles des études antérieures.

Ce dernier résultat sur les profils de vitesse plus plats nous a incité à nous poser les questions suivantes : pourquoi le profil de vitesse des LS est-il plus plat ? Est-ce dû au flux de mouvement plus continu en LS ? Est-ce lié aux postures du bras plus complexe en LS que les mouvements de pointage, geste d'atteinte de cible etc. ? En tout cas, ces valeurs semblent propres à la langue des signes en général, que le mouvement soit du type signe ou transition.

Comme les asymétries ont été modifiées dans les signes et non pas dans les transitions, l'explication logique est que la transmission des informations linguistiques dans les signes influence le principe des asymétries. Un seul résultat qui ne va pas dans ce sens et qui n'était pas compréhensible concerne la valeur du facteur DA/DM dans les mouvements de transition vers l'intérieur. La valeur trouvée était 0.55, ce qui n'est pas équivalent à celle connue en mouvements biologiques des études antérieures (autour de 0.49).

Afin de comprendre ce résultat sur les mouvements horizontaux vers l'intérieur qui nous a paru atypique, nous avons étudié plus en détails ces mouvements. Nous avons regardé les vidéos du corpus plus en détails en essayant de trouver une particularité dans ces mouvements vers l'intérieur. Cette tâche nous a permis de remarquer que ces mouvements se terminent souvent par un contact entre les mains ou avec le corps, ce qui n'est pas le cas pour les autres mouvements. La question qui se pose ici est : est-ce que ce résultat est lié aux contacts de fin du mouvement ? Dans l'intention de répondre à cette question, nous avons effectué une nouvelle analyse en enlevant tous les mouvements qui se terminent par un contact.

4.2.2.3 Analyse des mouvements sans contact

Dans cette nouvelle analyse, nous avons éliminé tous les mouvements qui se terminent par un contact entre les deux mains, quels que soient leurs directions et leur type. Nous avons analysé la cinématique des mouvements restants de la même manière, c.à.d., nous avons calculé les facteurs DA/DM et V_{max}/V_{moy} de toutes les directions des mouvements des signes et des

transitions. Les mesures ont été soumises à une analyse ANOVA à mesures répétées avec deux facteurs.

Pour enlever ces mouvements, nous avons appliqué un filtre qui élimine tous les mouvements dont la distance entre les deux mains à la fin du mouvement (les marqueurs placés sur les poignets) était inférieure à 12 cm. La contrainte de ce filtre est qu'il élimine aussi les mouvements qui se terminent avec les mains proches mais sans contact. Mais cette contrainte n'est pas gênante car il restait suffisamment de mouvements à étudier pour chaque sujet dans chaque direction.

4.2.2.3.1 Résultats

La figure 4-8 montre que les résultats des mouvements sans contact étaient quasiment les mêmes que ceux qui contiennent des contacts, quel que soit le type du mouvement (signe en bleu, et transition en rouge). Le seul résultat différent est celui des mouvements des transitions vers l'intérieur où les valeurs de DA/DM étaient 0.52. En effet, l'ANOVA a montré les mêmes effets que dans les mouvements avec contacts, c'est-à-dire un effet non significatif de l'interaction type*direction ($F_{3, 21} = 1.075$, $P = 0.23$). Par contre, elle a montré un effet significatif du type de mouvement ($F_{1, 7} = 74.939$, $P < 0.001$), et de la direction ($F_{3, 21} = 12.527$, $P < 0.001$).

Concernant le facteur V_{max}/V_{moy} , comme indiqué dans la figure 4-8, la seule valeur différente des mouvements avec contacts concerne celle des mouvements de transition vers l'intérieur. L'analyse ANOVA a montré les mêmes effets que dans les mouvements avec contacts, c'est-à-dire un effet significatif de l'interaction type*direction ($F_{3, 21} = 3.372$, $P < 0.05$) et de la direction ($F_{3, 21} = 28.078$, $P < 0.001$). Par contre, elle a montré un effet non significatif du type de mouvement ($F_{1, 7} = 0.814$, $P = 0.23$).

L'effet de la direction sur le facteur DA/DM a été analysé avec les comparaisons post-hoc qui ont donné le même résultat de significativité que dans les mouvements avec contacts. C'est-à-dire les mouvements ascendants étaient significativement différents ($P < 0.05$) des mouvements descendants dans les deux types de mouvements, et il n'y avait pas de différence significative entre les mouvements vers l'extérieur et vers l'intérieur ($P = 0.23$) dans les deux types de mouvement.

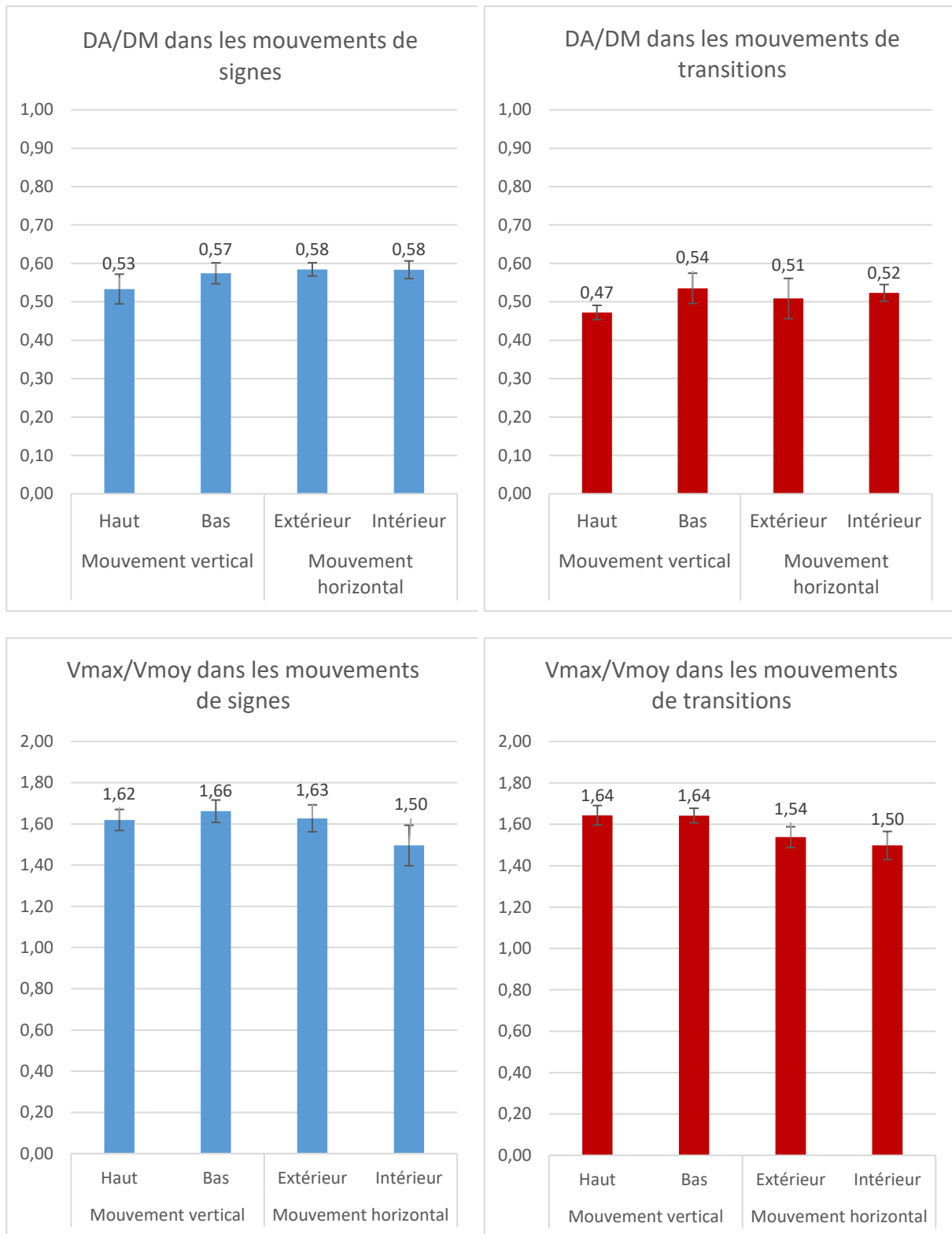


Figure 4-8 : Histogrammes montrant les effets de la direction de mouvement sur le facteur DA/DM dans les mouvements sans contact, en bleu : les mouvements de signes, en rouge : les mouvements de transitions.

L'effet du type de mouvement sur le facteur DA/DM a été aussi analysé avec les comparaisons post-hoc qui ont indiqué le même résultat de significativité pour les mouvements ascendants,

descendants et vers l'extérieur ($P < 0.01$, $P < 0.05$, et $P < 0.01$). Le changement de la valeur de DA/DM dans les mouvements vers l'intérieur a rendu l'effet du type de mouvement significatif, où la valeur est devenue significativement différente entre les signes et les transitions ($P < 0.001$) contrairement aux mouvements avec contacts.

Concernant le facteur V_{max}/V_{moy} , l'ANOVA a montré qu'il n'y avait pas d'effet significatif du type de mouvement sur le facteur ($P = 0.23$).

4.2.2.3.2 Discussion

En enlevant les mouvements qui se terminent avec un contact, les seules valeurs des facteurs DA/DM et V_{max}/V_{moy} qui ont changé étaient celles des mouvements de transition vers l'intérieur, où elles chutent de 0.55 à 0.52 et de 1.52 à 1.50 respectivement. Ainsi, les comparaisons post hoc sur les mouvements vers l'intérieur ont montré que les valeurs du facteur étaient significativement différentes entre les signes et les transitions, ce qui n'était pas le cas dans les mouvements avec contacts.

Ce résultat confirme l'hypothèse que nous avons posée sur les mouvements qui se terminent par un contact. La présence de contacts est une différence potentiellement importante avec les études de contrôle moteur classique qui s'intéressent à des pointages en direction de cibles mais sans besoin de les toucher. De cela, nous pouvons dire qu'il n'y a pas de différence dans le facteur DA/DM entre les mouvements horizontaux. Ce résultat montre aussi que la cinématique des mouvements horizontaux sans contact ne change pas selon la direction, ce qui est compatible avec la littérature. Cette subtilité d'analyse montre la difficulté et les précautions à prendre dans l'étude des mouvements de LSF.

Pour conclure cette discussion, nous pouvons dire que les asymétries haut/bas dans les profils de vitesses restent valides durant les mouvements des signes et des transitions, mais s'en trouvent modifiés dans les signes. Ainsi, dans les mouvements des signes, le temps relatif d'accélération est systématiquement plus grand que celui connu dans la littérature. Ces résultats renforcent notre hypothèse qui dit que la transmission d'informations linguistiques influence la cinématique des mouvements des signes. Nous avons trouvé que les signes tendent à se terminer de manière plus abrupte que les transitions ou les mouvements issus des expériences de contrôle moteur classiques. Finalement, les profils de vitesse dans les mouvements de la LSF sont plus plats (ont une forme écrasée) et les phases d'accélération/décélération sont moins prononcées que celles des études antérieures pour des raisons qui restent à élucider mais pourraient être liées à la continuité des mouvements en LSF.

4.2.3 La loi de puissance 2/3

La loi de puissance 2/3, une loi empirique indiquant une relation non linéaire inverse entre la vitesse tangentielle de la main et la courbure de sa trajectoire pendant le mouvement incurvé, est largement connue dans la littérature. Il a également été démontré qu'elle existait dans le mouvement des yeux, la locomotion et a même été démontré dans la perception et la prédiction du mouvement. Ici nous voulons la vérifier dans les mouvements des bras dans le contexte des signes, pour cela nous avons analysé les mouvements complets des signes et des transitions quelles que soient leurs directions.

Rappel : la loi de puissance 2/3 peut être formulée comme suit:

$$v(t) = \text{const} * k(t)^\beta \quad (1)$$

Ou, dans l'espace logarithmique :

$$\log(v(t)) = \text{const} + \beta \log(k(t)) \quad (2)$$

β est environ $-\frac{1}{3}$, l'équation (1) devient :

$$v(t) = \alpha * k(t)^{-\frac{1}{3}} \quad (3)$$

La validité de la loi a été testée séparément pour les mouvements des signes et des transitions. Au niveau qualitatif, cela a été fait en traçant la vitesse tangentielle $v(t)$ par rapport à la courbure $k(t)$. Selon l'équation 3, si le facteur de gain de vitesse α est constant par morceaux, les points de données se regrouperont le long d'un segment qui a une pente (l'exposant β) autour de $-1/3$.

Les mesures ont été soumises à un test student (t-test) entre le type de mouvements (signes et transitions). Une valeur de $P < 0.05$ indique que le résultat est significatif.

4.2.3.1 Résultats

Une analyse quantitative de la loi de puissance des deux tiers a également été réalisée en calculant pour chaque essai la régression linéaire de $\log(v)$ contre $\log(k)$. Ainsi que, des analyses séparées ont été faites pour chaque sujet. La figure 4-9 illustre la vitesse tangentielle $v(t)$ par rapport à la courbure $k(t)$ pour tous les mouvements des signes et des transitions de tous les sujets. La pente des lignes de régression estime l'exposant β de l'équation 1. L'intersection estime le logarithme du gain de vitesse k . Le tableau 4-5 résume les résultats.

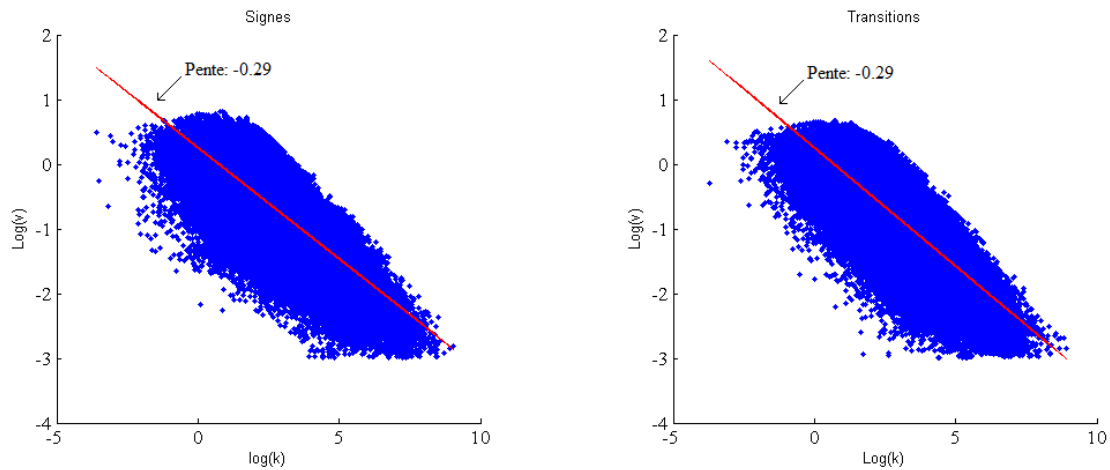


Figure 4-9 : Tracé du $\log(v)$ vs $\log(k)$ et la pente de régression pour tous les mouvements de tous les sujets, à gauche pour les mouvements des signes et à droite pour les mouvements des transitions.

Un t-test a été effectué pour comparer la pente des lignes de régression (l'exposant β), et le coefficient de détermination R^2 entre $\log(v)$ et $\log(k)$ entre les signes et les transitions. Le test sur le R^2 a révélé qu'il y avait une différence significative entre les signes ($M = 0.73$, $SD = 0.03$) et les transitions ($M = 0.69$, $SD = 0.03$); ($T_7 = 3.45$, $P = 0.01$). Par contre, sur l'exposant β le test a révélé un effet non significatif entre les signes ($M = -0.29$, $SD = 0.06$) et les transitions ($M = -0.29$, $SD = 0.06$); ($T_7 = 0.46$, $P = 0.23$). Ces résultats montrent que la pente des lignes de régression (l'exposant β) était similaire dans les deux types de mouvements, par contre la vitesse tangentielle $v(t)$ par rapport à la courbure $k(t)$ était plus corrélée dans les mouvements des signes.

Sujets	Les mouvements des signes		Les mouvements de transitions	
	L'exposant β	R^2	L'exposant β	R^2
S1	-0.33 ± 0.09	0.72 ± 0.18	-0.34 ± 0.08	0.67 ± 0.22
S2	-0.41 ± 0.09	0.73 ± 0.17	-0.39 ± 0.07	0.66 ± 0.21
S3	-0.27 ± 0.08	0.77 ± 0.19	-0.27 ± 0.09	0.70 ± 0.22
S4	-0.27 ± 0.11	0.69 ± 0.23	-0.27 ± 0.10	0.64 ± 0.24
S5	-0.24 ± 0.08	0.75 ± 0.21	-0.23 ± 0.09	0.70 ± 0.23
S6	-0.26 ± 0.09	0.74 ± 0.19	-0.28 ± 0.09	0.73 ± 0.19
S7	-0.28 ± 0.09	0.71 ± 0.23	-0.32 ± 0.08	0.67 ± 0.23
S8	-0.24 ± 0.09	0.69 ± 0.23	-0.23 ± 0.08	0.71 ± 0.21
Moy	-0.29 ± 0.06	0.73 ± 0.03	-0.29 ± 0.06	0.69 ± 0.03

Tableau 4-5: Analyse des données individuelles selon le modèle de loi de puissance des deux tiers

4.2.3.2 Discussion

Dans la présente étude, nous avons tenté de vérifier la validation de la loi de puissance $2/3$ dans les mouvements des signes et des transitions.

Un résultat important de cette étude est que la loi de puissance deux tiers reste valide durant les mouvements des signes et des transitions. Ce résultat correspond aux résultats des études antérieures (De'Sperati & Viviani, 1997; Kandel et al., 2000; Levit-Binnun et al., 2006; Perrier & Fuchs, 2008; Soechting et al., 1986; Tasko & Westbury, 2004; Paolo Viviani & Flash, 1995; P Viviani & Stucchi, 1992). La valeur de l'exposant β (la pente des lignes de régression) était la même (-0.29) dans les deux types de mouvements, et elle était équivalente à la valeur connue dans les études antérieures. On observe cependant une petite différence dans la corrélation entre $\log(v)$ et $\log(k)$ entre les mouvements des signes et des transitions.

Ce résultat permet de dire que la loi de puissance $2/3$ reste valide durant les mouvements de signes et des transitions avec des valeurs de l'exposant β et le coefficient de détermination R^2 assez proche de ce qui est connu dans la littérature.

4.3 Conclusion

Ce chapitre a été consacré à la présentation de la méthode de traitement de données suivi pour étudier les LCCM durant la LSF, ainsi que les résultats de ces études. Nous avons détaillé dans la première section la méthode utilisée pour calculer les paramètres de la cinématique du mouvement dans le logiciel Matlab, et dans la deuxième section, nous avons présenté les résultats des études des trois lois du contrôle moteur ainsi que leurs discussions.

Dans la première section, nous avons montré que l'étude des mouvements de la LSF est plus compliquée que l'étude d'autres mouvements (pointage, geste d'atteinte de cible, dessin de formes etc.), car les mouvements de LSF nécessitent l'annotation, la segmentation, ...etc. Nous avons présenté les deux méthodes de segmentation de mouvements que nous avons suivies. Finalement, nous avons présenté les filtres que nous avons appliqués sur ces mouvements afin de les comparer avec les études antérieures, et les paramètres cinématiques du mouvement que nous avons calculé pour étudier les LCCM.

Nous avons vu que les trois lois restent valides dans les signes et les transitions. Le résultat le plus important est que la transmission des informations linguistiques durant les mouvements de signes a modifié la loi des asymétries et le principe d'isochronie : dans les asymétries, les temps d'accélération étaient plus grands que ce qui est connu dans la littérature, et le principe

d'isochronie était plus fort dans les signes en le comparant avec la littérature. Cependant, dans les mouvements des transitions, ces deux lois n'ont pas changé par rapport aux études antérieures (pointage, geste d'atteinte de cible, ...etc.). D'autre part, la loi de puissance $2/3$ n'a pas été modifiée par le fait de la transmission des informations linguistiques durant les signes.

Ces résultats pourraient être utilisés pour bien modéliser les mouvements de LSF, de façon à avoir des modèles de génération de mouvements biologiques. Dans les mouvements de transitions les lois peuvent être utilisées directement, mais dans les mouvements de signes une adaptation de ces lois est nécessaire avant de les utiliser.

5 Perception du mouvement sur un signeur virtuel en LSF

Nous avons montré dans le chapitre précédent que les LCCM restent valides dans les mouvements de la LSF, mais la transmission d'informations dans les mouvements de signes modifie le principe d'isochronie et des asymétries. Avant de modéliser les mouvements de la LSF, il est important de valider cette hypothèse, en vérifiant la quantité d'informations qu'ils transmettent.

L'objectif de ce chapitre est de comprendre comment des individus sourds perçoivent la langue des signes en fonction de la quantité d'informations qui leur est disponible. Pour cela, nous avons mesuré la compréhension d'individus sourds et signeurs et d'individus non-sourds, non-signeurs faces à différents types de vidéos. Celles-ci présentaient soit une personne décrivant une photo en LSF, soit un avatar simplifié (sans doigts et sans expressions faciales) décrivant en LSF cette même photo.

La première section présente la méthodologie que nous avons suivie. Dans cette section, nous présentons les informations des participants à cette étude, le contenu que nous avons utilisé, et la procédure que nous avons suivie pour faire cette étude. Dans la deuxième section nous présentons les résultats obtenus et proposons des explications.

5.1 Méthodologie

Dans cette section, nous présentons les informations concernant les participants (section 5.1.1), le contenu (section 5.1.2) et la procédure suivie (section 5.1.3).

5.1.1 Participants

Les participants ont été recrutés par le biais d'internet où un lien vers un test en ligne était diffusé. La diffusion du test a été faite au LSF par un collègue sourd, pour toucher un public locuteur de la LSF. Le fait de posséder une connexion internet était une condition importante pour réaliser le questionnaire puisque celui-ci n'était disponible qu'en ligne. Les sessions où le test n'a pas été complété entièrement étaient exclues. Nous avons retenu 78 participants, tous de nationalité française, répartis en 3 groupes selon leur statut auditif (sourde ou entendant) et leur connaissance de la LSF²³ (signeur ou non signeur). L'ensemble des sujets a passé le même test en ligne.

²³ Aucune personne sourde et non signeuse n'a participé au test.

Le premier groupe était constitué de 22 Entendants Non Signeurs (ENS) : 9 hommes et 13 femmes (Moy-âge = 33.1, SD = 12.6 ans). Aucun n'avait de parents ou de proches utilisant la LSF et aucun n'avait suivi un apprentissage en LSF. Le deuxième groupe était constitué de 23 Entendants Signeurs (ES) : 5 hommes et 18 femmes (Moy-âge = 30.2, SD = 9.8 ans). La moyenne d'âge dont ils ont commencé à apprendre la LSF dans ce groupe était (Moy-âge = 20.6, SD = 7.4 ans). En moyenne les individus appartenant à ce groupe avaient une expérience en LSF de (M = 9.6 ans). Le troisième groupe était constitué de 33 Sourds Signeurs (SS) : 22 hommes et 11 femmes (Moy-âge = 39.4, SD = 9.7 ans). La moyenne d'âge dont ils ont commencé à apprendre la LSF dans ce groupe était (Moy-âge = 8.7, SD = 10.6 ans). En moyenne les individus appartenant à ce groupe avaient une expérience en LSF de (M = 30.7 ans).

5.1.2 Contenu

La même tâche du corpus que précédemment, qui consiste à la description de 25 photos (figure 5-1) a été utilisée dans cette étude. Nous avons choisi 8 photos parmi les 25. Ces descriptions ont été utilisées dans deux versions, où nous avons pris la vidéo d'une personne réelle (un signeur) et un avatar conçu à partir du squelette utilisé dans l'annotation. Dans le but de mesurer le taux d'informations transmises par les mouvements des bras et le reste du corps, nous avons utilisé un avatar sans les doigts et les expressions faciales.



Figure 5-1: Exemples de 4 scènes, décrites en Langue des Signes dans les vidéos Normales et Avatars. En haut à gauche : ville. En haut à droite : maison. En bas à gauche : magasin de vin. En bas à droite : dressing.

Nous avons obtenu 16 vidéos variant selon deux conditions : 8 vidéos où la description des scènes était réalisée par une personne réelle (condition Normale : CN) et 8 vidéos où un avatar (condition Avatar : CA) réalisait strictement les mêmes descriptions (figure 5-2). La dimension de chacune des vidéos était : 600 x 400 pixels. Les photos choisies appartenaient à 4 catégories différentes : ville, maison en construction, magasin de vin et dressing (figure 5-1). Pour chacune des catégories, 2 scènes différentes étaient présentées dans le test.

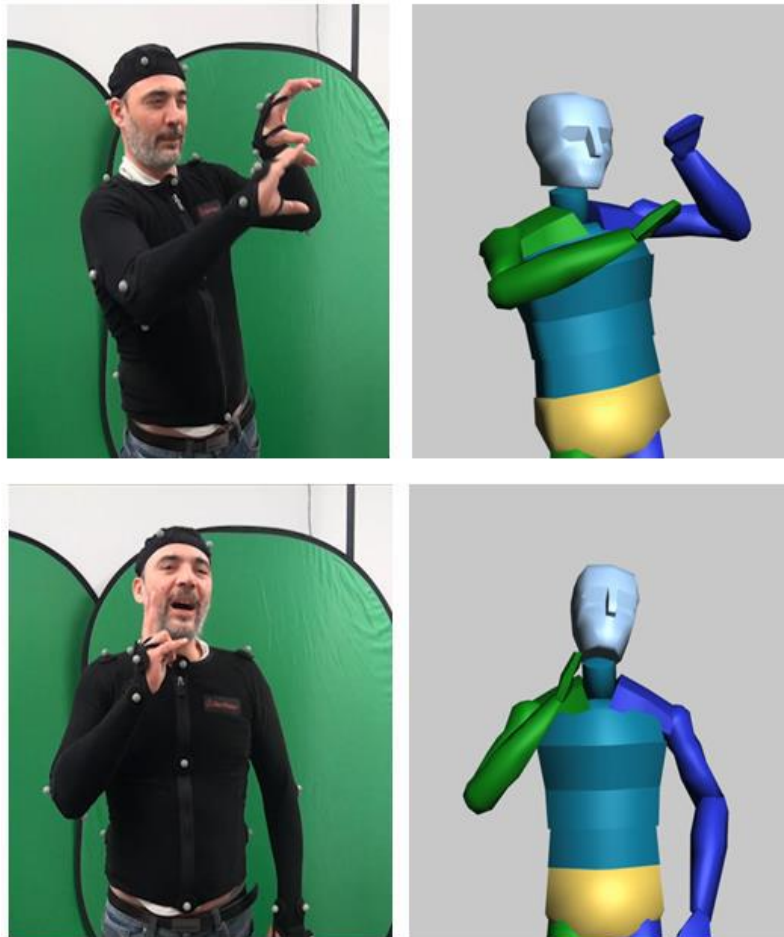


Figure 5-2: A gauche : extrait d'une vidéo Normale ; A droite : extrait d'une vidéo Avatar.

Un ensemble de 15 photos (figure 5-3) a été créé pour chaque vidéo (CN et CA) ce qui donne 16 ensembles. Ces ensembles ont été utilisés comme planche de réponse. Un programme conçu à l'aide du logiciel Matlab™ a permis de créer des ensembles de photos aléatoires à partir de photos libres de droits collectées sur internet²⁴. Quatre catégories de distracteurs ont été ajoutées (Mer, Bibliothèque, Forêt et Bois coupé) aux 4 catégories de base. Chaque catégorie

²⁴ <https://pixabay.com/fr/>

contenait 25 photos parmi lesquelles le programme sélectionnait aléatoirement celles utilisées pour créer la planche de réponse.

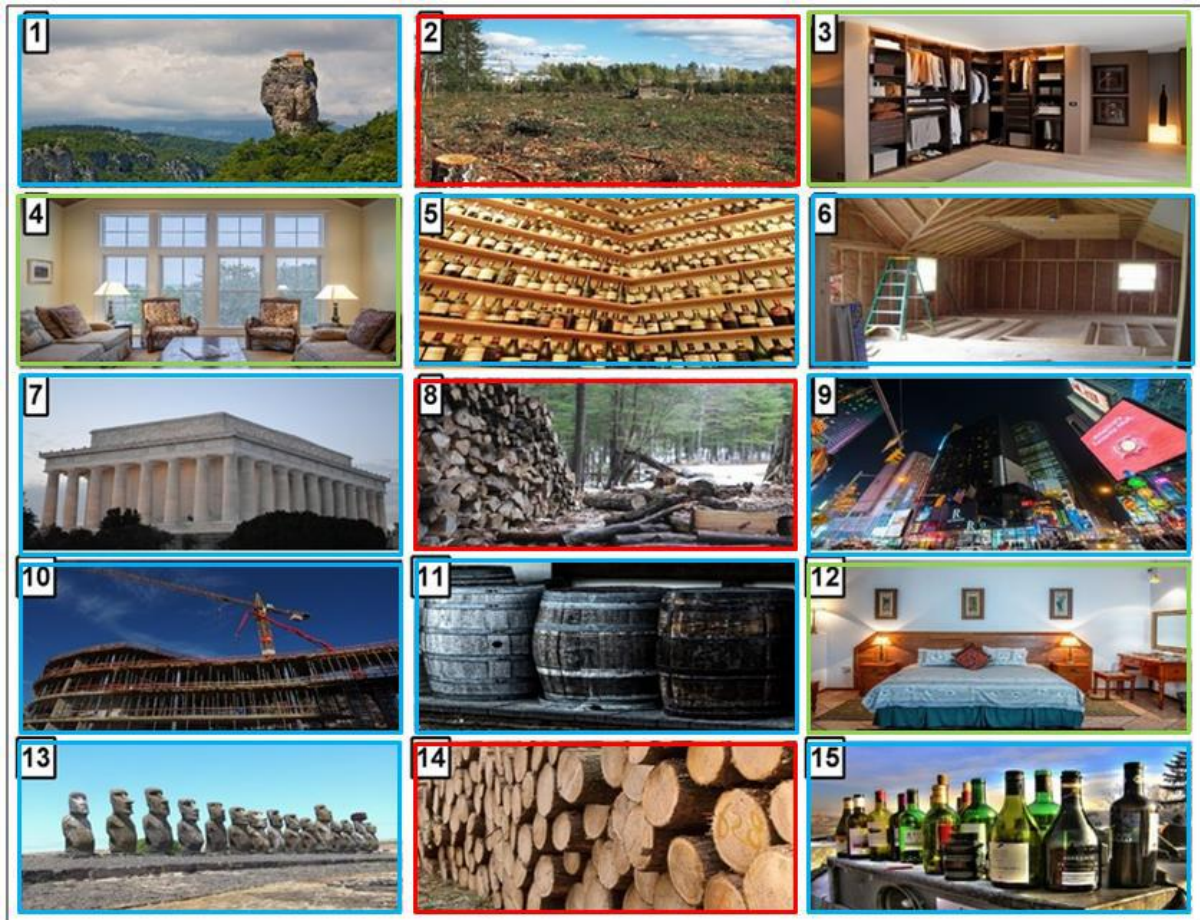


Figure 5-3: Exemple d'un ensemble de 15 photos présenté aux participants de l'étude. 3 photos en vert : Catégorie de la photo à retrouver, ici la photo 3 de la catégorie « pièce ». ; 9 photos en bleu : Autres catégories utilisées dans les autres essais (3 « Monument », 3 « Maison », 3 « Vin ») ; 3 photos en rouge : Catégorie « distractive », ici « Bois coupé ».

Ce programme permettait, dans un premier temps, de disposer aléatoirement 15 photos dans une même planche de réponse. Il sélectionnait 3 photos aléatoires à partir de 5 catégories de photos différentes choisies par l'expérimentateur, incluant la bonne photo qui correspond à la vidéo de description. L'emplacement de la bonne photo a été choisi aléatoirement pour chaque planche. Chacune des 16 planches était associée à sa vidéo.

Ainsi le questionnaire était composé de 16 vidéos auxquelles étaient associées 16 ensembles de photos (ou planche de réponse) identiques pour tous les participants.

5.1.3 Procédure

Le logiciel LimeSurvey™²⁵ a été utilisé pour réaliser le test (Schmitz, 2012). Le questionnaire conçu spécialement pour cette étude contenait les 16 vidéos descriptives (8 vidéos Avatars et 8 vidéos Normales) ainsi que les 16 planches de réponse. Afin de permettre une lecture des vidéos sur différents navigateurs internet (ex. Google Chrome, Mozilla Firefox et Internet Explorer), celles-ci ont été encodées sous différents formats (mp4, Webem, Ogv) et un code HTML spécifique était intégré dans le test. Ce test était disponible en ligne sur une période de trois semaines.

En accédant au questionnaire, des consignes étaient dispensées en français et en anglais écrit. Les consignes étaient rédigées de manière à être comprises par l'ensemble des participants, y compris par des personnes sourdes ayant un niveau de français écrit moyen.

Les participants remplissaient ensuite un ensemble de données les concernant (âge, nationalité, sexe). Selon que le participant rapportait pratiquer ou non la LSF, l'arborescence de questions n'était pas la même. Des questions supplémentaires sur l'âge d'apprentissage de la LSF, le lieu d'apprentissage de la LSF étaient proposées pour les individus ayant indiqué pratiquer la LSF.

Nous avons inclus dans le test deux essais de familiarisation, dans les deux conditions CA et CN, permettant aux participants de se familiariser rapidement à la tâche. Le questionnaire était constitué de 2 blocs (Avatar et Normal) présentés l'un à la suite de l'autre. Les 8 vidéos au sein de chaque bloc étaient présentées dans un ordre aléatoire selon les sujets.

Afin que les participants ne regardent la vidéo qu'une seule fois, puis passent à la page suivante, nous avons limité la répétition du visionnage. Les commandes de contrôle des vidéos (jouer, arrêter, barre d'avancement) étaient supprimées et le retour en arrière était bloqué volontairement.

Chaque test se déroulait pour les deux conditions CA et CN de la manière suivante (figure 5-4) : suite à un compte à rebours de 4 secondes, une vidéo démarrait automatiquement. Une fois la vidéo terminée, le participant devait cliquer sur le bouton « suivant ». La page suivante affichait la planche de la réponse associée à la vidéo. Le participant devait alors indiquer, à l'aide du clavier, le numéro de la photo dont il pensait avoir reconnue la description. Une fois la réponse donnée, une nouvelle vidéo apparaissait. Une fois le test terminé, les réponses de

²⁵ <http://www.limesurvey.org>

chaque participant étaient compilées au sein du logiciel LimeSurvey™ en ligne puis exportées dans un tableau Excel pour permettre leur analyse.

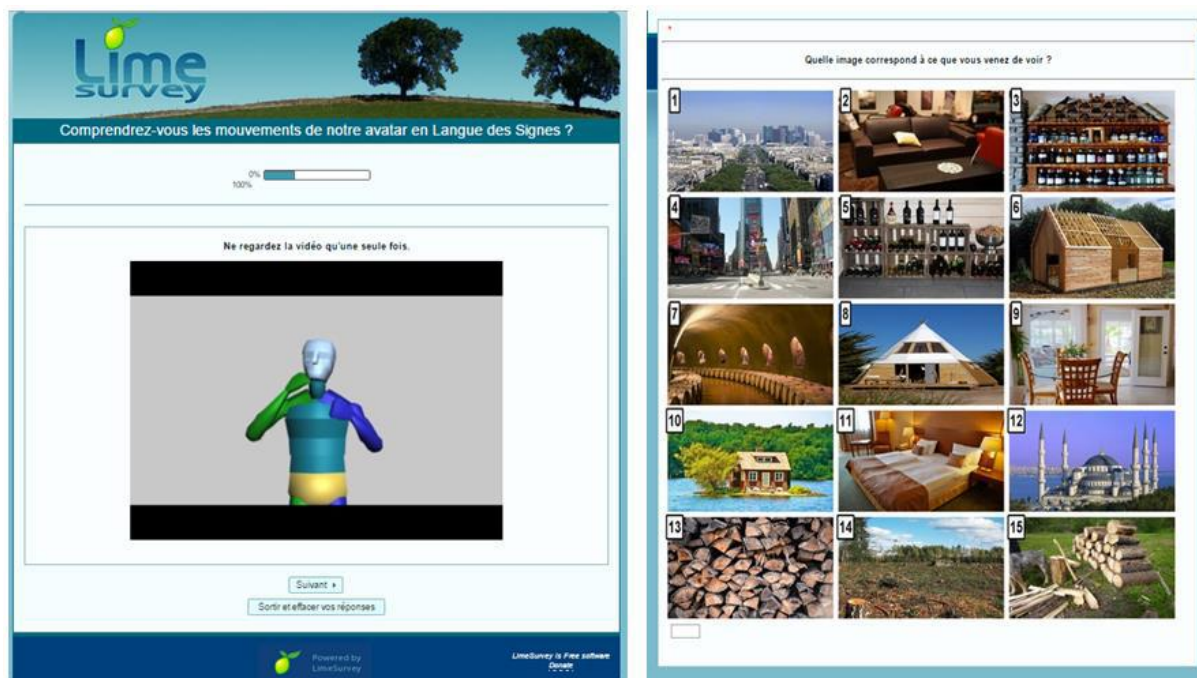


Figure 5-4: Illustration du déroulement d'un essai. A gauche : page sur laquelle était présentée la vidéo de description en LSF, ici en condition Avatar. L'onglet « suivant » en bas de la page permettait de passer à la page de réponse. A droite : L'ensemble de 15 photos contenant la bonne réponse, ici la photo 1. Le participant devait entrer sa réponse, à l'aide du clavier, dans le cadre vide en bas à gauche.

5.2 Résultats

La mesure principale de cette étude concerne la compréhension des vidéos par participants, dans les deux conditions CN et CA. Ainsi, la mesure de la performance a été évalué selon deux critères : le choix de la bonne photo, et le choix de la bonne catégorie mais pas de la bonne photo (c.à.d. la catégorie est bonne mais la photo est fausse).

Les mesures ont été soumises à une analyse de variance (ANOVA) à mesures répétées pour lesquelles le seuil de significativité était $\alpha = 0,05$. Nous avons également mené des post-hoc HSD de Tuckey.

Notre variable dépendante était la performance des participants, c'est-à-dire, le nombre de bonnes réponses obtenues à la fin du test. Les variables indépendantes étaient en facteur inter-sujet du groupe dans lequel étaient répartis les participants, c'est-à-dire Entendants Non-Signeurs, Entendants Signeurs, Sourds-signeurs (i.e., ENS, ES, SS) et en facteur intra-sujet les

deux types de vidéos (i.e., Avatars et Normales), 4 catégories et 2 Scènes décrites (par catégorie).

5.2.1 Effet du type de vidéo sur la performance

La moyenne du pourcentage de réussite dans la CN était plus grande que celle dans la CA avec une différence de 24%, comme montré dans la figure 5-5. L'analyse ANOVA a montré un effet significatif du type de vidéo sur la performance des participants au test ($F_{1, 75} = 134.37$, $p < 0,001$). Ainsi, les participants des 3 groupes ont de meilleures performances pour les vidéos en CN comparées aux vidéos en CA.

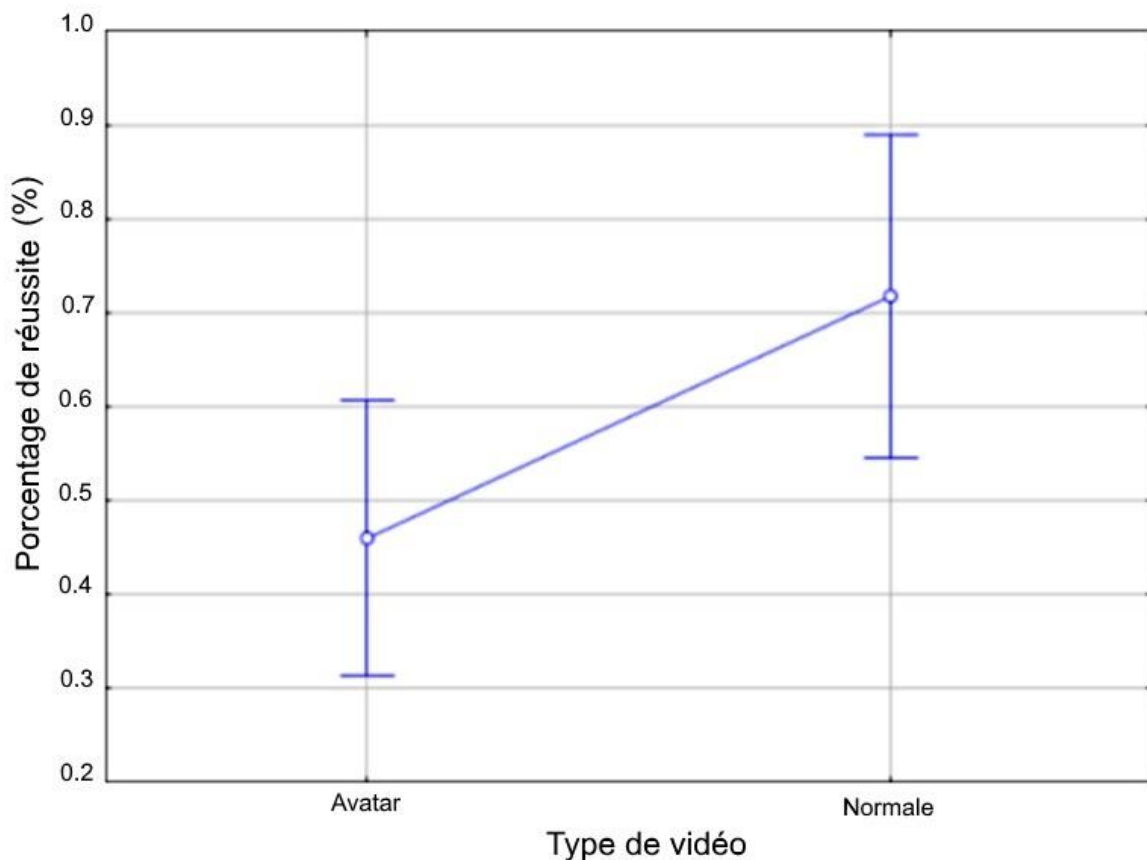


Figure 5-5: Effet du type de vidéo (Avatar ou Normale) sur la performance de l'ensemble des participants.

5.2.2 Effet principal du Groupe sur la performance

L'analyse ANOVA a montré un effet principal du groupe sur la performance des participants au test ($F_{2, 75} = 22,03$, $p < 0,0001$). Ainsi que, l'analyse post-hoc (HSD de Tuckey) a mis en évidence que les individus du groupe SS et ES ont de meilleurs performances comparées au groupe ENS. Cependant, il n'existe pas de différence significative entre les performances des groupes SS et ES (figure 5-6).

5.2.3 Interaction Groupe*Type de vidéo sur la performance

L'analyse ANOVA a mis en évidence l'existence d'une interaction significative Groupe*Type de vidéo sur la performance des participants ($F_{2,75} = 10,87, p < 0,0001$) (figure 5-7). L'analyse post-hoc a montré que pour chacun des trois groupes, les participants ont une performance plus importante pour les vidéos Normales comparées à celles Avatar (figure 5-7).

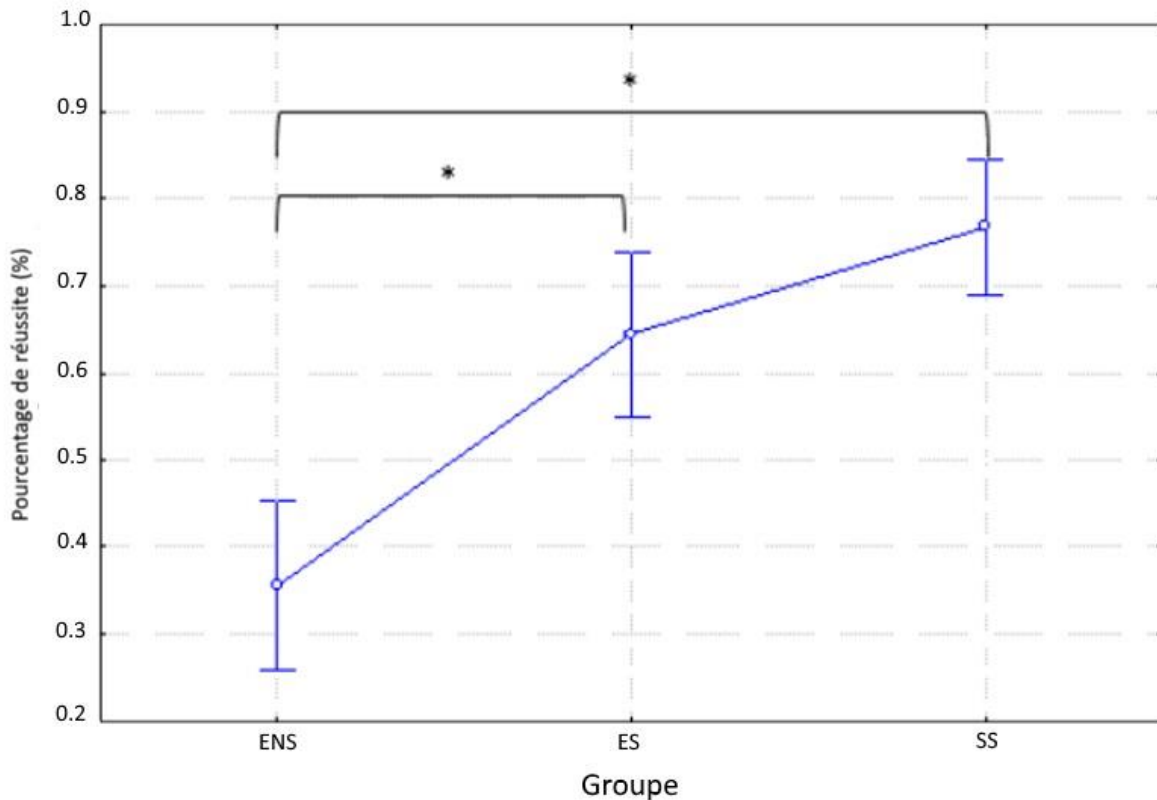


Figure 5-6: Performance globale au test en fonction du groupe en abscisse. Illustration de l'effet principal du Groupe sur la performance des participants

Cependant, l'ordre de présentation des vidéos en CA et CN au cours de l'étude est resté le même pour l'ensemble des participants : Avatars en premier puis Normales. Ainsi, nous pouvons difficilement déterminer si la différence de performance entre les deux types de vidéos est due au type de vidéo ou à l'ordre de présentation de ces différentes vidéos. Lors d'une étude ultérieure, il conviendrait de varier l'ordre de passage des vidéos Avatars et Normales pour tenter de déterminer précisément si l'ordre de présentation ou le type de vidéo ont une influence sur la performance des participants.

De ce fait, pour la suite des analyses nous avons affiné l'analyse sur la CA (Courbe bleue, figure 5-7) qui nous intéresse plus particulièrement pour cette étude.

5.2.4 Effet de la catégorie et des Scènes décrites sur la performance

L'analyse ANOVA a montré qu'il n'existe pas d'effet significatif des Scènes décrites sur la performance des participants ($F_{1, 75} = 3,30$).

Il existe cependant une interaction significative Catégorie*Scène décrite sur la performance ($F_{3, 225} = 3,12$). Une analyse post-hoc (HSD de Tuckey) met en évidence des différences significatives de la performance entre les 2 vidéos d'une même catégorie. Ainsi, nous avons retiré le facteur Catégorie des analyses suivantes.

Nous n'avons donc pas exploité la performance des individus au regard du choix de la bonne catégorie.

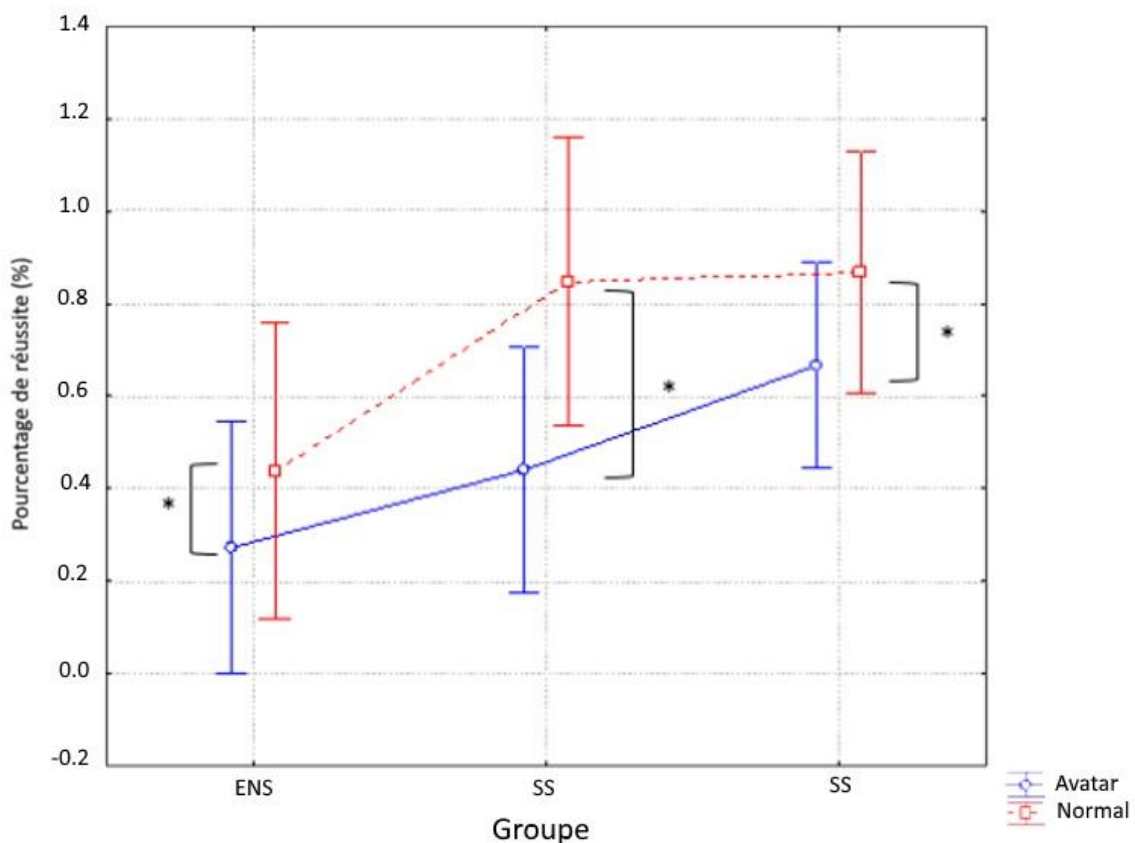


Figure 5-7: Effet d'interaction Groupe*Type de vidéo sur la performance des participants des groupes Entendants non Signeurs (ENS), Entendants Signeurs (ES), Sourds Signeurs (SS).

5.2.5 Effet principal du groupe sur la performance en Condition Avatar

L'ANOVA sur le pourcentage de bonnes réponses des participants des 3 groupes en CA a été menée (figure 5-7, Courbe Bleue) avec en facteur intra-sujet : 8 photos décrites (A, B, C, D, E, F, G, H) et inter-sujet : le Groupe (ENS, ES, SS). Nous avons observé un effet principal du

Groupe sur la performance des sujets lors de la présentation de vidéos de type Avatar ($F_{2,75} = 17,00, P < 0,0001$).

Suite à une analyse post-hoc, on observe une meilleure performance pour les individus Sourds Signeurs (SS) comparés aux deux autres groupes (ES et ENS). Les individus Entendants Signeurs (ES) ont une performance plus élevée comparée aux individus Entendants Non Signeurs (ENS).

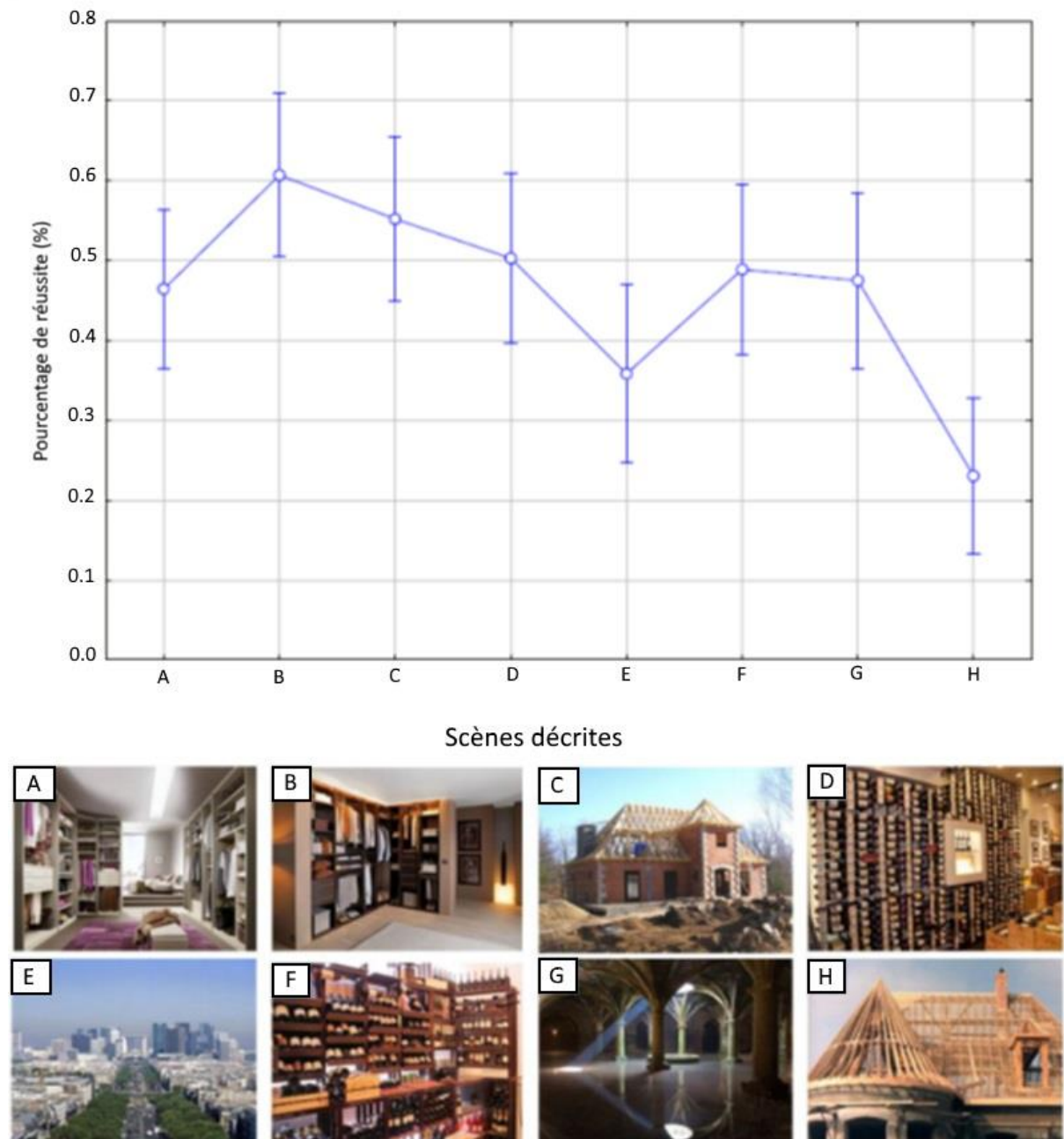


Figure 5-8: Performance des participants en fonction de la Scène décrite, seulement en condition Avatar.

5.2.6 Effet de la vidéo sur la performance en CA pour l'ensemble des participants

L'analyse ANOVA a montré qu'il n'existe pas un effet de la vidéo sur la performance pour le groupe ENS ($F_{7, 147} = 0,28$). Par contre, pour les deux autres groupes, Il existe un effet significatif, pour les groupes ES ($F_{7, 154} = 3,13$) et SS ($F_{7, 224} = 8,34$) (figure 5-8).

5.2.7 Effet d'interaction Vidéo*Groupe en condition Avatar

L'ANOVA menée sur la performance des participants en CA montre une interaction significative de la Scène décrite et du Groupe sur la performance des participants ($F_{14, 525} = 2,06$, $p = 0,013$) (figure 5-9). D'autres résultats obtenus par une analyse post-hoc montrent que pour le groupe ES, la performance pour la Scène H (le toit en construction) est significativement différente de celle des Scènes B (dressing), C (maison en construction) et G (intérieur avec des colonnes).

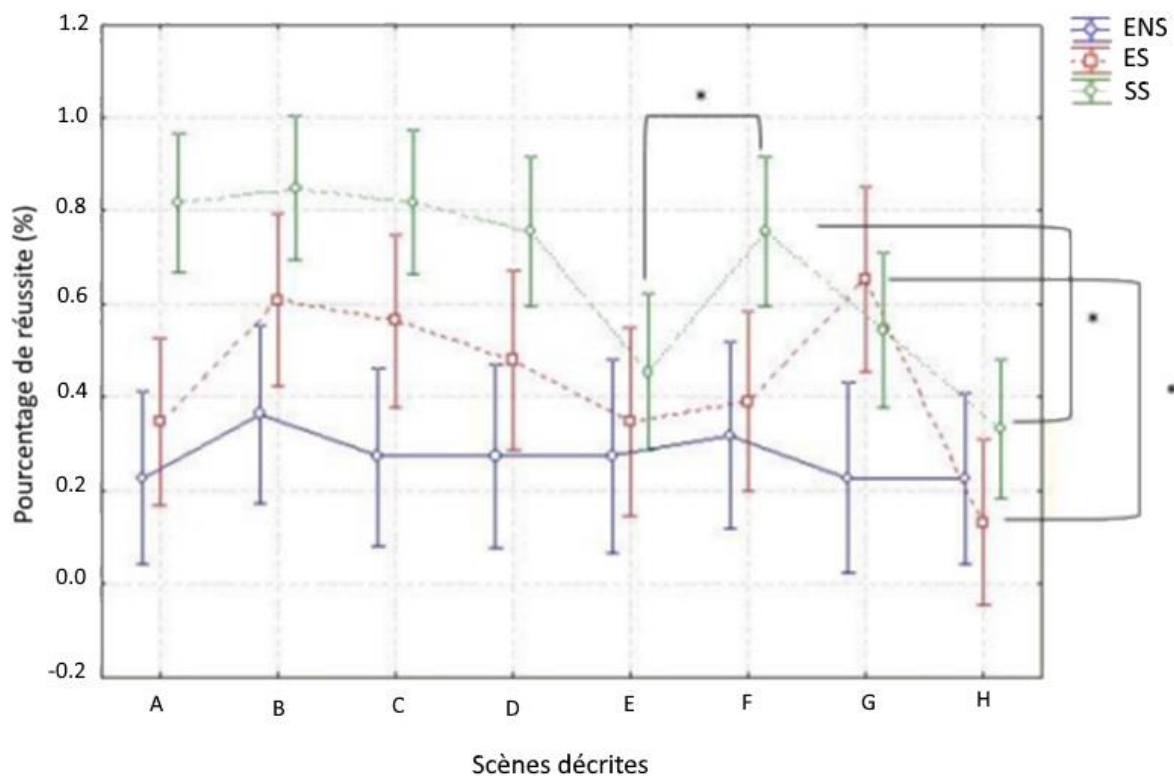


Figure 5-9: Performance des participants en fonction des Scènes décrites en abscisse et du Groupe en ligne.

Pour le groupe SS, la performance pour la Scène décrite H est significativement différente de toutes les autres excepté les Scènes E (ville avec monument) et G. Il en va de même pour la Scène E. Enfin, la performance pour la Scène G est significativement différente de celle pour

la Scène B. La performance du groupe ENS ne varie pas de manière significative pour les différentes Scènes décrites.

Pour tenter de comprendre pourquoi certaines Scènes décrites amenaient à des performances plus élevées ou plus faibles nous avons mené d'autres analyses. Nous avons utilisé l'annotation linguistique (voir chapitre 3) pour caractériser les 8 descriptions de scènes. L'analyse statistique par la corrélation de Pearson a été réalisée pour tester si le lien entre la quantité des différents indices linguistiques (i.e., Signes illustratifs, Signes Lexicaux, Pointage) pouvait être corrélée à la performance de chacun des groupes pour chaque Scène décrite. L'analyse n'a pas montré de corrélation significative entre les indices et la performance des participants.

5.2.8 Analyse descriptive et apports linguistiques

Une analyse descriptive et une approche linguistique nous permettent d'imaginer que les erreurs produites par les participants sur les vidéos seraient plutôt associées à la structure des autres scènes présentes sur la planche de réponse plutôt qu'aux catégories auxquelles elles appartiennent.

En effet, la Scène E, qui rapporte une performance plus faible des sujets sourds comparée aux autres scènes (figure 5-9), illustre cette idée. Au regard de la distribution des réponses des sujets sourds concernant cette scène (figure 5-10), 3 choix de photos se distinguent : la photo 1 (la bonne réponse), choisie par 16 sujets, la photo 10, choisie par 7 sujets et la photo 4, choisie par 3 sujets.

D'un point de vue linguistique le choix des photos 10 et 4 se rapporteraient notamment à leur structure, similaire à la Scène décrite recherchée (photo 1) : un élément central (une maison pour la photo 10, le quartier de la Défense pour la photo 1 et un immeuble pour la photo 4), un arrière-plan (e.g., des arbres, des éléments verticaux), un premier plan (e.g., un espace avant l'objet central, une route, de l'eau), de la perspective, etc. En l'absence d'informations manuelles (condition Avatar, sans doigts) interpréter des informations lexicales (le signe associé à un objet précis) semble plus complexe et ces signes lexicaux s'avèrent indispensables pour identifier une catégorie précise (tel que la catégorie Vin, Maison, etc.). Ainsi, en l'absence d'informations au niveau des doigts, les participants auraient interprété les mouvements des bras, beaucoup utilisés en LSF pour décrire la structure d'une scène. Les sujets auraient été amenés à se concentrer sur ces éléments structurels de la scène, ce qui peut expliquer leur erreur dans le cas de la description de la scène E, présentant une structure proche des deux autres photos (4 et 10) dans la planche de réponse.

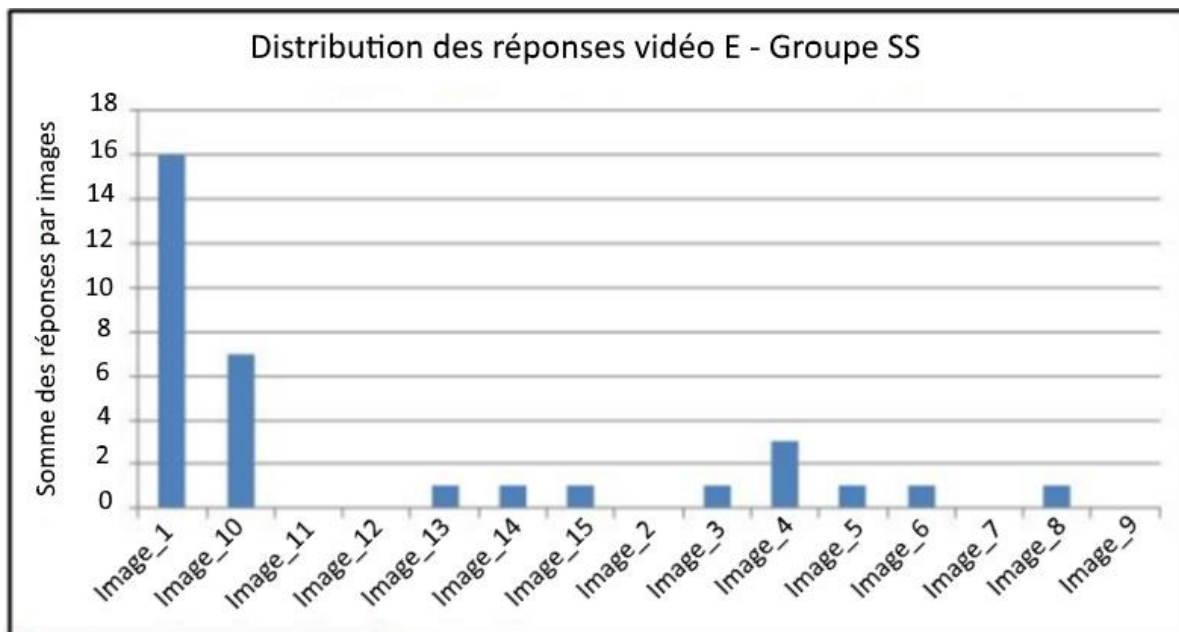


Figure 5-10: Analyse descriptive de la performance des sujets Sourds et Signeurs (SS) lors de l'observation de la description de la scène E (encadrée en rouge). En haut : Distribution du choix des sujets parmi les 15 photos. En Bas : Illustration de la planche de réponse vue par les sujet et composée des 15 photos en question.

5.3 Discussion

Cette étude avait pour objectif de comparer la compréhension des individus dans deux types de vidéo (un avatar signeur et une personne réelle). Les vidéos d'avatar ont été réalisées à partir de la mocap sur lequel des informations importantes en LSF (e.g., visage, doigts) ont été

supprimées. Ces sujets pouvaient avoir une connaissance de la LSF très variable, ils étaient soit sourds et locuteurs de la LSF, entendants et locuteurs de la LSF ou entendants et non locuteurs de la LSF.

Pour cela nous avons mis en place un test sur internet, au cours duquel les participants étaient confrontés à deux types de vidéos (CA et CN) (figure 5-3). A partir de celles-ci, ils devaient essayer de comprendre la description d'une Scène et indiquer celle qu'ils pensaient avoir reconnue parmi un ensemble de 15 photos (figure 5-4).

Les résultats montrent qu'il existe une relation entre la performance des participants à reconnaître ce qui leur est décrit, le groupe auquel ils appartiennent (SS, ES et ENS) et le type de vidéos (CA, CN) présentées. Ainsi, il apparaît que les performances pour les vidéos CN sont plus importantes comparées aux vidéos en CA pour les trois groupes de participants. Ces résultats sont en accord avec ceux d'autres études qui ont déjà démontré que la perception de la LS à partir d'un avatar ou d'une personne réelle peut amener des variations dans la manière dont l'individu prend l'information (Kacorri et al., 2013), ainsi que dans la reconnaissance des signes (Adamo-Villani et al., 2009). Dans notre étude, la compréhension des participants semble être affectée par l'apparence de la personne qui signe dans les vidéos (i.e., un avatar ou une personne réelle). Il est possible que les stratégies visuelles de prise d'informations en LS, notamment pour les groupes SS et ES, aient été perturbées. En effet, il a déjà été montré que les individus pratiquant la LS (débutants et experts) observaient principalement le visage d'un locuteur réel ou avatar (Emmorey et al., 2009; Kacorri et al., 2013; Muir & Richardson, 2005). En retirant les informations relatives au visage sur notre avatar, de nouvelles stratégies de prise d'informations pourraient avoir été mises en place rapidement pour tenter d'appréhender au mieux ce qui était décrit. Il est possible que les participants se soient alors concentrés sur des informations liées à la cinématique, la forme et les mouvements des bras par exemples.

Le gain d'informations (e.g. mains, expressions du visage) entre les vidéos de CA et de CN pourrait expliquer la différence de performance des participants entre ces deux types de vidéos. En effet, l'un des objectifs de cette étude était d'observer les effets de la suppression d'informations essentielles pour la compréhension de la LS sur un avatar et de voir dans quelle mesure le mouvement des membres supérieurs (hors mains et visage) et du tronc pouvait contribuer au message linguistique. L'étude menée par Tartter et Fischer sur la reconnaissance de signes manuels à partir de points lumineux ou d'une vidéo semble mener à des résultats similaires (Tartter & Fischer, 1982). En effet, les résultats obtenus par ces auteurs semblent

montrer une moins bonne localisation du mouvement, de l'orientation et de la forme des mains pour la condition points lumineux. Pour les auteurs, Ce résultat est lié au manque de redondance au niveau linguistique (e.g., élimination du visage, reste du corps). Ainsi, dans notre étude, la suppression d'informations manuelles et faciales sur notre avatar pourrait limiter la possibilité des participants de s'assurer complètement de la bonne compréhension de ce qui leur est décrit en LS. Cependant, la langue des signes met en œuvre de nombreuses composantes corporelles (e.g., mains, visage, buste) qui s'inscrivent dans des coordinations plus ou moins complexes pour fournir des informations. La somme de ces mouvements, porteurs d'informations, constituent alors un message global. Ces différentes informations pourraient créer une « redondance » chez l'observateur pour appréhender l'ensemble du message qui lui est transmis. Comme exemple, les informations du visage, par exemple, servent souvent à véhiculer des informations, par exemple dans la qualification de la taille d'un objet (e.g., gros, petit). Ainsi, même en supprimant des informations sur notre avatar, les participants sembleraient s'être appuyés sur d'autres informations pour comprendre ce qui leur a été décrit (e.g., mouvements des bras, mouvements des épaules, orientation de la tête). Cela rejoindrait une théorie proposée par Gibet, Marteau et Duarte pour qui les informations concernant les déplacements des articulations du corps seraient essentiellement relevées par les observateurs lors de la perception de la LS (Gibet, Marteau, & Duarte, 2012). De ce fait, la suppression d'informations sur l'avatar pourrait venir perturber la compréhension totale de la description, mais des informations importantes pourraient être encore utilisables par l'observateur. Cela pourrait expliquer, en partie, les résultats non nuls obtenus pour les vidéos de CA pour l'ensemble des participants.

Les résultats obtenus mettent également en évidence une variation de la performance des participants en fonction du groupe auquel ils appartiennent. Nous observons alors une performance plus faible pour les Entendants Non-Signeurs (ENS) comparés aux Entendants Signeurs (ES) et Sourds-Signeurs (SS). Cependant, nous n'observons pas de différence de performance entre les personnes ES et SS sur l'ensemble des vidéos (CA et CN). Il semble plutôt raisonnable de penser que cette différence de performance entre ces trois groupes est le fruit d'un apprentissage de la LSF, pour les groupes ES et SS, et d'une non connaissance de la LSF par le groupe ENS. L'étude (Emmorey, Kosslyn, & Bellugi, 1993) semble également aller en ce sens. En effet, dans leur étude, les auteurs ont testé la capacité d'individus SS, ES, et ENS à générer des photos mentales et à maintenir des photos dans leur mémoire à court terme. Ces habiletés pourraient être associées à la production et à la compréhension de la LSF. Les résultats obtenus dans cette étude tendent à montrer que les individus SS et ES réussiraient à générer

mentalement des photos plus complexes comparés aux individus ENS. Pour les auteurs, les individus signeurs développeraient des habiletés d'imagerie visuelle spécifiques peut être liées à des besoins linguistiques associés à la LSF. Par exemple, pour comprendre les signes réalisés par un locuteur, la personne qui observe doit en permanence retourner mentalement l'espace créé par le signeur qui lui fait face (e.g., position des objets, des personnes). Un changement de perspective permanent semble devoir être établi par la personne qui observe (passer du point de vue du signeur vers son propre point de vue). Ainsi, le niveau d'apprentissage de la LSF semble être déterminant pour garantir une meilleure compréhension de la LSF sur l'ensemble des vidéos (CA et CN). Les résultats comparables obtenus par les individus SS et ES pour les vidéos de CN peuvent amener la question de savoir si le fait d'être sourd peut également être un facteur déterminant pour une meilleure compréhension de la LSF.

Cependant, si l'on se concentre seulement sur les résultats obtenus pour les vidéos en CA (Courbe bleue, figure 5-7), les résultats indiquent une différence de performance entre chacun des groupes : les individus SS ont une performance supérieure à celle des individus ES et les individus ES ont une performance supérieure aux individus ENS. Il semblerait qu'ici, le fait d'être sourd et de signer permet une meilleure compréhension de la LSF sur un avatar sans visage et sans doigts.

La question de l'expérience en LSF peut alors se poser. En effet, il est possible que l'âge d'apprentissage de la LSF, le nombre d'années d'expérience en LSF puisse être déterminant pour améliorer la compréhension sur notre avatar. Dans cette étude, les participants du groupe SS ont une moyenne d'âge du début d'apprentissage de la LS inférieure à celle des individus du groupe ES. Ces derniers résultats semblent coïncider avec ceux de l'étude (Best, Mathur, Miranda, & Lillo-Martin, 2010). Les auteurs de cette étude ont étudié l'effet de l'expérience en LS sur la perception de contrastes entre des signes manuels proches (ici la lettre U et la lettre V) (figure 5-11) en fonction de 4 groupes : SS précoces, SS tardifs, ES tardifs, et ENS. Pour cela les auteurs ont créé un continuum de 11 photos représentant des étapes de transformation de la forme de la main pour passer du signe U (i.e., index et majeur tendus et collés l'un à l'autre) au signe V (i.e., index et majeurs tendus et totalement écartés). Leurs résultats semblent montrer que la perception de variation entre les deux signes serait affectée par l'expérience chez les SS (précoces ou tardifs) mais pas chez les individus ES. D'autres études ont également montré l'existence d'adaptations spécifiques, liées à la surdité, pour le traitement d'informations visuelles. Ainsi, les individus atteints de surdité auraient, par exemple, un seuil de détection du mouvement plus bas que des individus entendant (Shiell, Champoux, & Zatorre,

2014), une augmentation de l'attention visuelle (Hauthal et al., 2012), de la mémoire visuelle pour des formes abstraites (Cattani, Clibbens, & Perfect, 2007), de la détection de variation d'angles de la direction du mouvement (Hauthal, Thorne, Debener, & Sandmann, 2014). Ainsi, il est possible que les individus SS bénéficient de ces adaptations, possiblement associées à la surdit , ce qui leur permettraient de relever des informations plus pertinentes pour la compr hension de notre avatar compar s aux individus ES. De plus, une  tude men e par Skipper et al sugg re que, pour comprendre ce qui est exprim  par un locuteur en LS, les personnes signeurs utiliseraient leurs connaissances sur la fa on de produire des mouvements des mains et des bras pour d tecter des informations sur la s mantique des mouvements produits par un interlocuteur (Skipper, Goldin-Meadow, Nusbaum, & Small, 2009). Plusieurs  tudes ont d j  montr  de meilleures capacit s d'imagerie mentale chez des individus signeurs (sourds ou entendant) pour des t ches de transformation spatiale (Keehner & Gathercole, 2007), de rotation mentale (Emmorey et al., 1993) par exemple. Ainsi, les personnes signeurs simuleraient l'action du geste observ  comme s'ils produisaient eux-m mes l'action motrice. Il est donc possible qu'une meilleure capacit  de simulation mentale, avec l'exp rience notamment, permettent d'encoder des informations plus pr cises concernant le contenu linguistique (e.g., des formes des mains, des directions). Cela pourrait permettre aux individus SS et ES de mieux comprendre ce qui est d crit par un avatar priv  de l'ensemble des informations linguistiques en compl tant les informations manquantes par une simulation issue de leur propre exp rience. Ainsi il appara t que les individus SS seraient dans de meilleures dispositions pour appr hender les informations issues du mouvement sur un avatar appauvri en informations.

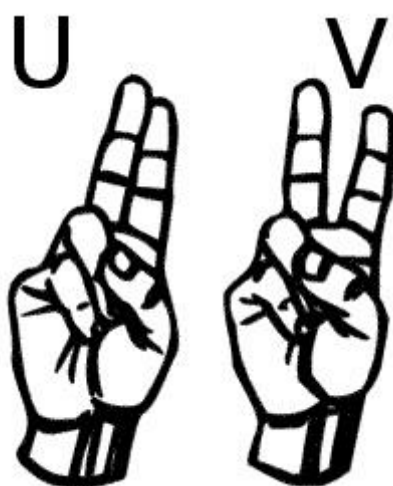


Figure 5-11 : Lettres U et V en ASL

Un autre résultat intéressant concerne les individus ENS. En effet, au regard de nos résultats nous observons que pour les vidéos en CA, bien que les performances soient plus faibles que pour les vidéos en CN, celles-ci ne sont pas nulles (note moyenne de 23 % de bonnes réponses) pour les individus ENS. Bien que n'ayant pas de connaissance de la LSF, ces individus réussissent à obtenir une performance moyenne constante et non-nulle pour l'ensemble des Scènes décrites (Courbe bleue, figure 5-9). Une étude avait déjà montré que des individus ENS réussissaient à décoder des émotions (i.e., colère, joie, neutralité) uniquement par des mouvements des mains en LS (Hietanen, Leppänen, & Lehtonen, 2004). Cela reprend d'autres résultats obtenus par des études menées sur la perception du mouvement biologique. A partir de simples points lumineux l'homme semble être capable de percevoir, par exemples, différentes caractéristiques sociales (e.g., sexe, identité, âge) (Brooks et al., 2008; Loula, Prasad, Harber, & Shiffrar, 2005; Montepare & Zebrowitz-Mcarthur, 1988), des actions (e.g marcher, sauter, peindre) (Vanrie & Verfaillie, 2004), un état émotionnel (Clarke, Bradshaw, Field, Hampson, & Rose, 2005), ou encore l'intention de communiquer et l'identification du geste de communication spécifique (e.g une action où « A demande à B de se rapprocher » et le geste associé était « un geste de la main », « A fait comprendre à B qu'il est en colère » et le geste associé était « une agitation nerveuse des bras et des mains ») (Manera, Schouten, Becchio, Bara, & Verfaillie, 2010). Ainsi, les résultats obtenus dans notre étude peuvent s'accorder avec l'hypothèse selon laquelle l'acquisition et le traitement du langage reposeraient sur les mêmes systèmes cognitifs et neuronaux que les systèmes engagés dans la perception et l'action (Barsalou, 1999). La présence de signes pouvant avoir une ressemblance avec la gestuelle utilisée avec le langage parlé pourrait avoir une influence sur la compréhension de ces vidéos en CA. En effet, pour Cuxac les LS comportent des structures linguistiques permettant de « montrer tout en disant » (Cuxac, 2000). Une analyse linguistique de notre corpus nous montre un nombre important de signes descriptifs indiquant la forme ou la taille de l'entité décrite, ou encore les relations spatiales entre entités. Il est possible que, même sans en connaître le sens, certains mouvements de LSF puissent activer des représentations mentales liées à un apprentissage de la langue (Thompson, Vinson, Woll, & Vigliocco, 2012), à l'expérience corporelle et personnelle de l'individu ENS.

5.4 Conclusion

Ce chapitre a été consacré à la présentation de l'étude de la perception du mouvement sur un signeur virtuel en LSF. Nous avons détaillé dans la première section la méthode que nous avons

mise en œuvre pour mettre le test en ligne, et dans la deuxième section, nous avons présenté les résultats et la discussion de cette étude.

En somme, nos résultats nous ont permis de valider notre première hypothèse et d'affirmer que les individus Sourds Signeurs (SS) ont des performances plus importantes que les individus Entendants Non-Signeurs (ENS) pour la compréhension des deux types vidéos (CA et CN). L'expérience de la LSF ainsi que des adaptations induites par l'absence d'audition peuvent être des facteurs à prendre en compte pour expliquer ces résultats.

Notre seconde hypothèse est également validée par nos résultats. Ainsi, les individus SS ont des performances plus faibles pour les vidéos en CA mais non nulles. Cela marque l'importance des informations dispensées au niveau du visage et des mains en LSF mais permet également de se rendre compte que les individus SS sont capables d'utiliser des informations à d'autres endroits sur cet avatar (mouvements des bras et du buste). Ces résultats ouvrent des perspectives de recherche pour déterminer où ces informations sont prélevées sur le locuteur et ainsi améliorer ces points cruciaux, en plus du visage et des mains, pour la compréhension des avatars signeurs. L'utilisation d'occulométrie pour de futures recherches peut être un moyen d'obtenir ces informations.

Enfin, un autre point intéressant soulevé par nos résultats est que les sujets Entendants et Non-Signeurs (ENS) ont, pour les vidéos en CA, un pourcentage de réussite constant et non-nul (différent de la réponse aléatoire). Il serait donc intéressant d'essayer de comprendre quels mécanismes cognitifs sont mis en jeu pour la compréhension, par des individus sans expertise de la LSF.

6 Conclusion générale

Dans cette thèse, nous avons présenté une première étude qui vérifie trois lois (le principe d'isochronie, la loi des asymétries et la loi de puissance $2/3$), dans le contexte des mouvements de LSF, et une deuxième étude sur la perception des mouvements de LSF sur un signeur virtuel. Dans les deux études, nous nous sommes basés sur un corpus de LSF composé de données 3D que nous avons créé et mis à disposition de la communauté scientifique.

6.1 Contributions de la thèse

Nous avons tout d'abord présenté le contexte dans lequel se place cette thèse qui s'articule sur trois domaines : la langue des signes, le contrôle moteur et la perception du mouvement humain, ces deux derniers champs constituant le domaine plus général de la cognition motrice. De plus, nous avons présenté l'intérêt de mener des recherches dans le domaine de la génération de la langue des signes basée sur des signeurs virtuels. Ensuite, nous avons justifié notre choix d'étudier les paramètres importants pour que la LSF générée automatiquement soit compréhensible, et pour cela le besoin d'étudier le mouvement de LSF afin de fournir des modèles de génération de mouvements plus naturels et plus compréhensibles. Nous avons aussi justifié le choix de créer un corpus de données 3D de LSF.

Sur la base de certaines études pertinentes soulignant l'importance des données 3D pour la recherche sur les LS, nous avons créé le corpus MOCAP1, un corpus de données 3D de LSF qui est mis à disposition pour la recherche pluridisciplinaire, notamment les études linguistiques, informatiques et en sciences du mouvement. Le corpus a un contenu riche qui contient plusieurs tâches, il est enregistré avec le système mocap et la vidéo, avec une durée de 2 heures et 20 minutes. Le corpus MOCAP1 a été annoté, et il est disponible pour les recherches sur la LSF sur demande. Il a déjà été utilisé dans des recherches en traitement d'images (Benseghir, Martinez, Gouiffès, & Braffort, 2016), et en linguistique (Collomb, 2017; Nogueira, 2016) au sein du laboratoire LIMSI, ainsi qu'en génération d'animation de signeur virtuel à DePaul University (Chicago).

L'étude bibliographique a mis en valeur l'importance du mouvement des segments corporels en LS, dans des différents types d'études. Les déplacements des mains semblent porteurs d'une partie non-négligeable de l'information linguistique bien qu'ils n'aient été que peu étudié dans le passé. Dans cette thèse, nous avons choisi de mieux comprendre les caractéristiques du mouvement de LSF et pour cela nous avons vérifié la robustesse de trois lois connues en

contrôle moteur. Nous avons trouvé que ces trois lois restent valides dans les deux types de mouvements « signe et transition », mais le principe d'isochronie et la loi des asymétries sont plus robustes dans les transitions. Notre hypothèse a été vérifiée : nous avons trouvé que la transmission des informations linguistiques durant les mouvements de signes a un impact sur ces lois. L'utilisation de ces résultats pourront permettre d'améliorer la génération des mouvements de LSF.

Un modèle de génération de LSF doit être capable de reproduire les différences entre signes et transitions et de générer des mouvements qui semblent naturels pour les signeurs. Ces travaux ont montré qu'il était globalement possible d'utiliser des modèles du contrôle moteur puisque les lois sont globalement vérifiées. Cependant, il y a quelques changements notables qui semblent liés à la présence d'informations linguistiques comme des arrêts de mouvements plus abruptes dans les signes.

En vue de bien comprendre la LSF générée à l'aide d'un signeur virtuel, et de mesurer le taux de transmission d'informations par les articulateurs du corps, nous avons mené une expérience de perception sur des vidéos d'une personne réelle et d'un signeur virtuel qui ne comportent pas le visage et les doigts. Nous avons trouvé que les individus Sourds Signeurs (SS) comprennent la LSF mieux que les Entendants Non Signeurs (ENS) dans les deux types de vidéo. De plus, les individus SS comprennent mieux les vidéos de la personne réelle. Ce résultat montre l'importance des informations transmises par le visage et les mains, mais permet également de se rendre compte que les individus SS sont capables d'utiliser les mouvements des bras et du buste pour reconnaître la description d'une image/scène.

6.2 Limites de cette thèse et perspectives

La route est encore longue avant d'aboutir à des modèles de générations de LS biologique et naturelle. Voici un certain nombre de perspectives ouvertes par cette thèse.

Les travaux de cette thèse ont consisté à étudier et vérifier les lois classiques du contrôle moteur dans les deux types de mouvements « signes » et « transitions ». Quoique nous ayons trouvé des différences entre ces deux types de mouvements, ces études peuvent être affinées. Premièrement, en étendant l'étude et en vérifiant les LCCM dans chaque type de signes et de transitions. En effet, les différents types de signes (signe lexical, signe illustratif, pointage, etc.) peuvent avoir des cinématiques de mouvements différentes, et les LCCM peuvent donner des résultats différents selon ces types. De plus, il y a deux types de transitions (transition inter-

signe et transition intra-signes), et il peut aussi y avoir des différences cinématiques et donc des différences dans les résultats sur les LCCM entre ces deux types. Il serait intéressant d'étudier les mouvements de LSF plus finement, mais cela nécessite un corpus encore plus fourni afin de détecter assez de mouvements dans chaque cas pour des analyses statistiques.

Ces études peuvent être aussi étendues pour étudier les postures complexes des bras durant les LS et leurs effets sur les deux types de mouvements. Des calculs de cinématique inverse permettraient de calculer les angles articulaires et d'effectuer d'autres analyses sur les coordinations inter-segmentaires. D'autres paramètres peuvent être étudiés comme les couples articulaires en considérant cette fois des calculs de dynamique inverse. Cela nécessite une connaissance de la distribution des masses des membres supérieurs. L'inertie pourrait d'ailleurs avoir une influence sur les mouvements de LS tout comme elle l'a sur les mouvements de contrôle moteur classiques. L'étude de ces paramètres permet de comprendre mieux la structure globale des mouvements de LS, et permet d'apporter plus d'informations aux modèles de générations comparé à notre étude qui s'est focalisée sur les mouvements des poignets.

Une étude intéressante envisageable dans le futur consisterait à vérifier ces lois du contrôle moteur sur la gestuelle de communication qui accompagne la parole des entendants. Cela permettrait de compléter et vérifier les résultats qui sont vraiment propres à la LSF comparé à de la gestuelle commune.

Les études réalisées dans le cadre de cette thèse nous ont permis de mieux comprendre les caractéristiques du mouvement en LSF au niveau des mains. Nous avons identifié des différences entre la cinématique des mouvements des signes et des transitions. Ainsi, nous envisagerons dans les futurs travaux d'utiliser les résultats de ces études pour proposer un modèle de génération de mouvements biologiques. Dans cette perspective, il serait utile d'étudier la pertinence de modèles de contrôle optimal. On pourrait imaginer utiliser des via points pour marquer les formes ou les changements de directions de signes, et des fonctions de coût différentes pour marquer les différences cinématiques entre signes et transitions. Les résultats d'une telle étude pourraient rendre l'animation des signeurs virtuels, dans le cadre de projet de génération automatique plus naturelle.

Le corpus que nous avons créé dans le cadre de cette thèse est assez riche pour les recherches pluridisciplinaires. Mais malgré la diversité de son contenu, il reste limité. Le manque de données 3D des mouvements de doigts ne permet pas de réaliser certaines études pour lesquelles la posture des doigts est nécessaire. Comme les mains et les doigts comportent une grande part

de l'information, la réalisation d'un corpus qui inclut les mouvements des doigts serait un grand plus pour la communauté scientifique qui s'intéresse aux LS.

Afin d'améliorer les résultats de l'étude de la perception des LS, il est indispensable de mieux déterminer les informations visuelles nécessaires pour la compréhension des signeurs virtuels. Dans cette perspective, une étude est en cours pour explorer l'effet du retrait de certaines informations sur la compréhension du message délivré par un signeur virtuel. Pour cette étude, nous utiliserons des vidéos de signeur virtuel issu du corpus MOCAP1 avec des conditions différentes (i.e. les mouvements de la tête et du buste du signeur virtuel bloqués, le signeur virtuel sans les mains, etc.). Cette étude permettrait de mesurer le taux de compréhension des signeurs virtuels sans les mains et le visage. Elle permettrait aussi de vérifier l'importance des mouvements de la tête et du buste pour la compréhension des signeurs virtuels.

Dans des études ultérieures, nous prévoyons d'étudier l'effet de l'expérience de la pratique de la LS, et l'adaptation induite par l'absence d'audition sur la compréhension de la LSF générée par des signeurs virtuels. Pour cela, une étude avec 2 groupes de sourds signeurs qui ont commencé à utiliser la LS à des âges différents (groupe 1 : dès la naissance, groupe 2 : après l'âge de 10 ans), et deux groupes d'entendants signeurs qui ont une expérience de la pratique de la LS différentes (groupe 1 : entre 3 et 8 ans d'expérience, groupe 2 : plus de 10 ans d'expérience).

Nous prévoyons aussi d'utiliser ces études de perception en ligne avec des sourds signeurs et des entendants signeurs pour évaluer les modèles de générations de la LS avec des signeurs virtuels.

Enfin, un approfondissement de nos connaissances nécessiterait d'étendre les disciplines réunies autour de ce sujet d'étude aux neurosciences et d'utiliser, par exemple, de l'imagerie par résonance magnétique fonctionnelle (IRMf)²⁶. Comme dans l'étude Newman et ses collaborateurs, qui ont cherché les zones du cerveau qui s'activent durant la langue des signes et les gestes non linguistiques (Newman, Supalla, Fernandez, Newport, & Bevelier, 2015). En suivant cette démarche, nous pourrions utiliser l'IRMf pour voir les zones du cerveau qui s'activent durant les mouvements de signes et les comparer avec celles qui s'activent durant les mouvements de transition. De plus, nous pourrions élargir ces études en utilisant l'IRMf pour

²⁶ Une application de l'imagerie par résonance magnétique permettant de visualiser, de manière indirecte, l'activité cérébrale

voir quelles zones s'activent lorsque les sourds signeurs observent la LSF générée à l'aide d'un signeur virtuel et lesquelles s'activent en observant un homme réel s'exprime en LSF et donc mieux appréhender les substrats neuronaux impliqués dans l'action et la perception de la LSF.

7 Bibliographie

- Adamo-Villani, N., Wilbur, R., Eccarius, P., & Abe-Harris, L. (2009). Effects of character geometric model on perception of sign language animation. In *Proceedings - 2009 2nd International Conference in Visualisation, VIZ 2009* (pp. 72–75). IEEE. <https://doi.org/10.1109/VIZ.2009.23>
- Alexanderson, S., & Beskow, J. (2015). Towards Fully Automated Motion Capture of Signs -- Development and Evaluation of a Key Word Signing Avatar. *ACM Transactions on Accessible Computing*, 7(2), 1–17. <https://doi.org/10.1145/2764918>
- Almohimeed, A., Wald, M., & Damper, R. (2010). An Arabic Sign Language Corpus for Instructional Language in School. *Computer*, 1–4. Retrieved from <http://eprints.ecs.soton.ac.uk/21106/>
- Angelaki, D. E. (2004). Eyes on target: what neurons must do for the vestibuloocular reflex during linear motion. *Journal of Neurophysiology*, 92(1), 20–35. <https://doi.org/10.1152/jn.00047.2004>
- Atkinson, A. P., Dittrich, W. H., Gemmell, A. J., & Young, A. W. (2004). Emotion Perception from Dynamic and Static Body Expressions in Point-Light and Full-Light Displays. *Perception*, 33(6), 717–746. <https://doi.org/10.1068/p5096>
- Augurelle, A.-S., Penta, M., White, O., & Thonnard, J.-L. (2003). The effects of a change in gravity on the dynamics of prehension. *Experimental Brain Research*, 148(4), 533–540. <https://doi.org/10.1007/s00221-002-1322-3>
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22(4), 577–609–60. <https://doi.org/10.1017/S0140525X99252144>
- Benseghir, K., Martinez, F., Gouiffès, M., & Braffort, A. (2016). *Descripteurs de mouvement pour l'annotation semi-automatique de vidéos de langue des signes. Journée FéDeV 2016 (FéDeV 2016)*. Evry, France.
- Bernstein, N. A. (1967). The co-ordination and regulation of movements: Conclusions towards the Study of Motor Co-ordination. *Biodynamics of Locomotion*. <https://doi.org/10.1097/00005072-196804000-00011>
- Berret, B., Bonnetblanc, F., Papaxanthis, C., & Pozzo, T. (2009). Modular control of pointing

- beyond arm's length. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 29(1), 191–205. <https://doi.org/10.1523/JNEUROSCI.3426-08.2009>
- Berret, B., Chiovetto, E., Nori, F., & Pozzo, T. (2011). Evidence for composite cost functions in arm movement planning: An inverse optimal control approach. *PLoS Computational Biology*, 7(10). <https://doi.org/10.1371/journal.pcbi.1002183>
- Berret, B., Darlot, C., Jean, F., Pozzo, T., Papaxanthis, C., & Gauthier, J. P. (2008). The inactivation principle: Mathematical solutions minimizing the absolute work and biological implications for the planning of arm movements. *PLoS Computational Biology*, 4(10). <https://doi.org/10.1371/journal.pcbi.1000194>
- Berret, B., & Jean, F. (2016). Why Don't We Move Slower? The Value of Time in the Neural Control of Action. *J Neurosci*, 36(4), 1056–1070. <https://doi.org/10.1523/JNEUROSCI.1921-15.2016>
- Berthier, F. (1852). *L'Abbé de l'Épée sa vie, son apostolat, ses travaux, sa lutte et ses succès*. Michel Lévy Frères. Retrieved from http://books.google.com/books?id=2I0t_Y0q2o4C&hl=&source=gbs_api
- Best, C. T., Mathur, G., Miranda, K., & Lillo-Martin, D. (2010). Effects of sign language experience on categorical perception of dynamic ASL pseudosigns. *Attention, Perception & Psychophysics*, 72(3), 747–762. <https://doi.org/10.3758/APP>
- Bézagu-Deluy, M. (1990). *L'abbé de L'Épée : instituteur gratuit des sourds et muets : 1712-1789*. Seghers.
- Binet, A., & Courtier, J. (1893). Sur la vitesse des mouvements graphiques. *Revue Philosophique de La France et de l'Étranger*, 35, 664–671. Retrieved from <https://philpapers.org/rec/BINSLV-2>
- Braffort, A., Blondel, M., Boutet, D., Boutora, L., Filhol, M., Dalle, P., & Leroy, E. (2016). *La langue des signes française (LSF) : modélisations, ressources et applications*. ISTE Editions.
- Braffort, A., Bolot, L., Chételat-Pelé, E., Choisier, A., Delorme, M., Filhol, M., ... Devos, N. (2010). Sign Language Corpora for Analysis, Processing and Evaluation. In *LREC* (pp. 453–456). Valetta, Malta. Retrieved from http://lrec.elra.info/proceedings/lrec2010/pdf/555_Paper.pdf

- Bresciani, J. P., Gauthier, G. M., Vercher, J. L., & Blouin, J. (2005). On the nature of the vestibular control of arm-reaching movements during whole-body rotations. *Experimental Brain Research*, *164*(4), 431–441. <https://doi.org/10.1007/s00221-005-2263-4>
- Broadbent, D. E. (1958). Perception and communication. *Education + Training*, *8*(6), 264–269. <https://doi.org/10.1108/eb015727>
- Brooks, A., Schouten, B., Troje, N. F., Verfaillie, K., Blanke, O., & van der Zwan, R. (2008). Correlated changes in perceptions of the gender and orientation of ambiguous biological motion figures. *Current Biology*. <https://doi.org/10.1016/j.cub.2008.06.054>
- Burger, B., & Toiviainen, P. (2013). Mocap Toolbox – a Matlab Toolbox for Computational Analysis of Movement Data. In R. Bresin (Ed.), *Proceedings of the Sound and Music Computing Conference 2013, SMC 2013, Logos Verlag Berlin, Stockholm, Sweden*, 172–178.
- Campbell, R., Capek, C. M., Gazarian, K., MacSweeney, M., Woll, B., David, A. S., ... Brammer, M. J. (2011). The signer and the sign: Cortical correlates of person identity and language processing from point-light displays. *Neuropsychologia*, *49*(11), 3018–3026. <https://doi.org/10.1016/j.neuropsychologia.2011.06.029>
- Catharine D. Barclay, James E. Cutting, L. T. K. (1978). Temporal and spatial factors in gait perception that influence gender recognition. *Perception and Psychophysics*, *23*(2), 145–152. <https://doi.org/10.3758/BF03208295>
- Cattani, A., Clibbens, J., & Perfect, T. J. (2007). Visual memory for shapes in deaf signers and nonsigners and in hearing signers and nonsigners: atypical lateralization and enhancement. *Neuropsychology*, *21*(1), 114–121. <https://doi.org/10.1037/0894-4105.21.3.399>
- Catteau, F., Blondel, M., Vincent, C., Guyot, P., & Boutet, D. (2016). Variation prosodique et traduction poétique (LSF/français) : Que devient la prosodie lorsqu'elle change de canal ? In *Journées d'Étude sur la Parole, Actes de la conférence conjointe JEP-TALN-RECITAL* (pp. 750–758). Paris, France. Retrieved from <https://hal-univ-evry.archives-ouvertes.fr/halshs-01354719/document>
- Clarke, T. J., Bradshaw, M. F., Field, D. T., Hampson, S. E., & Rose, D. (2005). The perception of emotion from body movement in point-light displays of interpersonal dialogue. *Perception*, *34*(10), 1171–1180. <https://doi.org/10.1068/p5203>

- Collomb, A. (2017). *Vers un traitement automatique de la référence en Langue des Signes Française : La modélisation syntaxique des proformes, Mémoire de master*. Université Paris Ouest Nanterre.
- Courtine, G., Papaxanthis, C., & Pozzo, T. (2002). Prolonged exposure to microgravity modifies limb endpoint kinematics during the swing phase of human walking. *Neuroscience Letters*, 332(1), 70–74. [https://doi.org/10.1016/S0304-3940\(02\)00909-6](https://doi.org/10.1016/S0304-3940(02)00909-6)
- Courtine, G., & Pozzo, T. (2004). Recovery of the locomotor function after prolonged microgravity exposure. I. Head-trunk movement and locomotor equilibrium during various tasks. *Experimental Brain Research*, 158(1), 86–99. <https://doi.org/10.1007/s00221-004-1877-2>
- Crasborn, O. A., & Zwitserlood, I. (2008). The Corpus NGT: An online corpus for professionals and laymen. In *Construction and Exploitation of Sign Language Corpora*. (pp. 44–49).
- Crevecoeur, F., Thonnard, J.-L., & Lefèvre, P. (2009). Optimal integration of gravity in trajectory planning of vertical pointing movements. *Journal of Neurophysiology*, 102(2), 786–96. <https://doi.org/10.1152/jn.00113.2009>
- Cuxac, C. (1996). *Fonctions et structures de l'iconicité des langues des signes*. Université Paris V.
- Cuxac, C. (2000). *La Langue des Signes Française (LSF). Les voies de l'iconicité. Faits de Langues*.
- d'Avella, A., Portone, A., Fernandez, L., & Lacquaniti, F. (2006). Control of Fast-Reaching Movements by Muscle Synergy Combinations. *Journal of Neuroscience*, 26(30), 7791–7810. <https://doi.org/10.1523/JNEUROSCI.0830-06.2006>
- d'Avella, A., Saltiel, P., & Bizzi, E. (2003). Combinations of muscle synergies in the construction of a natural motor behavior. *Nature Neuroscience*, 6(3), 300–308. <https://doi.org/10.1038/nn1010>
- Daniels, M. (1997). *Benedictine Roots in the Development of Deaf Education: Listening with the Heart*. Bergin & Garvey. Retrieved from https://books.google.fr/books/about/Benedictine_Roots_in_the_Development_of.html?id=9E5SOppb_rQC&redir_esc=y
- Dayan, E., Casile, A., Levit-Binnun, N., Giese, M. a, Hendler, T., & Flash, T. (2007). Neural

- representations of kinematic laws of motion: evidence for action-perception coupling. *Proceedings of the National Academy of Sciences of the United States of America*, 104(51), 20582–20587. <https://doi.org/10.1073/pnas.0710033104>
- De’Sperati, C., & Viviani, P. (1997). The relationship between curvature and velocity in two-dimensional smooth pursuit eye movements. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 17(10), 3932–3945.
- Derwort, A. (1938). Untersuchungen über den Zeitablauf figurierter Bewegungen beim Menschen [Investigations on the time course of tracing movements]. *Pflügers Archiv Für Die Gesamte Physiologie Des Menschen Und Der Tiere*, 240(6), 661–675. <https://doi.org/10.1007/BF01759969>
- Desloges, P. (1747-18. . ?. (1779). Observations d’un sourd et muèt, sur un cours élémentaire d’éducation des sourds et muèts, publié en 1779 par M. l’abbé Deschamps, chapelain de l’église d’Orléans. Retrieved from <http://gallica.bnf.fr/ark:/12148/bpt6k749465>
- Dreuw, P., Forster, J., & Ney, H. (2012). Tracking benchmark databases for video-based sign language recognition. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 6553 LNCS, pp. 286–297). https://doi.org/10.1007/978-3-642-35749-7_22
- Duarte, K., & Gibet, S. (2010). Heterogeneous Data Sources for Signed Language Analysis and Synthesis: The SignCom Project. In *International Conference on Language Resources and Evaluation (LREC)* (pp. 461–468). Valetta, Malta. Retrieved from <https://hal.inria.fr/hal-00503249/document>
- Efthimiou, E., Fotinea, S.-E., Hanke, T., Glauert, J., Bowden, R., Braffort, A., ... Goudenove, F. (2010). DICTA-SIGN: Sign Language Recognition, Generation and Modelling with application in Deaf Communication. Retrieved from <http://www.ilsp.gr/en/research/publications?view=publication&task=show&id=722>
- Efthimiou, E., Fotinea, S.-E., Hanke, T., Glauert, J., Bowden, R., Braffort, A., ... Lefebvre-Albaret, F. (2012). Sign Language technologies and resources of the Dicta-Sign project. In *LREC - 5th Workshop on the Representation and Processing of Sign Languages: Interactions between Corpus and Lexicon* (pp. 37–44).
- Emmorey, K., Kosslyn, S. M., & Bellugi, U. (1993). Visual imagery and visual-spatial

- language: Enhanced imagery abilities in deaf and hearing ASL signers. *Cognition*, 46(2), 139–181. [https://doi.org/10.1016/0010-0277\(93\)90017-P](https://doi.org/10.1016/0010-0277(93)90017-P)
- Emmorey, K., Thompson, R., & Colvin, R. (2009). Eye gaze during comprehension of American sign language by native and beginning signers. *Journal of Deaf Studies and Deaf Education*, 14(2), 237–243. <https://doi.org/10.1093/deafed/enn037>
- Encrevé, F. (2008). Réflexions sur le congrès de Milan et ses conséquences sur la langue des signes française à la fin du xix^e siècle. *Le Mouvement Social*, 223(2), 83. <https://doi.org/10.3917/lms.223.0083>
- Endres, D., Meirovitch, Y., Flash, T., & Giese, M. a. (2013). Segmenting sign language into motor primitives with Bayesian binning. *Frontiers in Computational Neuroscience*, 7(May), 68. <https://doi.org/10.3389/fncom.2013.00068>
- Filhol, M., & Hadjadj, M. N. (2016). Juxtaposition as a form feature – syntax captured and explained rather than assumed and modelled. In *LREC - 7th Workshop on the Representation and Processing of Sign Languages: Corpus Mining* (pp. 69–74).
- Filhol, M., Hadjadj, M. N., & Choisier, A. (2014). Non-manual features: the right to indifference. In *LREC - 6th Workshop on the Representation and Processing of Sign Languages: Beyond the Manual Channel* (pp. 49–54).
- Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 47(6), 381–391. <https://doi.org/10.1037/h0055392>
- Flash, T., & Hogan, N. (1985). The coordination of arm movements: an experimentally confirmed mathematical model. *The Journal of Neuroscience*, 5(7), 1688–1703. <https://doi.org/4020415>
- Flash, T., Meirovitch, Y., & Barliya, A. (2013). Models of human movement: Trajectory planning and inverse kinematics studies. *Robotics and Autonomous Systems*, 61(4), 330–339. <https://doi.org/10.1016/j.robot.2012.09.020>
- Forster, J., Schmidt, C., Hoyoux, T., Koller, O., Zelle, U., Piater, J., & Ney, H. (2012). RWTH-PHOENIX-Weather: A Large Vocabulary Sign Language Recognition and Translation Corpus. *International Conference on Language Resources and Evaluation*, 3785–3789. Retrieved from <http://www.lrec-conf.org/lrec2012>

- Forster, J., Schmidt, C., Koller, O., Bellgardt, M., & Ney, H. (2014). Extensions of the Sign Language Recognition and Translation Corpus. In *International Conference on Language Resources and Evaluation (LREC)* (pp. 1911–1916). Reykjavik, Island. Retrieved from <http://www.lrec-conf.org/proceedings/lrec2014/summaries/585.html>
- Freeman, F. N. (1914). Experimental analysis of the writing movement. *The Psychological Monographs*, *17*(4), 1–57. <https://doi.org/10.1037/h0093085>
- Freund, H. J., & Büdingen, H. J. (1978). The relationship between speed and amplitude of the fastest voluntary contractions of human arm muscles. *Experimental Brain Research*, *31*(1), 1–12. <https://doi.org/10.1007/BF00235800>
- Gachoud, J. P., Mounoud, P., Hauert, C. a, & Viviani, P. (1983). Motor strategies in lifting movements: a comparison of adult and child performance. *Journal of Motor Behavior*, *15*(3), 202–16. <https://doi.org/10.1080/00222895.1983.10735297>
- Gaveau, J., Berret, B., Angelaki, D. E., & Papaxanthis, C. (2016). Direction-dependent arm kinematics reveal optimal integration of gravity cues. *eLife*, *5*(NOVEMBER2016). <https://doi.org/10.7554/eLife.16394>
- Gaveau, J., Berret, B., Demougeot, L., Fadiga, L., Pozzo, T., & Papaxanthis, C. (2014). Energy-related optimal control accounts for gravitational load: comparing shoulder, elbow, and wrist rotations. *Journal of Neurophysiology*, *111*(1), 4–16. <https://doi.org/10.1152/jn.01029.2012>
- Gaveau, J., Paizis, C., Berret, B., Pozzo, T., & Papaxanthis, C. (2011). Sensorimotor adaptation of point-to-point arm movements after spaceflight: the role of internal representation of gravity force in trajectory planning. *J Neurophysiol*, *106*(2), 620–629. <https://doi.org/10.1152/jn.00081.2011>
- Gaveau, J., & Papaxanthis, C. (2011). The temporal structure of vertical arm movements. *PLoS ONE*, *6*(7). <https://doi.org/10.1371/journal.pone.0022045>
- Gentili, R., Cahouet, V., & Papaxanthis, C. (2007). Motor planning of arm movements is direction-dependent in the gravity field. *Neuroscience*, *145*(1), 20–32. <https://doi.org/10.1016/j.neuroscience.2006.11.035>
- Gibet, S., Marteau, P. F., & Duarte, K. (2012). Toward a motor theory of sign language perception. In *Lecture Notes in Computer Science (including subseries Lecture Notes in*

- Artificial Intelligence and Lecture Notes in Bioinformatics*) (Vol. 7206 LNAI, pp. 161–172). https://doi.org/10.1007/978-3-642-34182-3_15
- Gonzalez, M. (2012). Un système de segmentation automatique de gestes appliqué à la Langue des Signes. *JEP-TALN-RECITAL 2012, Atelier DEGELS 2012: Défi GEste Langue Des Signes*. Retrieved from http://degels2012.limsi.fr/actes/pdf/degels2012_7.pdf
- Grosjean, F. (1979). A study of timing in a manual and a spoken language: American sign language and English. *Journal of Psycholinguistic Research*, 8(4), 379–405. <https://doi.org/10.1007/BF01067141>
- Hauthal, N., Neumann, M. F., & Schweinberger, S. R. (2012). Attentional spread in deaf and hearing participants: face and object distractor processing under perceptual load. *Attention, Perception & Psychophysics*, 74(6), 1312–20. <https://doi.org/10.3758/s13414-012-0320-1>
- Hauthal, N., Thorne, J. D., Debener, S., & Sandmann, P. (2014). Source localisation of visual evoked potentials in congenitally deaf individuals. *Brain Topography*, 27(3), 412–424. <https://doi.org/10.1007/s10548-013-0341-7>
- Heimler, B., van Zoest, W., Baruffaldi, F., Rinaldi, P., Caselli, M. C., & Pavani, F. (2015). Attentional orienting to social and nonsocial cues in early deaf adults. *Journal of Experimental Psychology: Human Perception and Performance*, 41(6), 1758–1771. <https://doi.org/http://dx.doi.org/10.1037/xhp0000099>
- Heloir, A., Gibet, S., Multon, F., & Courty, N. (2006). Captured motion data processing for real time synthesis of sign language. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 3881 LNAI, pp. 168–171). https://doi.org/10.1007/11678816_19
- Hietanen, J. K., Leppänen, J. M., & Lehtonen, U. (2004). Perception of emotions in the hand movement quality of Finnish sign language. *Journal of Nonverbal Behavior*, 28(1), 53–64. <https://doi.org/10.1023/B:JONB.0000017867.70191.68>
- Huenerfauth, M., & Lu, P. (2010). Eliciting spatial reference for a motion-capture corpus of American Sign Language discourse. In *4th Workshop on the Representation and Processing of Signed Languages: Corpora and Sign Language Technologies*. Valetta, Malta. Retrieved from <http://huenerfauth.ist.rit.edu/pubs/huenerfauth-lu-2010-lrec.pdf>

- Huenerfauth, M., Lu, P., & Rosenberg, A. (2011). Evaluating importance of facial expression in american sign language and pidgin signed english animations. In *The 13th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS 2011)* (p. 99). Dundee, Scotland, United Kingdom. <https://doi.org/10.1145/2049536.2049556>
- Indovina, I., Maffei, V., Bosco, G., Zago, M., Macaluso, E., & Lacquaniti, F. (2005). Representation of visual gravitational motion in the human vestibular cortex. *Science (New York, N.Y.)*, *308*(5720), 416–9. <https://doi.org/10.1126/science.1107961>
- Jack, W. R. (1895). On The Analysis of Voluntary Muscular Movements by Certain New Instruments. *Journal of Anatomy and Physiology*, *29*(Pt 4), 473–478. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/17232148>
- Jantunen, T., Burger, B., Weerdt, D. De, Seilola, I., & Wainio, T. (2012). Experiences Collecting Motion Capture Data on Continuous Signing Collecting Continuous Mocap Data on Finnish Sign Language. In *LREC - 5th Workshop on the Representation and Processing of Sign Languages: Interactions between Corpus and Lexicon* (pp. 75–82).
- Jeannerod, M. (1984). The timing of natural prehension movements. *Journal of Motor Behavior*, *16*(3), 235–254. <https://doi.org/10.1080/00222895.1984.10735319>
- Jerde, T. E., Soechting, J. F., & Flanders, M. (2003). Coarticulation in fluent fingerspelling. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *23*(6), 2383–93. <https://doi.org/23/6/2383> [pii]
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, *14*(2), 201–211. <https://doi.org/10.3758/BF03212378>
- Kacorri, H., Harper, A., & Huenerfauth, M. (2013). Comparing native signers' perception of American Sign Language animations and videos via eye tracking. *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility - ASSETS '13*, 1–8. <https://doi.org/10.1145/2513383.2513441>
- Kandel, S., Orliaguet, J. P., & Viviani, P. (2000). Perceptual anticipation in handwriting: the role of implicit motor competence. *Perception & Psychophysics*, *62*(4), 706–716. <https://doi.org/10.3758/BF03206917>
- Keehner, M., & Gathercole, S. E. (2007). Cognitive adaptations arising from nonnative experience of sign language in hearing adults. *Memory & Cognition*, *35*(4), 752–761.

<https://doi.org/10.3758/BF03193312>

- Kipp, M. (2001). ANVIL A Generic Annotation Tool for Multimodal Dialogue. *Proceedings of the 7th European Conference on Speech Communication and Technology*, 1367–1370.
- Kipp, M. (2008). Spatiotemporal Coding in ANVIL. *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC)*, 3–6. Retrieved from <http://www.lrec-conf.org/proceedings/lrec2008/>
- Kipp, M. (2012). Multimedia Annotation, Querying, and Analysis in Anvil. In *Multimedia Information Extraction: Advances in Video, Audio, and Imagery Analysis for Search, Data Mining, Surveillance, and Authoring* (pp. 351–367). <https://doi.org/10.1002/9781118219546.ch21>
- Kipp, M. (2014). ANVIL: The Video Annotation Research Tool. In *Handbook of Corpus Phonology*.
- Kipp, M., Heloir, A., & Nguyen, Q. (2011). Sign language avatars: Animation and comprehensibility. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 6895 LNAI, pp. 113–126). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-23974-8_13
- Kyle, F. E., & Harris, M. (2006). Concurrent correlates and predictors of reading and spelling achievement in deaf and hearing school children. *Journal of Deaf Studies and Deaf Education*, 11(3), 273–288. <https://doi.org/10.1093/deafed/enj037>
- Lackner, J. R., & DiZio, P. (2000). Human orientation and movement control in weightless and artificial gravity environments. *Experimental Brain Research.*, 130(1), 2–26. <https://doi.org/10.1007/s002210050002>
- Lacquaniti, F., Terzuolo, C., & Viviani, P. (1983). The law relating the kinematic and figural aspects of drawing movements. *Acta Psychologica*, 54(1–3), 115–130. [https://doi.org/10.1016/0001-6918\(83\)90027-6](https://doi.org/10.1016/0001-6918(83)90027-6)
- Latash, M. L. (2008). *Synergy*. *Synergy*. <https://doi.org/10.1093/acprof:oso/9780195333169.001.0001>
- Lebourque, T. (1998). *Spécification et génération de gestes naturels. Application à la Langue des Signes Française*. Université Paris 11.

- Lefebvre-Albaret, F. (2010, October 7). *Traitement automatique de vidéos en LSF Modélisation et exploitation des contraintes phonologiques du mouvement*. Université Paul Sabatier - Toulouse III. Retrieved from <https://tel.archives-ouvertes.fr/tel-00608768/>
- Lefebvre-Albaret, F., & Patrice Dalle. (2009). Body posture estimation in sign language videos. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 5934 LNAI, pp. 289–300). https://doi.org/10.1007/978-3-642-12553-9_26
- Levit-Binnun, N., Schechtman, E., & Flash, T. (2006). On the similarities between the perception and production of elliptical trajectories. *Experimental Brain Research*, 172(4), 533–555. <https://doi.org/10.1007/s00221-006-0355-4>
- Loula, F., Prasad, S., Harber, K., & Shiffrar, M. (2005). Recognizing people from their movement. *Journal of Experimental Psychology. Human Perception and Performance*, 31(1), 210–220. <https://doi.org/10.1037/0096-1523.31.1.210>
- Lu, P., & Huenerfauth, M. (2010). Collecting a Motion-Capture Corpus of American Sign Language for Data-Driven Generation Research. *Proceedings of the NAACL HLT 2010 Workshop on Speech and Language Processing for Assistive Technologies*, (June), 89–97. Retrieved from <http://www.aclweb.org/anthology/W/W10/W10-1312>
- Lu, P., & Huenerfauth, M. (2011). Collecting an American sign language corpus through the participation of native signers. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 6768 LNCS, pp. 81–90). https://doi.org/10.1007/978-3-642-21657-2_9
- Lu, P., & Huenerfauth, M. (2012). CUNY American Sign Language Motion-Capture Corpus : First Release. In *LREC - 5th Workshop on the Representation and Processing of Sign Languages: Interactions between Corpus and Lexicon* (pp. 109–116).
- Lu, P., & Huenerfauth, M. (2014). Collecting and evaluating the CUNY ASL corpus for research on American Sign Language animation. *Computer Speech and Language*, 28(3), 812–831. <https://doi.org/10.1016/j.csl.2013.10.004>
- Luckner, J. L., & Handley, M. (2008). A Summary of the Reading Comprehension Research Undertaken With Students Who Are Deaf or Hard of Hearing. *American Annals of the Deaf*, 153(1), 6–36. <https://doi.org/10.1353/aad.0.0006>

- MacKenzie, C. L., & Iberall, T. (1994). *The Grasping Hand. Advances in Psychology* (Vol. 104). [https://doi.org/10.1016/S0166-4115\(08\)61577-0](https://doi.org/10.1016/S0166-4115(08)61577-0)
- Malaia, E., Borneman, J., & Wilbur, R. B. (2008). Analysis of ASL Motion Capture Data towards Identification of Verb Type. In *Proceedings of the 2008 Conference on Semantics in Text Processing* (pp. 155--164). <https://doi.org/10.3115/1626481.1626494>
- Malaia, E., & Wilbur, R. B. (2012). Kinematic Signatures of Telic and Atelic Events in ASL Predicates. *Language and Speech*, 55(3), 407–421. <https://doi.org/10.1177/0023830911422201>
- Malaia, E., Wilbur, R. B., & Milković, M. (2013). Kinematic parameters of signed verbs. *Journal of Speech, Language, and Hearing Research*, 56(October), 1677–1688. [https://doi.org/10.1044/1092-4388\(2013/12-0257\)We](https://doi.org/10.1044/1092-4388(2013/12-0257)We)
- Manera, V., Schouten, B., Becchio, C., Bara, B. G., & Verfaillie, K. (2010). Inferring intentions from biological motion: a stimulus set of point-light communicative interactions. *Behavior Research Methods*, 42(1), 168–78. <https://doi.org/10.3758/BRM.42.1.168>
- Maoz, U., Portugaly, E., Flash, T., & Weiss, Y. (2005). Noise and the two-thirds power law. *Advances in Neural Information Processing Systems 18*, 18(2), 851–858. Retrieved from <http://eprints.pascal-network.org/archive/00002157/>
- Massey, J. T., Lurito, J. T., Pellizzer, G., & Georgopoulos, a P. (1992). Three-dimensional drawings in isometric conditions: relation between geometry and kinematics. *Experimental Brain Research. Experimentelle Hirnforschung. Experimentation Cerebrale*, 88(3), 685–690. <https://doi.org/10.1007/BF00228198>
- Matthes, S., Hanke, T., Storz, J., Efthimiou, E., Dimiou, N., Karioris, P., ... Safar, E. (2010). Elicitation Tasks and Materials designed for Dicta-Sign's Multi-lingual Corpus. In . *4th Workshop on Representation and Processing of Sign Languages: Corpora and Sign Language Technologies/LREC conf* (pp. 158–163). Valetta, Malta.
- McDonald, J., Wolfe, R., Moncrief, R., & Baowidan, S. (2014). Analysis for Synthesis: Investigating Corpora for Supporting the Automatic Generation of Role Shift. In *LREC 2014 - NINTH INTERNATIONAL CONFERENCE ON LANGUAGE RESOURCES AND EVALUATION* (pp. 117–122).
- McDonald, J., Wolfe, R., Schnepf, J., Hochgesang, J., Gorman Jamrozik, D., Stumbo, M., &

- Berke, L. (2013). TOWARD LIFELIKE ANIMATIONS OF ASL: ACHIEVING NATURAL MOTIONS FROM THE MOVEMENT- HOLD MODEL. *Third International Symposium on Sign Language Translation and Avatar Technology (SLTAT 2013)*. Retrieved from <http://sltat.cs.depaul.edu/slides/McDonald.pdf>
- McIntyre, J., Zago, M., Berthoz, a, & Lacquaniti, F. (2001). Does the brain model Newton's laws? *Nature Neuroscience*, 4(7), 693–694. <https://doi.org/10.1097/00001756-200112040-00004>
- Méary, D., Kitromilides, E., Mazens, K., Graff, C., & Gentaz, E. (2007). Four-day-old human neonates look longer at non-biological motions of a single point-of-light. *PLoS ONE*, 2(1). <https://doi.org/10.1371/journal.pone.0000186>
- Michel, F. (1971). Etude experimentale de la vitesse du geste graphique. *Neuropsychologia*, 9(1), 1–13. [https://doi.org/10.1016/0028-3932\(71\)90055-8](https://doi.org/10.1016/0028-3932(71)90055-8)
- Montepare, J. M., & Zebrowitz-Mcarthur, L. (1988). Impressions of People Created by Age-Related Qualities of Their Gaits. *Journal of Personality and Social Psychology*, 55(4), 547–556. <https://doi.org/10.1037/0022-3514.55.4.547>
- Morrissey, S., Somers, H., & Smith, R. (2010). Building a Sign Language corpus for use in Machine Translation. In *4th Workshop on the Representation and Processing of Sign Languages at LREC*. Valetta, Malta. Retrieved from http://nclt.computing.dcu.ie/~sdandapat/publication/conf_paper/lrec_cslt10.pdf%5Cnhttp://rian.ie/en/item/view/48508.html
- Muir, L. J., & Richardson, I. E. G. (2005). Perception of sign language and its application to visual communications for deaf people. *Journal of Deaf Studies and Deaf Education*, 10(4), 390–401. <https://doi.org/10.1093/deafed/eni037>
- Neidle, C., & Vogler, C. (2012). A New Web Interface to Facilitate Access to Corpora : Development of the ASLLRP Data Access Interface (DAI). In *LREC - 5th Workshop on the Representation and Processing of Sign Languages: Interactions between Corpus and Lexicon* (pp. 137–142). Retrieved from <https://pdfs.semanticscholar.org/8d10/0e631c5b1443cf8b9bddbe1f89b2741aa5c4.pdf>
- Neri, P., Morrone, M. C., & Burr, D. C. (1998). Seeing biological motion. *Nature*, 395(6705), 894–6. <https://doi.org/10.1038/27661>

- Newman, A. J., Supalla, T., Fernandez, N., Newport, E., & Bevelier, D. (2015). Neural systems supporting linguistic structure, linguistic experience, and symbolic communication in sign language and gesture. *Proceedings of the National Academy of Sciences*, *112*(37), 11684–11689. <https://doi.org/10.1073/pnas.1510527112>
- Nogueira, S. (2016). *Annotation et analyse de signes iconiques dans un corpus de LSF, Mémoire de Licence*. Université d'Orléans.
- Papaxanthis, C., Dubost, V., & Pozzo, T. (2003). Similar planning strategies for whole-body and arm movements performed in the sagittal plane. *Neuroscience*, *117*(4), 779–783. [https://doi.org/10.1016/S0306-4522\(02\)00964-8](https://doi.org/10.1016/S0306-4522(02)00964-8)
- Papaxanthis, C., Pozzo, T., & McIntyre, J. (2005). Kinematic and dynamic processes for the control of pointing movements in humans revealed by short-term exposure to microgravity. *Neuroscience*, *135*(2), 371–383. <https://doi.org/10.1016/j.neuroscience.2005.06.063>
- Papaxanthis, C., Pozzo, T., Popov, K. E., & McIntyre, J. (1998). Hand trajectories of vertical arm movements in one-G and zero-G environments. Evidence for a central representation of gravitational force. *Experimental Brain Research*, *120*(4), 496–502. <https://doi.org/10.1007/s002210050423>
- Papaxanthis, C., Pozzo, T., & Schieppati, M. (2003). Trajectories of arm pointing movements on the sagittal plane vary with both direction and speed. *Experimental Brain Research. Experimentelle Hirnforschung. Experimentation Cerebrale*, *148*, 498–503. <https://doi.org/10.1007/s00221-002-1327-y>
- Papaxanthis, C., Pozzo, T., & Stapley, P. (1998). Effects of movement direction upon kinematic characteristics of vertical arm pointing movements in man. *Neuroscience Letters*, *253*(2), 103–106. [https://doi.org/10.1016/S0304-3940\(98\)00604-1](https://doi.org/10.1016/S0304-3940(98)00604-1)
- Papaxanthis, C., Pozzo, T., Vinter, A., & Grishin, A. (1998). The representation of gravitational force during drawing movements of the arm. *Experimental Brain Research*, *120*(2), 233–242. <https://doi.org/10.1007/s002210050397>
- Perrier, P., & Fuchs, S. (2008). Speed-curvature relations in speech production challenge the 1/3 power law. *Journal of Neurophysiology*, *100*(3), 1171–1183. <https://doi.org/10.1152/jn.01116.2007>

- Poom, L., & Olsson, H. (2002). Are mechanisms for perception of biological motion different from mechanisms for perception of nonbiological motion? *Perceptual and Motor Skills*, 95(3), 1301–1310. <https://doi.org/10.2466/pms.2002.95.3f.1301>
- Pozzo, T., Papaxanthis, C., Stapley, P., & Berthoz, A. (1998). The sensorimotor and cognitive integration of gravity. *Brain Research Reviews*. [https://doi.org/10.1016/S0165-0173\(98\)00030-7](https://doi.org/10.1016/S0165-0173(98)00030-7)
- PRESNEAU, J. R. (1998). *Signes et institution des sourds : XVIIIe-XIXe siècle*. Champ Vallon. Retrieved from https://books.google.fr/books/about/Signes_et_institution_des_sourds.html?id=LslIs60s4wC&redir_esc=y
- Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology*, 9(2), 129–154. <https://doi.org/10.1080/713752551>
- Qi, S., & Mitchell, R. E. (2012). Large-scale academic achievement testing of deaf and hard-of-hearing students: Past, present, and future. *Journal of Deaf Studies and Deaf Education*. <https://doi.org/10.1093/deafed/enr028>
- Rohlfing, K., Loehr, D., Duncan, S., Brown, A., Franklin, A., Kimbara, I., ... Sandra, T. (2006). Comparison of multimodal annotation tools — workshop report. *Gesprächsforschung - Online-Zeitschrift Zur Verbalen Interaktion*, 7, 99–123.
- Rosenbaum, D. A. (1991). *Human motor control*. Academic Press. Retrieved from https://books.google.fr/books/about/Human_Motor_Control.html?id=w3kVQBqSIEQC&redir_esc=y
- Runeson, S., & Frykholm, G. (1983). Kinematic specification of dynamics as an informational basis for person-and-action perception: Expectation, gender recognition, and deceptive intention. *Journal of Experimental Psychology: General*, 112(4), 585–615. <https://doi.org/10.1037/0096-3445.112.4.585>
- Rutkowski, P., Łacheta, J., Mostowski, P., Filipczak, J., & Łozińska, S. (2013). The corpus of Polish Sign Language (PJM): Methodology, procedures and impact. Retrieved from <https://ses.library.usyd.edu.au/handle/2123/9841>
- Saleh, M., Takahashi, K., & Hatsopoulos, N. G. (2012). Behavioral/Systems/Cognitive Encoding of Coordinated Reach and Grasp Trajectories in Primary Motor Cortex. *The*

- Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 32(4), 1220–32. <https://doi.org/10.1523/JNEUROSCI.2438-11.2012>
- San-Segundo, R., Pardo, J. M., Ferreiros, J., Sama, V., Barra-Chicote, R., Lucas, J. M., ... García, A. (2010). Spoken Spanish generation from sign language. *Interacting with Computers*, 22(2), 123–139. <https://doi.org/10.1016/j.intcom.2009.11.011>
- Santello, M., Baud-Bovy, G., & Jörntell, H. (2013). Neural bases of hand synergies. *Frontiers in Computational Neuroscience*, 7(April), 23. <https://doi.org/10.3389/fncom.2013.00023>
- Santello, M., Flanders, M., & Soechting, J. F. (2002). Patterns of Hand Motion during Grasping and the Influence of Sensory Guidance. *The Journal of Neuroscience*, 22(4), 1426–1435. <https://doi.org/22/4/1426> [pii]
- Sartori, L., Camperio-Ciani, A., Bulgheroni, M., & Castiello, U. (2013). Reach-to-grasp movements in macaca fascicularis monkeys: The isochrony principle at work. *Frontiers in Psychology*, 4(MAR), 1–5. <https://doi.org/10.3389/fpsyg.2013.00114>
- Schmitz, C. (2012). Limesurvey: an open source survey tool. *Limesurvey Project, Hamburg, Germany*. Retrieved from <http://www.limesurvey.org>
- Schnepp, J., Wolfe, R., McDonald, J., & Toro, J. (2013). Generating Co-occurring facial nonmanual signals in synthesized American sign language. In *GRAPP 2013 IVAPP 2013 - Proceedings of the International Conference on Computer Graphics Theory and Applications and International Conference on Information Visualization Theory and Applications* (pp. 407–416).
- Scott, S. S. H. (2004). Optimal feedback control and the neural basis of volitional motor control. *Nature Reviews. Neuroscience*, 5(7), 532–46. <https://doi.org/10.1038/nrn1427>
- Segouat, J., & Braffort, A. (2009). Toward the study of sign language coarticulation: Methodology proposal. In *Proceedings of the 2nd International Conferences on Advances in Computer-Human Interactions, ACHI 2009* (pp. 369–374). <https://doi.org/10.1109/ACHI.2009.25>
- Shiell, M. M., Champoux, F., & Zatorre, R. J. (2014). Enhancement of Visual Motion Detection Thresholds in Early Deaf People. *PLoS ONE*, 9(2), e90498. <https://doi.org/10.1371/journal.pone.0090498>
- Skipper, J. I., Goldin-Meadow, S., Nusbaum, H. C., & Small, S. L. (2009). Gestures Orchestrate

- Brain Networks for Language Understanding. *Current Biology*, 19(8), 661–667.
<https://doi.org/10.1016/j.cub.2009.02.051>
- Soechting, J. F., Lacquaniti, F., & Terzuolo, C. A. (1986). Coordination of arm movements in three-dimensional space. Sensorimotor mapping during drawing movement. *Neuroscience*, 17(2), 295–311. [https://doi.org/10.1016/0306-4522\(86\)90248-4](https://doi.org/10.1016/0306-4522(86)90248-4)
- Sommerville, J. A., & Decety, J. (2006). Weaving the fabric of social interaction: articulating developmental psychology and cognitive neuroscience in the domain of motor cognition. *Psychonomic Bulletin & Review*, 13(2), 179–200. <https://doi.org/10.3758/BF03193831>
- Sperry, R. W. (1952). Neurology and the mind-brain problem. *American Scientist*.
<https://doi.org/10.2307/27826433>
- Tarter, V. C., & Fischer, S. D. (1982). Perceiving minimal distinctions in ASL under normal and point-light display conditions. *Perception & Psychophysics*, 32(4), 327–334.
<https://doi.org/10.3758/BF03206238>
- Tarter, V. C., & Knowlton, K. C. (1981). Perception of sign language from an array of 27 moving spots. *Nature* (Vol. 289). <https://doi.org/10.1038/289676a0>
- Tasko, S. M., & Westbury, J. R. (2004). Speed-curvature relations for speech-related articulatory movement. *Journal of Phonetics*, 32(1), 65–80.
[https://doi.org/10.1016/S0095-4470\(03\)00006-8](https://doi.org/10.1016/S0095-4470(03)00006-8)
- Thelen, E., & Fisher, D. M. (1983). The Organization of Spontaneous Leg Movements in Newborn Infants. *Journal of Motor Behavior*, 15(4), 353–377.
<https://doi.org/10.1080/00222895.1983.10735305>
- Thompson, R. L., Vinson, D. P., Woll, B., & Vigliocco, G. (2012). The road to language learning is iconic: Evidence from British Sign Language. *Psychological Science*, 23(12), 1443–1448. <https://doi.org/10.1177/0956797612459763>
- Toro, J. A. (2004). Automated 3D Animation System to Inflect Agreement Verbs. In *Proceedings of the Sixth High Desert Linguistics Conference*. Albuquerque, New Mexico, November 4-6.
- Toro, J. A. (2005). *Automatic verb agreement in computer synthesized depictions of American Sign Language*. ProQuest Dissertations and Theses. Retrieved from <http://login.proxy.library.vanderbilt.edu/login?url=http://search.proquest.com/docview/3>

05356297?accountid=14816%5Cnhttp://sfx.library.vanderbilt.edu/vu?url_ver=Z39.88-2004&rft_val_fmt=info:ofi/fmt:kev:mtx:dissertation&genre=dissertations+%26+theses&sid=P

- Turvey, M. T. (2007). Action and perception at the level of synergies. *Human Movement Science*, 26(4), 657–697. <https://doi.org/10.1016/j.humov.2007.04.002>
- Tyrone, M. E., & Mauk, C. E. (2012). Phonetic reduction and variation in American Sign Language: A quantitative study of sign lowering. *Laboratory Phonology*, 3(2), 425–453. <https://doi.org/10.1515/lp-2012-0019>
- Tyrone, M. E., Nam, H., Saltzman, E. L., Mathur, G., & Goldstein, L. (2010). Prosody and movement in American Sign Language: A task-dynamics approach. *Speech Prosody*, 1–4. Retrieved from http://20.210-193-52.unknown.qala.com.sg/archive/sp2010/papers/sp10_957.pdf
- Vanrie, J., & Verfaillie, K. (2004). Perception of biological motion: a stimulus set of human point-light actions. *Behavior Research Methods, Instruments, & Computers : A Journal of the Psychonomic Society, Inc*, 36(4), 625–629. <https://doi.org/10.3758/BF03206542>
- Vinjamuri, R., Mao, Z. H., Scwabassi, R., & Sun, M. (2007). Time-varying synergies in velocity profiles of finger joints of the hand during reach and grasp. In *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings* (pp. 4846–4849). <https://doi.org/10.1109/IEMBS.2007.4353425>
- Vinjamuri, R., Sun, M., Chang, C.-C., Lee, H.-N., Scwabassi, R. J., & Mao, Z.-H. (2010). Dimensionality reduction in control and coordination of the human hand. *IEEE Transactions on Bio-Medical Engineering*, 57(2), 284–95. <https://doi.org/10.1109/TBME.2009.2032532>
- Viviani, P., Burkhard, P. R., Chiuvé, S. C., Dell’Acqua, C. C., & Vindras, P. (2009). Velocity control in Parkinson’s disease: A quantitative analysis of isochrony in scribbling movements. *Experimental Brain Research*, 194(2), 259–283. <https://doi.org/10.1007/s00221-008-1695-z>
- Viviani, P., & Flash, T. (1995). Minimum-jerk, two-thirds power law, and isochrony: converging approaches to movement planning. *Journal of Experimental Psychology. Human Perception and Performance*, 21(1), 32–53. <https://doi.org/10.1037/0096->

- Viviani, P., & McCollum, G. (1983). The relation between linear extent and velocity in drawing movements. *Neuroscience*, *10*(1), 211–218. [https://doi.org/10.1016/0306-4522\(83\)90094-5](https://doi.org/10.1016/0306-4522(83)90094-5)
- Viviani, P., & Schneider, R. (1991). A developmental study of the relationship between geometry and kinematics in drawing movements. *Journal of Experimental Psychology. Human Perception and Performance*, *17*(1), 198–218. <https://doi.org/10.1037/0096-1523.17.1.198>
- Viviani, P., & Stucchi, N. (1992). Biological movements look uniform: evidence of motor-perceptual interactions. *Journal of Experimental Psychology. Human Perception and Performance*, *18*(3), 603–623. <https://doi.org/10.1037/0096-1523.18.3.603>
- Viviani, P., & Terzuolo, C. (1982). Trajectory determines movement dynamics. *Neuroscience*, *7*(2), 431–437. [https://doi.org/10.1016/0306-4522\(82\)90277-9](https://doi.org/10.1016/0306-4522(82)90277-9)
- von Agris, U., Knorr, M., & Kraiss, K.-F. (2008). The significance of facial features for automatic sign language recognition. *2008 8th IEEE International Conference on Automatic Face & Gesture Recognition*, 1–6. <https://doi.org/10.1109/AFGR.2008.4813472>
- Watanabe, K., Matsuda, T., Nishioka, T., & Namatame, M. (2011). Eye gaze during observation of static faces in deaf people. *PLoS ONE*, *6*(2). <https://doi.org/10.1371/journal.pone.0016919>
- Wauters, L. N., Van Bon, W. H. J., & Tellings, A. E. J. M. (2006). Reading comprehension of Dutch deaf children. *Reading and Writing*, *19*(1), 49–76. <https://doi.org/10.1007/s11145-004-5894-0>
- White, O., McIntyre, J., Augurelle, A.-S., & Thonnard, J.-L. (2005). Do novel gravitational environments alter the grip-force/load-force coupling at the fingertips? *Experimental Brain Research*, *163*(3), 324–334. <https://doi.org/10.1007/s00221-004-2175-8>
- Wilbur, R. B. (1999). Stress in ASL: empirical evidence and linguistic issues. *Language and Speech*, *42* (Pt 2-), 229–250. <https://doi.org/10.1177/00238309990420020501>
- Wilbur, R. B. (2009). Effects of varying rate of signing on ASL manual signs and nonmanual markers. *Language and Speech*, *52*(2–3), 245–285.

<https://doi.org/10.1037/a0030561>.Striving

- Wilbur, R., & Schick, B. (1987). The effects of lexical stress on ASL signs. *Language and Speech*. <https://doi.org/10.1177/002383098703000402>
- Wise, S. P., & Shadmehr, R. (2002). Motor Control. *Encyclopedia of the Human Brain*, 3(1), 1–21. <https://doi.org/10.1016/B0-12-227210-2/00216-8>
- Wolfe, R., Cook, P., McDonald, J., & Schnepf, J. (2011). Linguistics as structure in computer animation: Toward a more effective synthesis of brow motion in American Sign Language. *Sign Language & Linguistics*, 14(1), 179–199. <https://doi.org/10.1075/sll.14.1.09wol>
- Wolfe, R., McDonald, J., Moncrief, R., Baowidan, S., & Stumbo, M. (2015). Inferring biomechanical kinematics from linguistic data : A case study for role shift. In *SLTAT - Sign Language Translation and Avatar Technology*. <https://doi.org/doi.org/10.13140/RG.2.2.28994.79044>
- Young, S. J., Pratt, J., & Chau, T. (2009). Target-Directed Movements at a Comfortable Pace: Movement Duration and Fitts's Law. *Journal of Motor Behavior*, 41(4), 339–346. <https://doi.org/10.3200/JMBR.41.4.339-346>
- Zafrulla, Z., Brashear, H., Hamilton, H., & Starner, T. (2010). A novel approach to American Sign Language (ASL) phrase verification using reversed signing. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, CVPRW 2010* (pp. 48–55). <https://doi.org/10.1109/CVPRW.2010.5543268>
- Zhao, W., Chai, J., & Xu, Y.-Q. (2012). Combining marker-based mocap and RGB-D camera for acquiring high-fidelity hand motion data. *Eurographics ACM SIGGRAPH Symposium on Computer Animation 2012*. <https://doi.org/10.2312/SCA/SCA12/033-042>
- Zupan, L. H., Merfeld, D. M., & Darlot, C. (2002). Using sensory weighting to model the influence of canal, otolith and visual cues on spatial orientation and eye movements. *Biological Cybernetics*, 86(3), 209–230. <https://doi.org/10.1007/s00422-001-0290-1>

8 Annexe 1 : Catégories des signes et attributs de l'annotation

Cette annexe présente les différentes catégories de signes (section 1), et définit les différents attributs que nous avons utilisés dans l'annotation du corpus (section 2).

8.1 Catégories des signes

Selon les théories linguistiques, la notion de signe est très variable. Christian Cuxac distingue plusieurs catégories de signes. Nous présentons ici une de théorie en usage en France. Nous définissons ici ces catégories :

- **Signe lexical** : qui se rapporte au lexique ou qui fonctionne comme une unité lexicale.
- **TTF** : Transfert de taille ou de forme, ces transferts réfèrent à des lieux, représentent des objets, des personnages particuliers de l'histoire racontée. Le narrateur construit avec ses mains, dans l'espace devant lui ou sur un endroit de son corps, des formes de base stable (configuration de la main) avec un mouvement et une orientation signifiant le déploiement de ses formes dans l'espace, pour en décrire la taille et la forme.
- **TS** : Transfert situationnel, ces transferts décrivent les déplacements d'un objet ou d'un personnage par rapport à un repère locatif fixe : le locatif est réalisé par la main dominée, la main dominante effectuant le déplacement de l'actant.
- **Pointage** : action de pointer avec la main dominante, le regard et l'épaule.

8.2 Les attributs de l'annotation du corpus

Cette section a pour objet de définir les différents attributs utilisés dans l'annotation linguistique et du mouvement du corpus.

L'attribut de structure temporelle :

- **Répétition** : il y a une ou plusieurs répétitions du signe.
- **Composition** : il y a une composition de signe.
- **Complexe** : il y a une composition et une répétition dans un signe.

L'attribut de relation entre les deux mains :

- Identique : les deux mains font le même mouvement dans la même direction.
- Symétrique : les deux mains font des mouvements opposés (par exemple : une main fait un mouvement vers l'intérieur, tandis que l'autre fait un mouvement vers l'extérieur).
- Alterné : les deux mains font des mouvements alternés (par exemple : une main fait un mouvement vers le haut, tandis que l'autre fait un mouvement vers le bas).
- Landmark : la main dominante (droite) fait un mouvement, tandis que la main dominée (gauche) est fixe dans un endroit précis.
- Buoy (ancree) : la même chose que dans Landmark, mais la seule différence est que la main dominée est fixe dans l'endroit où le signe précédent s'est terminé.
- Pointant : la main dominante pointe vers la main dominée.

Les directions des mouvements des mains :

- Ascendant (Haut) : mouvement de la main vers le haut selon l'axe longitudinal.
- Descendant (Bas) : mouvement de la main vers le bas selon l'axe longitudinal.
- Extérieur ou latéral : mouvement de la main vers l'extérieur du corps selon l'axe transversal.
- Intérieur ou médial : mouvement de la main vers l'intérieur du corps selon l'axe transversal.
- Avant ou antérieur : mouvement de la main vers l'extérieur du corps selon l'axe antéro-postérieur.
- Arrière ou postérieur : mouvement de la main vers l'intérieur du corps selon l'axe antéro-postérieur.

La figure 8-1 montre le tracé de l'axe antéro-postérieur, l'axe transversal l'axe longitudinal et les directions sur un squelette.

L'axe antéro-postérieur : passe horizontalement de l'avant à l'arrière. Lorsqu'une gymnaste effectue une roue, son corps tourne autour de cet axe.

L'axe transversal : passe horizontalement de droite à gauche. Lors d'une roulade ou d'un salto, le corps tourne autour de cet axe.

L'axe longitudinal : passe verticalement de bas en haut. En danse classique, lors d'une pirouette, le corps tourne autour de cet axe.

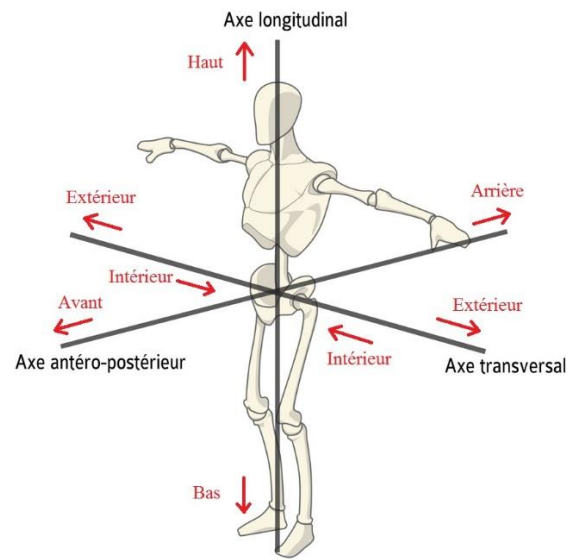


Figure 8-1: L'axe antéro-postérieur, transversal, et longitudinal.

9 Annexe 2 : Prétraitement des données du corpus

Cette annexe montre la méthode suivie pour importer les données (section 2.1), et pour extraire les données de déplacement du poignet (section 2.2).

9.1 Importation des données

Dans cette section, nous détaillons la méthode d'importation et de lecture dans Matlab, les données des marqueurs (section 2.1.1), d'annotation (section 2.1.2), et de synchronisation (section 2.1.3).

9.1.1 Importation et lecture des données de capture de mouvement

La première étape consiste à importer les données de la mocap à Matlab pour les analyser. Les données ont été enregistrées en trois formats CSV, C3D et BVH. Pour nos analyses, nous avons besoin de toutes les données des marqueurs, donc nous avons choisi les fichiers dans le format C3D.

Dans le cadre de cette thèse, nous avons analysé les 25 fichiers de données correspondant à la première tâche que nous avons enregistrée, c.à.d. la description des 25 images. Les données de chaque sujet ont été analysées séparément.

Premièrement, nous avons importé les données mocap. Ensuite, nous avons utilisé la fonction « mcread » de la Toolbox de Burger et Toiviainen (Burger et Toiviainen, 2013) pour les lire, cette fonction lit un fichier de mocap de format TSV, C3D, MAT, et WII. Elle retourne une structure de mocap contenant des valeurs de paramètres et des données. Les principales valeurs et paramètres qu'elle retourne sont : le nom, le nombre de frames, le nombre de caméras utilisées dans l'enregistrement, le nombre de marqueurs, la fréquence utilisée, les noms des marqueurs, et les données (les positions des marqueurs dans l'espace).

Afin de lire ces fichiers, nous avons créé une cellule de taille 25. Cette cellule sert à contenir les 25 structures mocap. Le stockage de toutes les données des 25 fichiers dans une seule cellule a facilité l'utilisation des données mocap. La figure 9-1 montre un exemple de lecture des données mocap.

```

% Importation et lecture des données mocap
mymocap = cell(1, numfiles);
for k = 1:numfiles
    myfilename1 = sprintf('i%d.c3d', k);
    mymocap{k} = mcread(myfilename1);
end

```

Figure 9-1: Importation et lecture des données mocap

9.1.2 Importation des données d'annotation

L'annotation a été enregistrée dans des fichiers textes. Ainsi, l'annotation du mouvement et l'annotation linguistique sont dans des fichiers séparés, cette séparation a facilité leurs utilisations dans le traitement. Autrement dit, chaque sujet à 50 fichiers d'annotation (25 fichiers d'annotation de mouvement et 25 fichiers annotation linguistique). Nous avons importé et lisons les annotations, puis les enregistrer dans une cellule (figure 9-2).

```

% Importation et lecture des données d'annotation
% Annotation du mouvement
mydata = cell(1, numfiles);
for k = 1:numfiles
    myfilename = sprintf('m%d.txt', k);
    file = fopen(myfilename);
    mydata{k} = textscan(file, '%s %s %f %f %f %s %s %s %s %s', 'headerlines', 1);
    fclose(file);
end
celldisp(mydata);

% Annotation linguistique
mydata2 = cell(1, numfiles);
for k = 1:numfiles
    myfilename2 = sprintf('s%d.txt', k);
    file2 = fopen(myfilename2);
    mydata2{k} = textscan(file2, '%s %s %f %f %f %s %s %s %s %s', 'headerlines', 1);
    fclose(file2);
end
celldisp(mydata2);

```

Figure 9-2: Importation et lecture des données d'annotation

9.1.3 Importation des données de synchronisation

L'annotation du corpus a été faite par rapport à la vidéo et non pas par rapport au squelette 3D, mais nous avons sauvegardé la synchronisation entre la vidéo et les données 3D (l'enregistrement des données a été manuel, ce qui a créé un décalage entre les vidéos et les données mocap, ensuite les vidéos et les données mocap ont été synchronisées lors de l'annotation, voire chapitre 3 section II.4.1). Afin d'utiliser l'annotation sur les données de mocap, nous avons besoin des décalages de la synchronisation. Ces décalages ont été récupérés et enregistrés dans un fichier texte.

Chaque sujet a un fichier de synchronisation, ce fichier contient les 25 décalages entre les données mocap et les annotations (nous avons 25 enregistrements de données qui correspondent à 25 vidéos et à 25 données 3D, entre chaque vidéo et ces données 3D il y a un décalage, au total il y a 25 décalages pour toutes les données). Ce fichier a été lu et utilisé pour synchroniser les données (figure 9-3).

```
% Importation et lecture des données de synchronisation
file = fopen('dec.txt');
dec = textscan(file, '%s %f', 'headerlines', 1);
fclose(file);
```

Figure 9-3: Importation et lecture des données de synchronisation

9.2 Extraction des données de déplacement

La deuxième tâche se compose de trois étapes :

- Trouver l'emplacement (le numéro) du marqueur dans la structure est la première étape, pour cela nous avons cherché avec le nom du marqueur dans l'endroit où sont stockés les noms des marqueurs. Nous avons utilisé une fonction qui compare le nom du marqueur voulu avec la liste des noms des marqueurs et donne l'endroit où se trouve le marqueur voulu. La figure 9-4 est un exemple pour trouver l'emplacement de l'un des deux marqueurs placés sur le poignet du bras droit, ici l'emplacement est dans RWRA.

```
% Pour trouver l'emplacement du marqueur du poignet du bras droite
g1{i} = strcmp(mymocap{i}.markerName, 'Right_arm_RWRA');
[r1(i), RWRA(i)] = max(g1{i});
```

Figure 9-4: l'emplacement de l'un des deux marqueurs poignet du bras droit

- La deuxième étape consiste à extraire les données des marqueurs des poignets des deux bras. Pour la faire, il suffit d'utiliser la fonction « mconcatenate » de la Toolbox que nous avons citée avec les numéros de marqueurs. Cette dernière extrait les données des marqueurs provenant des différentes structures de données mocap, et retourne ces données dans une structure mocap. La figure 9-5 montre l'exemple d'extraction des données de l'un des deux marqueurs placés sur le poignet droit.

```
% L'extraction des données des marqueurs des poignets
donr{i} = mconcatenate(mymocap{i}, RWRA(i));
```

Figure 9-5: Exemple d'extraction des données des marqueurs

- La séparation des données de déplacement par rapport aux 3 axes X, Y et Z est la troisième étape. Elle consiste à extraire les données de déplacement du marqueur RWRA de la structure mocap, et à mettre le déplacement sur chaque axe dans un vecteur.

10 Annexe 3 : Contenu du corpus

Dans cette annexe, nous présentons les photos du matériel d'élicitation du corpus que nous n'avons pas montré dans le chapitre 3 (section 3.1.1). La figure 10-1 présente les 25 photos décrites dans la première tâche du corpus. La figure 10-2 présente le matériel d'élicitation utilisé pour raconter l'histoire du Titanic.

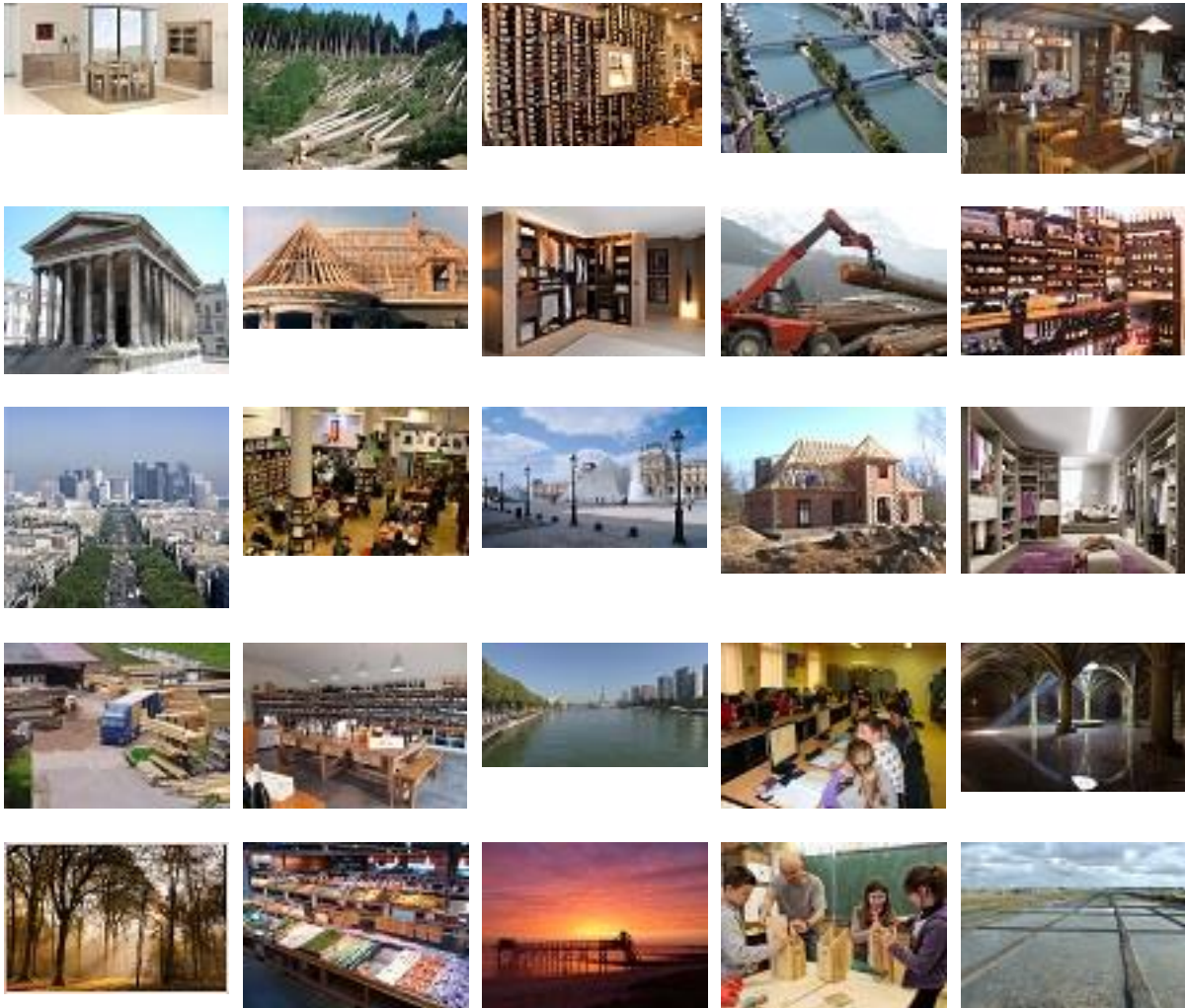


Figure 10-1: Photos décrites dans la première tâche du corpus



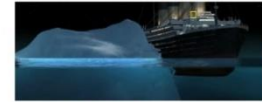
Son premier départ
le 10 avril 1912 à 12 h 15.



Itinéraire prévu



Titanic heurte un iceberg au large de l'Angleterre



Positionnement du Titanic lors de l'accident



Le Titanic disposait de 20 canots, dont 16 en bois, et 4 plâtrés.

Vue intérieure d'un des canots

Plus de 700 personnes participèrent au sauvetage du Titanic dont 200 passagers de première classe. Au sein de l'équipage 24% du total ont été secourus.



Il coula entre 23h00 et 2 h 20 le 15 avril au large de Terre-Neuve

causa 895 et 1 529 victimes (1 529 hommes, 109 femmes et 93 enfants)



Hommage au Titanic



Figure 10-2: Graphes relatant l'histoire du Titanic.

Titre : Contribution à l'analyse des mouvements 3D de la Langue des Signes Française (LSF) en Action et en Perception.

Mots clés : Contrôle Moteur, Langue des Signes, Perception, Mouvements Humain, Capture de Mouvement, Corpus.

Résumé : La langue des signes (LS) est encore une langue peu décrite, en particulier ce qui concerne le mouvement des articulateurs. La recherche sur la LS s'est concentrée sur la compréhension et la modélisation des propriétés linguistiques. Peu d'études ont essayé de comprendre la cinématique et la dynamique du mouvement en lui-même et ce que cela apporte à la compréhension de la LS générée par des modèles. Cette thèse porte sur l'analyse du mouvement en Langue des Signes Française (LSF) tant des points de vue de sa production et de sa compréhension par les individus sourds.

Mieux appréhender le mouvement en LS nécessite la création de nouvelles ressources pour la communauté scientifique étudiant les LSF. Dans cette perspective, nous avons créé et annoté un corpus de données de mouvements 3D de la partie supérieure du corps et du visage, à l'aide du système de capture de mouvement.

Le traitement de ce corpus a permis de spécifier la cinématique du mouvement en LSF durant les signes et les transitions.

La question posée dans la première partie de cette thèse était de quantifier dans quelle mesure certaines lois classiques connues en contrôle moteur restaient valides durant les mouvements de LS, afin de savoir si les connaissances acquises en contrôle moteur pouvaient être réutilisées en LS.

Trouver quelles informations du mouvement sont cruciales pour la compréhension de la LS a constitué la deuxième partie de cette thèse. L'intérêt était de savoir quels aspects du mouvement des modèles de production de LS devraient reproduire en priorité. Dans cette démarche, nous avons étudié dans quelle mesure des individus sourds, signeurs ou non, parvenaient à comprendre la LS en fonction de la quantité d'informations qui leur est disponible.

Title: Contribution to the analysis of the 3D movements of the French Sign Language (LSF) in Action and Perception.

Keywords : Motor Control, Sign Language, Perception, Human Movements, Motion-Capture, Corpus

Abstract: Nowadays, Sign Language (SL) is still little described, particularly for what concerns the movement of articulators. Research on SL has focused on understanding and modeling linguistic properties. Few investigations have been carried out to understand the kinematics and dynamics of the movement itself and what it brings to understand the LS SL generated by models. This thesis deals with the analysis of movement in the French Sign Language LSF with a main focus on its production as well as its understanding by deaf people.

Better understanding the movement in SL requires the creation of new resources for the scientific community studying SL. In this framework, we have created and annotated a corpus of 3D motion data from the upper body and face, using a motion capture system.

The processing of this corpus made it possible to specify the kinematics of the movement in SL during the signs and the transitions.

The first contribution of this thesis was to quantify to what extent certain classical laws, known in motor control, remained valid during the movements of SL, in order to know if the knowledge acquired in motor control could be exploited in SL.

Finding relevant information of the movement that is crucial for understanding SL represented the second part of this thesis. We were basically interested to know which aspects of the movement of SL production models should be replicated as a priority. In this approach, we have examined to what extent deaf individuals, whether signers or not, were able to understand SL according to the amount of information available to them.

