

Méthodologie d'exploration de corpus et de formalisation de règles grammaticales pour les langues des signes

Michael Filhol Annelies Braffort

LIMSI-CNRS, Campus d'Orsay bat 508, BP133, 91403 Orsay cx
michael.filhol@limsi.fr, annelies.braffort@limsi.fr

RÉSUMÉ

Cet article présente une méthodologie visant, à partir d'une observation de corpus vidéo de langue des signes, à repérer puis formaliser les régularités de structure dans les constructions linguistiques. Cette méthodologie est applicable à tous les niveaux du langage, du sub-lexical à l'énoncé complet. En s'appuyant sur deux exemples, il présente une application de cette méthodologie ainsi que le modèle AZee qui, intégrant la souplesse nécessaire en termes de synchronisation des articulateurs, permet une formalisation des règles repérées.

ABSTRACT

Methodology for corpus exploration and grammatical rule building in Sign Language

This paper presents a methodology for Sign Language video observation to extract and then formalise observed linguistic structure. This methodology is relevant to all linguistic layers from sub-lexical to discourse as a whole. Relying on two examples, we apply this methodology and describe the AZee model, which integrates the required flexibility for synchronising articulators, hence enables a specification of any new systematic rule observed.

MOTS-CLÉS : Langue des signes, analyse de corpus, modèle grammatical, synchronisation.

KEYWORDS : .Sign Language, corpus analysis, grammatical models, synchronisation.

1 Introduction

Notre objectif général est de représenter de manière formelle le fonctionnement des Langues des Signes (LS), ses éléments et ses règles. Ces représentations doivent nous permettre de générer des énoncés et produire automatiquement des animations en LS via un signeur virtuel (personnage virtuel en 3d s'exprimant en LS), et à terme d'envisager la traduction d'une langue écrite vers la LS (Filhol, 2011).

Les LS sont des langues peu dotées, dont les ressources (dictionnaires, livres de grammaire, méthodes pédagogiques, corpus...) sont très limitées. En France, la LSF n'est reconnue comme langue à part entière que depuis 2005¹. Les quelques descriptions existantes² sont sommaires et nous avons pu observer pour certaines d'entre-elles qu'elles ne résistent pas à la vérification sur corpus. C'est pourquoi notre démarche comporte

¹Loi 2005-102

²La langue des signes, Histoire et grammaire, IVT (ed) 1998 ; La LSF mode d'emploi, M. Companys (ed) 2003.

l'annotation et l'analyse d'un corpus de LS pour identifier des régularités de forme qui conduiront à la description de règles grammaticales. Ensuite, nous élaborons des modèles formels permettant de représenter ces règles.

Cet article décrit tout d'abord la méthodologie employée pour mettre en évidence les phénomènes systématiques d'une LS puis décrit le nouveau formalisme « AZee » proposé pour les représenter, en s'appuyant sur deux exemples.

2 La conception des règles

Cette section présente la méthodologie et le résultat d'une étude multilingue sur les LS française (LSF), grecque (GSL), anglaise (BSL) et allemande (DGS), réalisée pendant le projet européen Dicta-Sign³. Une structure linguistique identifiée est celle que nous nommons la structure de « qualification/désignation ». C'est un exemple représentatif du type de règle à représenter.

2.1 Méthodologie : de la forme à la fonction, et réciproquement

Deux approches sont possibles pour déterminer une règle systématique entre une structure ou une relation sémantique d'une part, et une production de surface (phonétique) d'autre part : soit à partir de la fonction sémantique, soit à partir de la forme de la surface. La structure présentée ici a été découverte au moyen de la deuxième approche, qui a comporté trois étapes :

(1) Choix des occurrences à collecter dans le corpus : Nous avons tout d'abord repéré un grand nombre d'occurrences où la posture de la main dominée était maintenue pendant que les gestes de la main dominante continuait, sans que les deux mains ne soient en relation pour des raisons géométrique ou topologique (comme lorsque la main dominante pointe vers la main dominée). Ceci nous a conduit à définir le critère de repérage de ces structures de forme, nommé « persistance indépendante de la main dominée », comme suit :

Un signe bimanuel s0, suivi par un ou plusieurs signes monomanuels de la main dominante, pendant que la posture finale de s0 est maintenue par la main dominée.

Main dominante : | _ s0 _ | | _ signes monomanuels s' _ _ _
Main dominée : | _ s0 _ _ _ _ _ maintenue _ _ _

(2) De la forme à la fonction en LSF : Dans la partie LSF du corpus, nous avons recueilli un minimum de 150 occurrences claires de la forme de la surface décrite en (1), et nous avons constaté que toutes correspondaient à l'une des deux catégories ci-dessous :

- a) *Qualification/dénomination* : La suite s' réalisée par la main dominante qualifie le signe s0 tel un adjectif, le nomme avec un « nom-signe » (nom propre en LS), ou encore épelle un mot (avec la dactylogogie) pour l'identifier. Cela peut être une combinaison de ces réalisations.
- b) *Conservation de l'activation* : s0 est tenu par la main dominée parce qu'il est à

³<http://www.dictasign.eu/>

nouveau nécessaire, après la séquence monomanuelle de la main dominante (s0 est souvent répété ensuite). Cela peut être considéré comme une parenthèse dans un discours, au cours de laquelle s0 doit être conservé « actif »

(3) De la fonction à la forme en différentes LSs : L'étape suivante a été de commencer un processus de vérification multilingue sur les parties LSF, DGS et GSL (respectivement LS française, allemande et grecque) du corpus Dicta-Sign. Toutes les langues ont été fouillées pour trouver des occurrences de la fonction sémantique de qualification/dénomination (2a) et les formes correspondantes observées. Les LS ont été analysées indépendamment par des experts de chaque LS et les résultats nous ont permis de confirmer nos observations sur la LSF et de proposer la règle suivante, commune aux trois LSs :

Lorsque s0 est un signe bimanuel suivi par un ou plusieurs signes monomanuels de qualification ou de dénomination, la main dominée a tendance à garder de manière ferme la dernière posture de s0, tandis que les autres signes sont effectués avec la main dominée.

Dans l'exemple LSF montré figure 1, il s'agit d'une qualification suivie d'une dénomination d'une ligne de métro sur un plan : le locuteur identifie la ligne jaune U3. La main dominée garde très clairement la posture finale du premier signe et est maintenue fermement tout au long des trois signes suivants, jusqu'à ce que les deux mains soient relâchées.



FIGURE 1 – Combinaison des quatre signes LIGNE JAUNE U 3.

2.2 Discussion

Ces premiers résultats doivent être affinés, que ce soit sur les formes associées à la fonction ou sur les fonctions associées à la forme. Mais dès à présent, ils nous permettent de mettre en lumière un certain type de contraintes pouvant s'exercer sur les événements manuels et qui sont probablement accompagnés d'autres contraintes à découvrir, sur des éléments non-manuels par exemple.

Nous avons remarqué la présence de ces structures (2a) dans la base de données de lexique de LSF construite pendant le projet. Dans cette base de données, chaque entrée est une unité lexicale lemmatisée, associée à un ou plusieurs concepts. Une entrée possédant une telle structure doit-elle être considérée comme une unité lexicale, ou s'agit-il d'une construction « syntaxique » à laquelle on peut associer un concept ? Le signe est-il une étape de la lexicalisation d'une structure, et comment trancher ? De plus, il est possible que des articulateurs non manuels soient porteurs d'une structure, comme

chez certaines observées dans le projet DictaSign, et leur synchronisation peut devenir d'autant plus complexe.

Pour en revenir à notre motivation initiale de concevoir des modèles informatiques, ces considérations plaident en faveur de représentations qui ne sont ni organisées autour de l'activité manuelle a priori, ni limités à des niveaux linguistiques spécifiques (lexique, syntaxe, etc.) mais proposent un point de vue global.

3 La représentation des règles

Cette section présente un état des lieux des modèles existants, le cahier des charges auquel selon nous doit répondre un modèle de description de la LS, puis le nouveau formalisme que nous proposons nommé AZee.

3.1 Modèles à composante temporelle

Le projet le plus abouti reste celui élaboré durant le projet européen ViSiCAST, basé sur HPSG (Marshall, 2004). Il est intégré dans un système de génération automatique d'énoncés en LS qui définit des séquences de mouvements ou de signes élémentaires séparés par des transitions de même nature (Elliott, 2004). Ce type de représentation n'intègre pas de système de synchronisation suffisant pour représenter les phénomènes liés à la multi-linéarité de la LSF.

Deux modèles font tout de même apparaître la multi-linéarité dans les descriptions. Liddell & Johnson (1989) ont montré que les signes étaient divisibles en unités temporelles où les articulateurs du corps se synchronisaient en *postures*, séparées par des unités de *transition*, alternant sur une ligne temporelle de description. Ce modèle reste en revanche comme ses prédécesseurs porté sur l'activité manuelle et la description lexicale dans sa forme de citation (dictionnaire), or les LS permettent la création spontanée et sémantiquement productive d'unités non répertoriées qui contrastent avec le vocabulaire figé (« standard ») en cela qu'elles mettent souvent en jeu de nombreux articulateurs non manuels (épaules, buste, muscles faciaux, etc.) qu'il faut synchroniser.

Le modèle P/C de Huenerfauth (2006) permet de diviser par endroits une ligne de temps en deux lignes parallèles pour spécifier deux activités simultanées. L'énoncé peut se représenter sous la forme d'un arbre où les feuilles sont des signes lexicaux et les nœuds intermédiaires sont chacun :

- soit de type C (constituant), dont les enfants sont des sous-parties de l'énoncé à concaténer ;
- soit de type P (partition), dont les enfants sont des sous-parties de l'énoncé à paralléliser.

Le problème est alors que les nœuds P et C partagent systématiquement les mêmes bornes temporelles et ne peuvent se chevaucher librement à moins d'utiliser des nœuds spéciaux « Ø » qui ne représentent rien linguistiquement et rendent les descriptions fastidieuses.

3.2 Clés pour un nouveau modèle

Nous proposons un nouveau formalisme de description nommé Azee. En utilisant deux méthodes de synchronisation combinées, AZee peut décrire n'importe quel motif de synchronisation des articulateurs du corps en LS.

Dans le cas général, un groupe d'articulateurs dans une production signée a une période d'activité pendant laquelle ils concourent à l'énoncé et hors de laquelle ils sont ou retournent dans une position de repos. Cette période est appelée « intervalle » et notée « TI » (time interval). Par exemple, le schéma suivant montre 5 TI synchronisés sur un axe temporel qui correspondant à l'exemple de la figure 1.

Main dominante : |_LIGNE_| |_JAUNE_| |_U_| |_3_|
Main dominée : |_LIGNE_|_____ --

Chaque production linguistique met en jeu un certain nombre de TI qu'il faut synchroniser, et chaque TI contient une partie de la signation qu'il faut spécifier.

Notons en outre que l'observation d'un corpus de vidéos montre que pour une construction linguistique donnée, tous les locuteurs ne synchronisent pas nécessairement les TI de manière rigoureusement identique. On remarque que le maintien de la configuration finale du signe LIGNE par la main dominée peut varier dans sa durée, mais que la synchronisation initiale des deux mains au début du signe LIGNE reste identique pour tous les signeurs.

À propos de cette variabilité et en vue de spécifier la structure linguistique, nous posons les trois objectifs suivants :

- toute variabilité dans la production n'entraînant pas de modification du sens doit rester possible (pas de sur-spécification) ;
- tout changement entraînant une modification du sens de l'énoncé fait l'objet d'un paramètre de la règle ;
- toute spécification valable quelle que soit le contexte et le signeur doit être fixé par la règle (on appelle ces éléments les invariants de la structure).

Pour traiter ce problème, nous proposons :

- le recours à des ensembles minimaux de contraintes (gestuelles et temporelles) suffisantes pour énoncer une règle sans contraindre trop la signation ;
- la possibilité pour les éléments de spécification de dépendre de variables contextuelles non fixées par la règle mais qui prendront une valeur selon leur utilisation.

3.3 Le modèle AZee

Le modèle Azee permet de représenter les contraintes nécessaires et suffisantes (CNS) de synchronisation et de réalisation d'un énoncé en LS. Il est composé de deux modèles, Zebedee et Azalee.

Zebedee est un langage de description qui implémente des CNS ainsi que des dépendances contextuelles pour donner aux séquences posture-transition ces mêmes propriétés. Il a été initialement conçu pour décrire les unités lexicales de la LS. Nous ne détaillons pas ce formalisme ici mais une page web lui est dédiée⁴.

Azalee est un formalisme capable de décrire tous types de synchronisation entre TI. En Azalee, les TI, généralement superposés, doivent être agencés sur la ligne de temps selon des contraintes temporelles à déterminer, puis chaque TI doit être spécifié, séparément, spécifiant ainsi la totalité de la structure.

Soit un ensemble de TI numérotés TI1, TI2, etc. concourant à une structure linguistique. Azalee décrit cette structure en un « azalisting », en les encapsulant comme suit :

```
[[
    règle de synchro, %% Liste des contraintes temporelles
    règle de synchro, %% nécessaires et suffisantes agencant
    règle de synchro... %% les TI sur l'axe temporel
|| TI1 :
    bloc de spécification
|| TI2 :
    bloc de spécification
|| ... :
    %% etc. (un bloc pour chaque TI apparaissant dans le bloc de synchro)
]]
```

À l'instar des CNS de Zebedee, les TI sont agencés sur l'axe temporel avec un ensemble minimal de contraintes temporelles nécessaires sur les bornes des intervalles (relations <, =, ≥...) ou sur les intervalles tout entiers (Allen, 1983). Ces contraintes apparaissent dans la première section de l'azalisting, encadrées par « [[» et le premier séparateur « || ». Nous en donnons quelques exemples ci-dessous :

- |gaze <|pt → le début du TI nommé « gaze » précède immédiatement le début du TI « pt » (pointage manuel) ;
- gaze| < pt| → la fin de « gaze » précède celle de « pt » ;
- pt|d| md → le TI « pt » est inclus dans « md » (p. ex. la tenue de la main dominée) – le « d » est pour « during » ;
- |A = B| ~ C| → « A » débute entre la fin de « B » et la fin de « C ».

Chaque TI doit ensuite être spécifié dans son propre bloc de spécification. Cela peut représenter un simple regard sur une cible, une suite de signes manuels, un geste des épaules, du buste, des sourcils ou une combinaison de ceux-ci. Formellement, cela peut représenter :

- une simple liste de contraintes qui sera à maintenir durant toute sa durée, en utilisant des contraintes articulatoires élémentaires – ces contraintes ciblent les articulateurs du corps comme un os du squelette ou un muscle du visage, éventuellement paramétrés par un numéro d'ordre (p.ex. la « n-ième » phalange

⁴<http://perso.limsi.fr/filhol/zebedee>

du doigt) et/ou par un côté du corps (gauche/droit ou dominant/dominé) ;

- une synchronisation de postures séparées par des transitions, en utilisant le langage Zebedee prévu à cet effet ;
- une structure temporellement plus complexe, à savoir un azalisting imbriqué, répartissant ainsi les TI contenus sur le morceau de l'axe temporel dédié au TI englobant.

Les deux formalismes Azalee et Zebedee, et ce faisant les deux stratégies de synchronisation, se combinent et permettent l'imbrication libre de structures réutilisables.

Voici un exemple complet d'azalisting pour une structure activant une zone de l'espace de signation par un pointage dont le schéma général est décrit ci-dessous :

- le regard précède toujours d'un temps très court le signe du pointage ;
- le regard et le pointage ciblent tous deux le même point de l'espace qui dépend de l'énoncé ;
- si le regard est maintenu, il ne dépasse jamais la rétractation du pointage manuel.

AZOP "activation + pointage de l'espace"

%% Ci-dessous, déclaration d'une dépendance au contexte

DEP spaceloc : Point %% représente l'emplacement activé

[[

 |gaze = <|pt, %% regard débute juste avant pointage

 gaze| < pt| %% regard termine avant la fin du pointage

 || gaze:

 LOOK at [spaceloc] %% regarder l'emplacement à activer

 || pt:

 SEQ "pointage" WITH %% ce même emplacement comme cible

 target = [spaceloc]

 END

]]

END

où "pointage" est une « zebedescription » définie par ailleurs avec une dépendance contextuelle nommée « target » qui représente la cible du pointage.

4 Conclusion et perspectives

Cet article a présenté une méthodologie visant, à partir d'une observation de corpus vidéo, à repérer puis formaliser les régularités de structure apparaissant entre le sens et la forme de constructions linguistiquement motivées. Cette méthodologie se veut le moins possible empreinte de courant linguistique a priori, et applicable à tous les niveaux du langage, du sub-lexical à l'énoncé complet. Le repérage vidéo s'organise autour de deux démarches inverses et complémentaires : rechercher une forme et généraliser sur le sens ou rechercher une valeur sémantique et en extraire les invariants

surfaciques. La formalisation de règles de production à partir de ces régularités observées est rendue possible grâce au modèle AZee, dont nous avons présenté les bases. Celui-ci permet une combinaison des deux stratégies de synchronisation qu'offrent les langages Zebedee et Azalee, respectivement la synchronisation par postures et celle par agencement contraint d'intervalles temporels sur une ligne de temps.

Si la partie Zebedee de ce modèle est déjà bien évaluée (2000 signes LSF décrits), nous n'avons pour l'instant exploré qu'une dizaine de structures. Les cinq heures de corpus DictaSign annotées⁵ devraient nous permettre de recueillir plus de ces structures, et plus d'occurrences pour chacune, et ainsi augmenter notre échantillon d'étude.

Nous comptons poursuivre ce travail pour multiplier le nombre de structures ainsi décrites et à terme pouvoir conclure sur la capacité d'AZee à couvrir l'ensemble des structures linguistiques mises en évidence. Notre hypothèse est qu'une fois combinées, ces règles pourront permettre de décrire des énoncés complets en LS. Ceci sera le point de départ d'une évaluation approfondie du modèle AZee et de sa mise en œuvre au sein d'applications dédiées à la génération et LS et à plus long terme à la traduction du texte vers la LS, en utilisant les entrées de ces règles comme sortie d'un système d'analyse textuelle. Ces règles pourront aussi faire suite au travail déjà entamé avec Zebedee (Gonzalez, 2012), pour étendre les travaux de reconnaissance des signes lexicaux aux structures plus larges des énoncés.

5 Références

- ALLEN, J. F. (1983). Maintaining Knowledge about Temporal Intervals. *In Communications of the ACM* 26:11, pp. 832–843, New-York, USA.
- ELLIOTT, R., GLAUERT, J., JENNINGS V., KENNAWAY, R. (2004). An overview of SiGML notation and SiGMLSigning software system. *In Proceedings of LREC 2004 (Language Resources and Evaluation Conference) workshop RPSL (Representation and Processing of Sign Languages) : From SignWriting to Image Processing. Information techniques and their implications for teaching, documentation and communication*, Lisbon, Portugal.
- FILHOL, M. (2011). Text-sign parallel corpus study to start designing an automatic translation system. *In proceedings of SLTAT workshop 2011 (Sign Language Translation and Avatar Technology)*, Dundee, Scotland.
- GONZALEZ, M., FILHOL, M., COLLET, C. (2012). *Semi-automatic Sign Language corpora annotation using lexical representations of signs*, Language Resource and Evaluation Conference, Istanbul.
- HUENERFAUTH, M. (2006). Generating American Sign Language classifier predicates for English-to-ASL machine translation. *PhD thesis*, University of Pennsylvania, USA.
- LIDDELL, S. K., JOHNSON, R. E. (1989). American Sign Language, the phonological base. *Sign Language studies* 64, Cambridge University press.
- MARSHALL, I., SÁFÁR, É. (2004). Sign Language Generation in an ALE HPSG . *In proceedings of HPSG-11*, Leuven, Belgique.

⁵<http://www.sign-lang.uni-hamburg.de/dicta-sign/portal/>