HAL
open science

Utilisation d'outils oculométriques dans le repérage d'indices visuels de la LSF en perception

Leila Boutora, Brigitte Meillon

To cite this version:

Leila Boutora, Brigitte Meillon. Utilisation d'outils oculométriques dans le repérage d'indices visuels de la LSF en perception. Traitement Automatique des Langues Naturelles , ATALA/AFCP, Jul 2010, Montréal, Canada. 8 p. hal-01803209

HAL Id: hal-01803209

https://hal.science/hal-01803209v1

Submitted on 1 Jun 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.

TALS 2010, Montréal, 23 juillet 2010

Utilisation d'outils oculométriques dans le repérage d'indices visuels de la LSF en perception

Leïla Boutora[1] Brigitte Meillon[2]

(1) LPL (Laboratoire Parole et Langage) / Université de Provence - UMR 6057
BP 80975 - 13604 Aix en Provence, France
(2) LIG (Laboratoire d'Informatique de Grenoble) - Equipe Multicom
BP 53 - 38041 GRENOBLE CEDEX 9, France
leila.boutora@lpl-aix.fr, Brigitte.Meillon@imag.fr

Résumé L'utilisation d'outils oculométriques constitue un nouveau tremplin dans l'exploration linguistique des langues des signes. Nous présentons ici deux expérimentations pilotes qui ont été menées à des niveaux distincts de l'analyse linguistique, phonétique et morpho-syntaxique, et évaluons les possibilités offertes par ces outils à partir de mesures de fixations et de trajets oculaires chez des signeurs et des non-signeurs.

Abstract Using eye-tracking tools is a new platform in exploring sign language linguistics. In this paper, we present two pilot experiments which were conducted at different levels of language analysis, phonetic and morpho-syntactic, and assess the potential of these tools from measurements of fixations and eye movements for signers and non-signers.

Mots-clés : LSF, perception, mesure de trajectoires oculaires, indices visuels
Keywords: LSF, perception, scanpath study, visual cues

1 Introduction

La présente contribution n'a pas la prétention de présenter les résultats d'une étude achevée. Son objectif est de soulever et de soumettre à la réflexion des chercheurs un certain nombre de questions concernant l'utilisation d'outils oculométriques dans l'étude des langues des signes (LS). Dans cette présentation, nous nous focaliserons sur la dimension perceptive de la langue, c'est-à-dire sur la perception mise en œuvre par le destinataire d'un message signé, et non sur la production du signeur comme c'est souvent le cas dans les études consacrées aux LS. En effet, cet angle d'approche est privilégié pour dégager à terme au sein de la langue – ici la langue des signes française (LSF) – les unités pertinentes que le destinataire utilise pour décoder le message signé. Nous posons donc l'hypothèse que le regard de l'interlocuteur portera son attention sur les éléments du discours qui sont utilisés dans le décodage de l'information à un niveau ou à un autre de l'organisation du langage.

A court terme, nous souhaitons évaluer les possibilités des oculomètres dans l'exploration de la perception d'un message en LSF, par des sourds signeurs et des entendants non signeurs en vue d'une comparaison des stratégies. A moyen terme, nous envisageons de fournir des données mesurables qui nous permettront à plus long terme de développer des outils descriptifs et d'investigation qui aujourd'hui font défaut pour continuer de manière efficace et objective la description linguistique des LS. Studdert-Kennedy & Lane (1980) ou encore Corina & Hildebrandt (2002) ont clairement mis en évidence la nécessité de développer ce type d'outils.

Nous présentons ici deux expérimentations pilotes qui ont été menées à des niveaux distincts de l'analyse linguistique. La première expérimentation vise à apporter des éléments de description de type phonétique. La seconde porte sur les éléments du discours de type morpho-syntaxique. Dans les deux cas, nous présentons les objectifs, la méthodologie et les données produites, mais ne concluons pas quant aux résultats de ces expériences. Ces expérimentations pilotes nous permettent avant tout de proposer ou préciser des pistes de recherche pour des études futures. Avant cela, nous commencerons par présenter le principe de fonctionnement des oculomètres et les possibilités offertes par le modèle utilisé dans ces pilotes : TOBII 1750.

2 Oculométrie : présentation

La technique d'*eye tracking* consiste à enregistrer en temps réel le parcours du regard de l'utilisateur, reflétant ainsi les opérations cognitives de l'individu. La succession des saccades (saut du regard vers un centre d'intérêt) et des fixations (pauses du regard) met en évidence la séquence de prise d'informations et l'ordre dans lequel l'utilisateur les traite. Les trajets oculaires de l'utilisateur sont enregistrés lors de la lecture, de la recherche d'information ou de l'inspection d'une scène visuelle, que ce soit sur une interface ou dans un environnement réel. L'oculométrie apporte des informations quantitatives sur l'efficacité de la prise d'information, les difficultés de traitement et les stratégies d'inspection du sujet. On pourra se reporter à Duchowski (2007) pour approfondir les éléments de définition proposés dans la section suivante.

2.1 Définition des concepts clés

Les **fixations** sont des pauses d'une durée moyenne de 200 ms, témoignant des traitements cognitifs de l'information perçue. Les fixations sont caractérisées par une durée et une position, qui sont calculées à partir des points bruts d'acquisition (points de mesure).

Une **saccade** est un saut très rapide d'une fixation à une autre, pendant lequel l'œil est « aveugle » et ne traite donc pas l'information. Les saccades relèvent de mécanismes perceptifs et oculomoteurs et sont caractérisées par une vitesse, une latence et une précision.

Un ***scanpath*** est l'ensemble du trajet oculaire, formé d'une séquence de saccades et de fixations, qui se représente visuellement superposé au stimulus. Les fixations sont généralement matérialisées par des cercles de diamètre proportionnel à leur durée.

Le **champ visuel** est décomposable en plusieurs régions. Pendant une fixation, l'œil a accès à la région fovéale (environ 2 degrés autour du point de fixation, pour une distance de 50 cm entre l'œil et la scène regardée), la région parafovéale (5 degrés), et la région périphérique, correspondant au champ visuel au-delà de la région parafovéale. L'information en vision fovéale et parafovéale peut être traitée pendant une fixation.

2.2 L'*eye tracker* TOBII 1750

L'*eye tracker* Tobii 1750 a été développé par la société suédoise Tobii. Il comprend un écran de 17 pouces dans lequel sont intégrées une caméra filmant l'œil du sujet (60 Hz) et une source de lumière infrarouge se reflétant dans l'œil. Les stimuli présentés aux sujets peuvent être des séquences d'images statiques, de la navigation Internet, des films, ou toute activité possible sur un écran d'ordinateur. Un logiciel d'acquisition et de traitement des données complète le système.

La technique utilisée est celle du reflet cornéen, basée sur le fait que la direction du regard peut être déterminée par la position relative du centre de la pupille par rapport au centre du reflet cornéen engendré par l'illumination de la pupille par la source de lumière infrarouge.

2.3 Données produites

Les données quantitatives fournies par l'*eye tracker* se présentent sous deux formes :

- soit les données d'acquisition brutes, sous forme de fichier texte, correspondant chacune à un *time code* suivi des coordonnées du point observé dans l'image, pour chaque image prise par la caméra (60 images par seconde)

- soit les données regroupées en points de fixation, par un algorithme interne. Dans ce cas, les données produites sont la séquence des fixations, chaque fixation étant caractérisée par un *time code*, les coordonnées du centre de la fixation et la durée de la fixation.

D'autre part, une représentation visuelle est disponible, soit sous forme de « rejeu » dynamique du parcours oculaire sur le stimulus regardé, soit sous forme de *scanpath* statique montrant l'ensemble du trajet superposé au stimulus.

Concernant le traitement des données, dans le cas de stimuli dynamiques, il est nécessaire de découper les films en séquences relativement statiques avant de définir graphiquement des zones d'intérêt (AOI pour *Area of Interest*) pour, selon les objectifs de l'étude, quantifier la prise d'information, repérer les zones explorées, connaître les durées de traitement, déceler les ré-inspections. A l'intérieur de ces zones d'intérêt, on peut calculer les indicateurs classiques en oculométrie, tels que le nombre des fixations par AOI, la durée totale et la durée moyenne de ces fixations, les transitions entre les AOI. La comparaison de

ces indicateurs suivant les conditions expérimentales permet d'identifier différents types de stratégies et de comportements cognitifs et d'aider à leur compréhension.

3 Expérimentation pilote 1 : formes manuelles

L'objectif de ce test pilote était d'évaluer la capacité du système Tobii à nous fournir des données d'un niveau phonétique ou descriptif de bas niveau concernant les formes manuelles de la LSF. Lors d'un travail antérieur qui portait sur l'étude expérimentale de la perception catégorielle des configurations manuelles en LSF (Boutora, 2008), nous avons pu pointer les limites et les lacunes des outils descriptifs du bas niveau à la disposition des chercheurs. En d'autres termes, nous manquons actuellement d'informations suffisantes sur les aspects phonétiques (de type articulatoires, perceptifs et acoustiques) concernant les LS. Nous souhaitons donc poursuivre le type de recherches initiées par Lane *et al.* (1976) sur la caractérisation en traits des éléments manuels des LS, caractérisation posée comme préliminaire à toute analyse de la parole (Jakobson *et al.*, 1952), qu'elle soit parlée ou signée.

3.1 Méthodologie du pilote 1

Le test a porté sur deux *continua* de formes manuelles UV (index et majeurs tendus et serrés qui s'écartent progressivement l'un de l'autre) et VX (index et majeur tendus et écartés qui se plient) définis et élaborés dans le cadre de l'étude expérimentale sur la perception catégorielle des configurations manuelles en LSF (Boutora, 2008). Chaque continuum a constitué une tâche séparée. Les sujets, deux sourds signeurs et un entendant non-signeur, devaient indiquer pour chaque continuum s'il voyait apparaître une nouvelle forme manuelle au fur et à mesure du défilement des *continua.* Ainsi, le but de cette tâche était de faire apparaître des catégories de formes manuelles *a posteriori* – et non pas données *a priori* comme c'est le cas dans les études expérimentales portant sur la perception catégorielle dans les LS (Emmorey *et al*. 2003, Baker *et al*. 2005), mais surtout de mettre en évidence les indices visuels qui permettaient aux sujets de prendre leur décision. La consigne était donc « Cliquez quand vous pensez voir une autre forme ».

3.2 Données produites

L'étude qualitative des trajets oculaires dynamiques des sujets (fig. 1) a permis de se faire une première idée sur les zones les plus regardées et donc de déterminer, *a posteriori*, les zones d'intérêt les plus pertinentes en regard de l'étude.

Pour le stimulus UV, un premier découpage distinguait 3 zones, respectivement formées par le pouce, le groupe index-majeur et le groupe annulaire auriculaire. La quasi-totalité des fixations se trouvant dans le groupe index-majeur, cette zone a été redécoupée en 6 zones, permettant d'affiner les résultats.

Pour le stimulus VX, le film a été scindé en 2 séquences, l'une dans laquelle les ongles sont encore invisibles (début du repliement), et l'autre dans laquelle les ongles sont visibles. Pour chacune de ces séquences, des zones d'intérêt ont été définies.

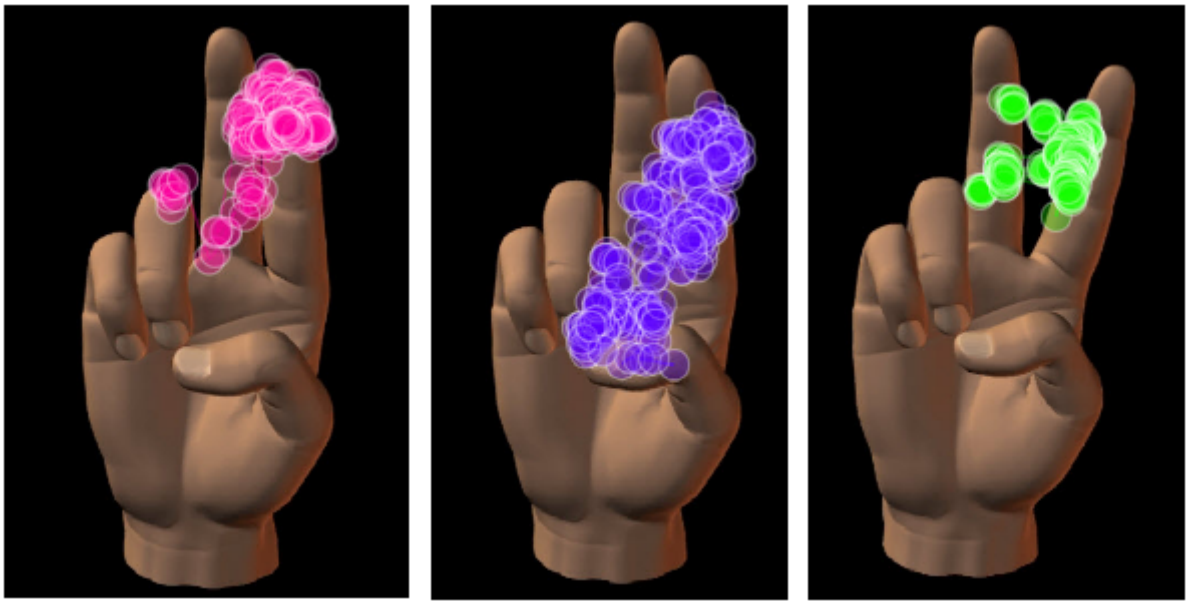

Figure 1: Répartition des fixations pour 3 sujets

Une fois ces zones définies (fig. 2), il est donc possible de calculer le nombre de fixations, leurs durées et de comparer ces indicateurs pour les différents sujets.

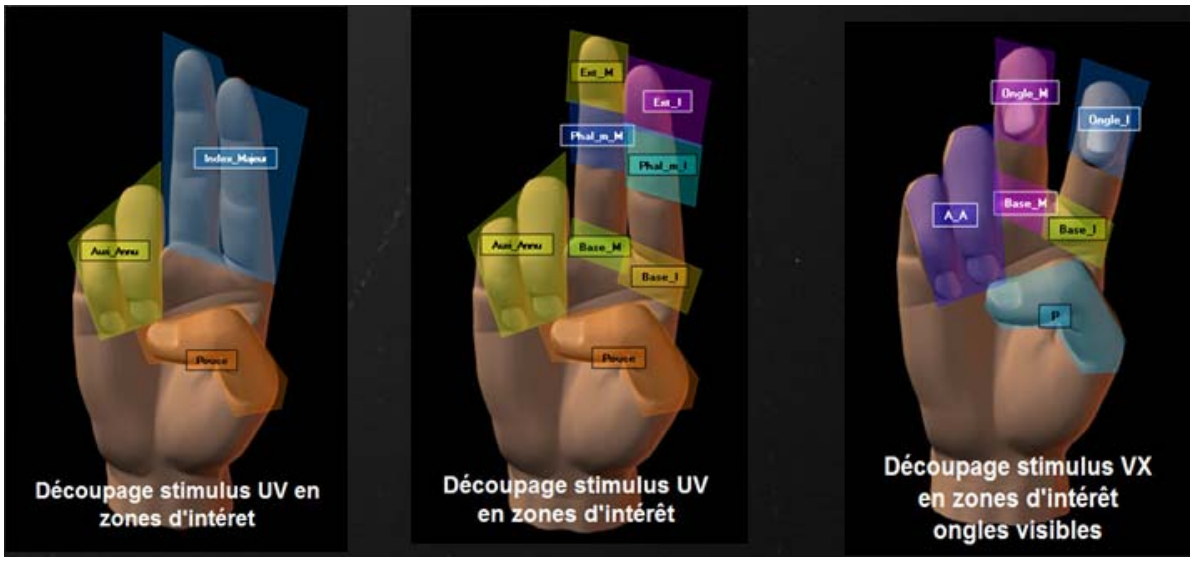

Figure 2 : Découpage en zones d'intérêt

4 Expérimentation pilote 2 : éléments morpho-syntaxiques et du discours

L'objectif de ce second test pilote était de vérifier la capacité du système Tobii à nous fournir des données utilisables sur les éléments morpho-syntaxiques ou du discours issus d'une production dialogique en LSF.

4.1 Méthodologie du pilote 2

Le test a porté cette fois sur un film présentant un dialogue impliquant un homme et une femme, tous deux locuteurs sourds de la LSF. Cette vidéo est extraite du corpus développé à l'occasion de l'édition 2005 de l'atelier TALS (Jardino, 2005). Il s'agissait de relever la destination du regard du sujet : son regard se focalise-t-il sur les yeux ou le visage du signeur ; les mains du signeur ; ou encore les zones de l'espace pointées par le signeur ? Les mêmes sujets ont participé à ce second test. La consigne donnée a été du type « regardez attentivement cette discussion, vous expliquerez ensuite ce qui s'est dit à votre ami ».

4.2 Données produites

Concernant cette expérimentation, la visualisation des *scanpaths* a montré une forte différence entre les trajets oculaires de sourds visualisant la production d'un signeur en LSF et d'un entendant non signeur visualisant la même séquence. On peut constater que les fixations des personnes sourdes sont complètement centrées sur le visage, alors que les fixations de la personne entendante sont beaucoup plus diffuses et sont réparties entre le visage du signeur et ses mains.

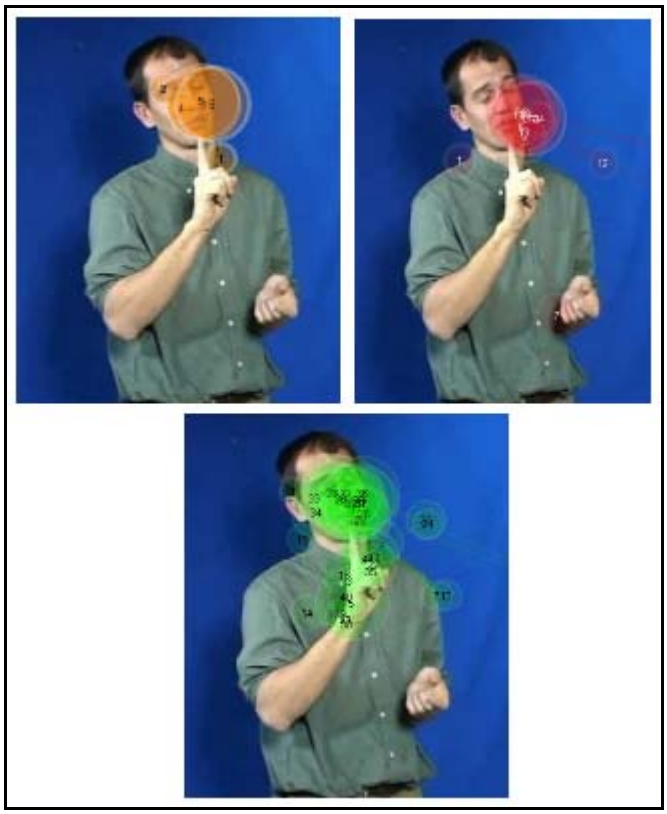

Figure 3 : Fixations : signeur 1 (haut gauche) ; signeur 2 (haut droite) ; non-signeur (en bas)

L'étude qualitative de ces trajets nous a porté à définir des zones d'intérêt sur le visage et les mains du signeur.

Les données quantitatives ont alors permis de montrer d'une part une absence de fixations en dehors du visage du signeur pour les 2 sujets sourds, et d'autre part une durée moyenne de fixation beaucoup plus importante (4 secondes environ) pour ces 2 sujets que pour le sujet entendant (une seconde environ).

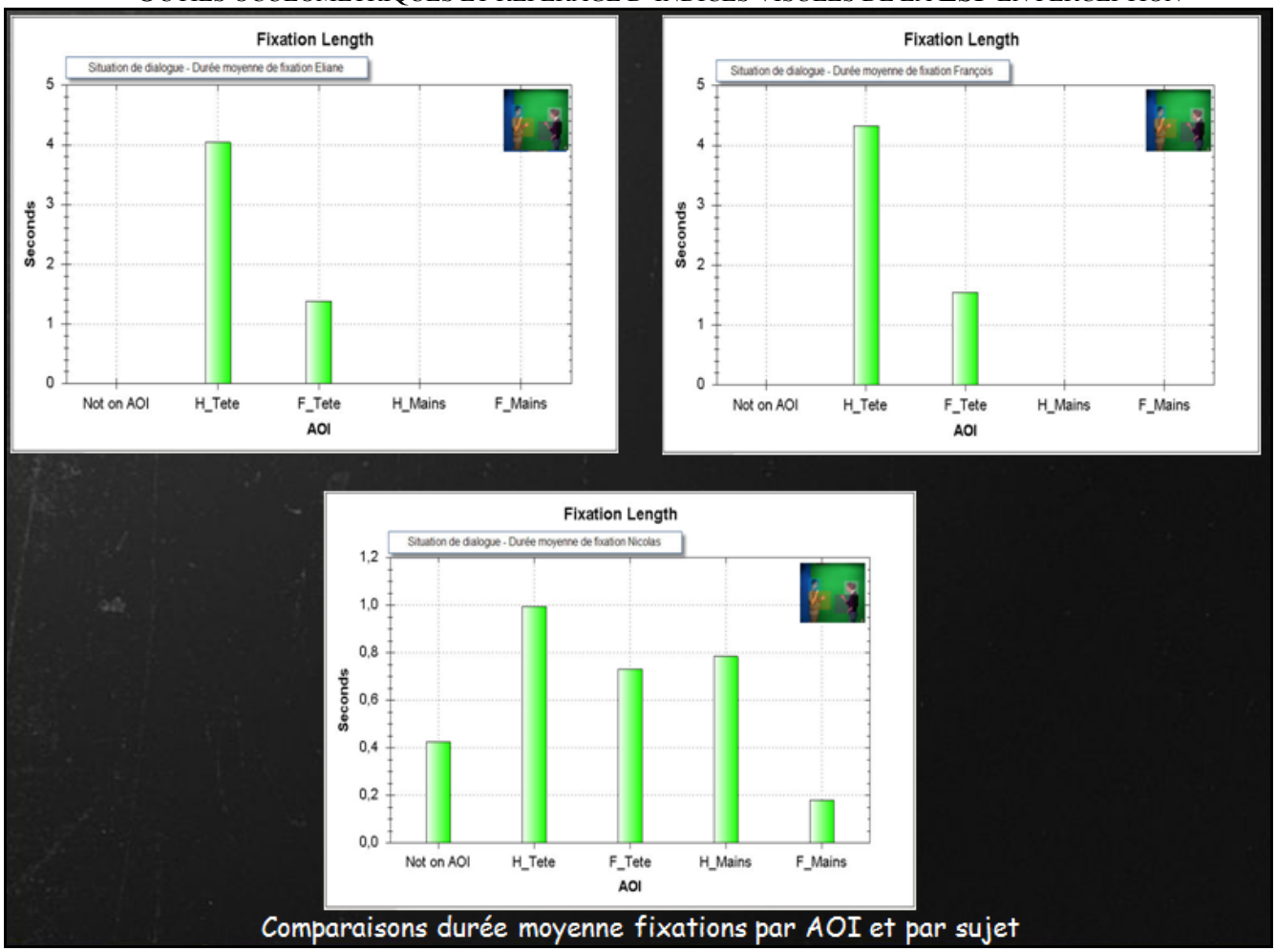

Figure 4 : Données quantitatives sur les éléments morpho-syntaxiques

Bien entendu, ces premiers résultats ne peuvent constituer qu'une piste de travail et doivent être confortés par une expérimentation plus complète avec un nombre de sujets des deux catégories plus significatif et un protocole expérimental plus complet.

5 Discussion et perspectives

Plusieurs éléments ont relevé notre attention et sont à prendre en compte dans le développement d'expérimentations futures. Dans le pilote 1, la question du « naturel » se pose de manière évidente : le mouvement des doigts apparaît-il naturel au signeur, ou bien contraint ? La texture même de l'image de synthèse gêne-t-elle le signeur et dans ce cas doit-on envisager le recours à des stimuli naturels ?

Concernant les zones d'intérêt, est-on sûr que les mêmes zones sont pertinentes à l'écran *vs* en situation dialogique réelle ? Comment réunir les conditions d'une situation dialogique réelle tout en mesurant les trajets oculaires de manière non invasive ?

Nous espérons que ces premières expérimentations ouvrent la voie à un large champs d'investigation dans la description des LS, mais pas seulement. Nous envisageons de poursuivre ces recherches pour aboutir à l'élaboration d'outils descriptifs de type phonétique qui nous permettront de développer les descriptions sur des bases fiables, quels que soient les niveaux d'analyse visés. D'autres utilisations des oculomètres

sont envisageables, comme par exemple l'évaluation des signeurs virtuels ; la didactique des LS ; la didactique de l'écrit chez l'enfant sourd ; la gestion des sous-titrages entre autres.

Remerciements

Nous tenons ici à remercier Eliane, François et Nicolas qui se sont spontanément prêtés au jeu et ont accepté de participer à ces tests totalement exploratoires. Nous remercions également toutes les personnes qui ont participé d'une manière ou d'une autre à ces tests, à leur mise en place ou à la réflexion qui les a précédés et, naturellement, à celle qui les suit actuellement.

Références

Remplacer toutes les perluètes par des virgules : (ex : BAKER S., ISARDI W., MICHNICK-GOLINKOFF R., PETTITO L.-A.)

DUCHOWSKI A. T. (2007). *Eye Tracking methodology. Theory and Practice*. 2e édition. Edition Springer

EMMOREY K., MCCULLOUGH S. & BRENTARI D. (2003). Categorical perception in American Sign Language. *Language and Cognitive Processes*, 18 (1), 21–45.

JAKOBSON R., FANT G. & HALLE M. (1952 [1963]). *Preliminaries to Speech Analysis. The Distinctive Features and their Correlates*. Cambridge : MIT Press.

JARDINO M. (ed.) (2005). Atelier TALS. Actes de *TALN 2005* (Traitement automatique des langues naturelles), Dourdan : ATALA LIMSI : http://tals.limsi.fr/.

LANE H., BOYES-BRAEM P. & BELLUGI U. (1976). Preliminaries to a distinctive feature analysis of handshapes in American Sign Language. *Cognitive Psychology*, 8, 263–289.

STUDDERT-KENNEDY M. & LANE H. (1980). Clues from the Difference between Signed and Spoken Language. In BELLUGI. & STUDDERT-KENNEDY (eds), *Signed and Spoken Languages : Biological Constraints on Linguistic Form*, pp. 29–40, Weinheim : Dahlem Konferenzen Verlag Chemie GmbH.

Site Internet Tobii : www.tobii.com