

Grammaire réursive non linéaire pour les langues des signes

Michael Filhol
LIMSI-CNRS, B. P. 133, 91403 Orsay cedex
michael.filhol@limsi.fr

Résumé. Cet article propose une approche pour la formalisation de grammaires pour les langues des signes, rendant compte de leurs particularités linguistiques. Comparable aux grammaires génératives en termes de récursivité productive, le système présente des propriétés nouvelles comme la multi-linéarité permettant la spécification simultanée des articulateurs. Basé sur l'analyse des liens entre formes produites/observées et fonctions linguistiques au sens large, on observe un décloisonnement des niveaux traditionnels de construction de la langue, inhérent à la méthodologie employée. Nous présentons un ensemble de règles trouvées suivant la démarche présentée et concluons avec une perspective intéressante en traduction automatique vers la langue des signes.

Abstract. This article presents a formal approach to Sign Language grammars, with the aim of capturing their specificities. The system is comparable to generative grammar in the sense that it is recursively productive, but it has quite different properties such as multilinearity, enabling simultaneous articulator specification. As it is based on the analysis of systematic links between observable form features and interpreted linguistic functions in the general sense, the traditionally separate linguistic levels all end up covered by the same model. We present the results found for a set of linguistic structures, following the presented methodology, and conclude with an interesting prospect in the field of text-to-Sign machine translation.

Mots-clés : Grammaire formelle, multi-linéarité, langue des signes.

Keywords: Formal grammar, multilinearity, Sign Language.

1 Introduction : langue des signes

La communauté s'est longtemps intéressée aux signes à un niveau lexical. Du côté de la linguistique : inventaires de signes pour l'élaboration de dictionnaires de vignettes (cf. fig. 1) ou représentation "paramétrique" de signes isolés (Stokoe, 1960), sorte de système phonologique selon lequel un signe est vu comme la combinaison de valeurs données à un ensemble prédéfini de paramètres (1 signe = 1 configuration + 1 orientation + 1 emplacement + 1 mouvement des mains). En TAL : conception de systèmes d'animation de signes placés en séquence (Hanke, 2002), par analogie assimilatrice aux séquences de mots des phrases écrites. À défaut de modèles plus adaptés, la langue des signes a peu été étudiée autrement que comme un alignement de signes manuels réputés lexicaux. Aujourd'hui, nous manquons de connaissances et (donc) de modèles formels capables de représenter l'organisation du discours signé.



FIGURE 1 – Vignette de dictionnaire LS pour le signe "parking" (Moody, 1997)

La difficulté est due à ce qu'en langue des signes, langue orale par excellence, beaucoup d'articulateurs simultanés contribuent à la construction de l'énoncé (sourcils, doigts, buste, ligne des épaules...), tous à des fréquences, niveaux et moments différents (cf. fig. 2). Aussi, la nature lexicale de certaines unités et leur compositionnalité phonologique ne

fait pas consensus chez les linguistes. La pensée assimilatrice et générativiste (Marshall & Sáfár, 2004) suppose que les niveaux linguistiques (phonétique, phonologique, morphologique, lexical, syntaxique, etc.) identifiés pour décrire les langues écrites et, déjà dans une moindre mesure, vocales, sont valides pour les LS sans quoi leur statut de langue pourrait être remis en cause. Au contraire, d'autres observent sur corpus de nombreux éléments pour lesquels il est impossible de trancher s'il s'agit d'éléments lexicaux ou de constructions syntaxiques, ou encore si un geste est linguistique ou "co-verbal", c'est-à-dire à considérer comme hors du système linguistique. Par exemple, l'expression de relations spatiales entre entités du discours se fait par un usage direct de l'espace de signation, respectant la topologie de la relation (Cuxac, 2000). Ce type de construction, fortement iconique¹ ne fait pas partie des lexiques car leur nombre est infini, mais n'est pas reléguable au non-verbal. Doit-on y voir une construction syntaxique alors qu'elle ne comporte pas d'éléments lexicaux identifiés et que tout est simultané ?



FIGURE 2 – LS : de multiples articulateurs simultanés

Aussi, l'emploi plus que récurrent de ces structures fortement iconiques en LS amène souvent à questionner les hypothèses traditionnellement admises liées à l'arbitraire du signe. Ces structures consistent en des unités non stabilisées, au sens où aucun inventaire ne les ferait apparaître comme des unités de construction, mais utilisant les propriétés du canal gestuel (trois dimensions, relations topologiques dans l'espace de signation, angles entre articulateurs et formes des mains) pour créer de nouvelles unités au fil du discours de manière extrêmement productive. Contrairement au mime, celles-ci sont pourtant bien linguistiquement régies et ne peuvent être improvisées. Si elles sont clairement efficaces pour les relations spatiales et la description de formes, l'utilisation symbolique de l'espace est tout aussi récurrente pour les concepts abstraits. L'iconicité fait donc véritablement partie intégrante de la langue et doit être selon nous prise en compte dans tout effort de modélisation des LS.

2 Une grammaire fonction–forme

Sans vouloir arbitrer le débat linguistique, nous avons ces dernières années proposé un modèle formel adapté à l'utilisation en informatique de ces langues naturelles, qui prend en compte leurs dimensions orale – même uniquement orale donc fondamentalement non linéaire, où plusieurs réalisations simultanées sont possibles au contraire de l'écrit – et iconique dans l'espace gestuel.

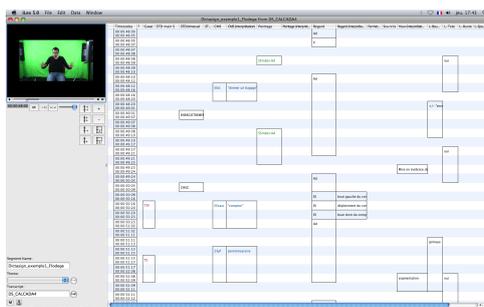
2.1 Forme, fonction et règles

Pour appréhender la LS sans influence de l'écrit, nous avons d'une part basé nos études sur des corpus de vidéos de langue attestée par des locuteurs natifs de LS, et fait le choix de revenir à une hypothèse linguistique de base, moins contraignante, selon laquelle la langue permet un lien entre :

- des formes produites, que l'on peut annoter dans un logiciel d'annotation en pistes (fig. 3), e.g. en LS la direction du regard, les mouvements des épaules, du buste, la durée d'une pause, les rotations de la tête, positions des mains, la flexion d'un doigt ;
- et leurs fonctions, interprétées à la réception du message, p. ex. dater un événement, séparer deux parties d'un discours, intensifier une fonction adverbiale, marquer l'objectivité d'une mesure...

L'idée est alors d'identifier, à partir de données réelles (corpus de langue), les liens systématiques qu'il peut exister, quitte à les paramétrer, entre forme et fonction, et de les formaliser. Toute fonction interprétée systématiquement pour un

1. L'*iconicité* est la propriété pour un signe, un signifiant ou toute forme linguistique produite de ressembler ou de rappeler le sens, le signifié, le concept associé.

FIGURE 3 – Annotation en pistes (ici verticales avec le logiciel *iLex*)

critère de forme donné constitue pour nous une *règle d'interprétation* de la forme observée. Inversement, la description des invariants de forme dans les occurrences d'une fonction donnée permet la formalisation d'une *règle de production* pour cette fonction. Dans notre but particulier de synthèse de LS, ce sont ces dernières qui vont constituer la base d'une grammaire formelle de génération propre à la langue des signes observée, permettant d'animer un personnage virtuel en synthétisant les formes correspondant aux fonctions qui seraient données.

2.2 Méthodologie pour la recherche de règles

Pour élaborer les règles de production, la méthodologie utilisée vise à identifier les liens entre fonction et forme en opérant un va-et-vient entre les deux, associant à chaque fois des descriptions de formes (resp. fonctions interprétées) aux occurrences d'un critère de fonction (resp. de forme) :

- dans le sens de l'*observation*, repérer toutes les occurrences d'un critère fonctionnel et en décrire les formes (articulations et précédences temporelles) ;
- dans le sens de l'*interprétation*, repérer toutes les occurrences d'un critère de forme et en décrire la fonction (interprétation de l'intention sémantique, au sens large, du locuteur une fois le message compris).

Dans les deux cas, il s'agit de dégager les éléments communs dans les descriptions élaborées, et procéder au sens inverse à partir du critère ainsi dégagé. Le va-et-vient entre observation et interprétation permet d'affiner les critères de fonction jusqu'à converger vers un invariant paramétré de forme à produire pour une fonction de la langue.

Par exemple, à partir du critère de forme illustré par la figure 4 (un pointage de l'index d'une main sur l'autre main en configuration chiffrée), nous avons pu dérouler la recherche suivante.

1. Recherche en interprétation, critère de forme défini ci-dessus (fig. 4) :
→ ici, toutes les occurrences faisaient partie d'une liste énumérée.
2. Recherche en observation, critère fonctionnel "énumération" :
→ ici, un grand nombre d'occurrences ont fait apparaître le même mouvement de tête vers l'en avant, sur chaque item énuméré.
3. Recherche en interprétation, critère de forme "mouvement de tête vers l'avant" :
→ ici, toutes les occurrences marquent des éléments non mutuellement exclusifs d'énumérations non-exhaustives.
4. Recherche en observation, critère fonctionnel "énumération non-exhaustive d'items non mutuellement exclusifs" – nous appellerons ces énumérations "listes ouvertes" :
→ retrouvant la même liste qu'à l'étape précédente, on propose de s'arrêter directement ici ; le critère fonctionnel "liste ouverte" donne ici déjà systématiquement lieu à la même forme, ce qui permet une règle de production.



FIGURE 4 – Illustration d'une forme initiale

Accessoirement, on note une utilisation totalement variable de la forme initiale avec pointage de l'index dans les énumérations qui l'utilisent, tout à fait optionnelle et sans effet. Le critère de départ a donc disparu dans le processus itératif. De manière générale, si un choix intuitif au départ peut accélérer la convergence vers une règle, celui-ci n'est pas décisif. Aussi, l'exemple ici montre une entrée par la forme, mais l'entrée par la fonction est également possible, le processus conduisant à l'alternance des points de vue.

3 Résultats récents et validation de l'approche

3.1 Système de règles temporelles

Suivant la méthodologie décrite en partant de fonctions temporelles (événements en séquence, duratifs, etc.), nous avons déjà pu établir un système cohérent de huit règles stables dont les fonctions sont (Filhol *et al.*, 2013) :

- (r1) Succession de deux événements séparés par une période de moins de 10 jours
Arguments : la durée *dur* (optionnelle), les événements *pre* et *post*
- (r2) Séquence chronologique de dates, périodes/durées et événements
Arguments : N éléments *item_i* dans l'ordre chronologique, $i \in [1..N]$
- (r3) Période de 10 jours minimum
Arguments : la durée *dur* de la période
- (r4) Un événement dure au moins 10 jours
Arguments : la durée *dur* et l'événement *event*
- (r5) Un événement dure moins de 10 jours
Arguments : la durée *dur* et l'événement *event*
- (r6) Événement daté
Arguments : la *date* et l'événement *event*
- (r7) Un événement s'écoule du temps de l'énonciation à une borne temporelle
Arguments : la date de fin *d2* et l'événement *event*
- (r8) Un événement s'écoule entre deux bornes temporelles dont la première est différente du temps de l'énonciation
Arguments : les bornes temporelles *d1* et *d2*, et l'événement *event*

Les formes associées sont ensuite décrites avec le formalisme "AZee", non développé ici mais présenté ailleurs (Filhol *et al.*, 2014). Ces formes peuvent être schématisées comme sur la figure 5, où :

- les rectangles délimitent sur l'axe horizontal les intervalles temporels durant lesquels la forme qu'ils contiennent est produite (l'unité de temps reste ici fictive mais des mesures par capture de mouvement nous permettront de préciser les spécifications sur cet axe) ;
- l'arrangement des rectangles sur l'axe vertical est arbitraire et n'est pas à interpréter comme les pistes d'une annotation où l'on trouverait un articulatoire annoté par ligne ;
- les éléments en italique sont les noms des arguments des règles, en d'autres termes les boîtes à remplir ;
- les articulations décrites peuvent être de tout type et viser tout articulatoire, ou faire appel à une autre règle, imbriquée, par exemple : 'hd: fwd' = tête déplacée en avant, 'eg: s-sp' = regard sur l'espace de signation, 'el: semi-cl' = paupières semi-fermées, 'hd: nod-up' = menton relevé, etc.

3.2 Discussion

Ce résultat est pour le moins intéressant, et ce pour trois raisons que nous présentons ci-dessous.

Compositionnalité récursive Nous avons produit un système de règles, chacune produisant la signation (forme) de la fonction portée, et pouvant comporter des dépendances (paramètres). Par exemple, la forme décrite par (r7) dépend de celle des éléments en italique, à savoir un événement et une date. Nous observons que ces signations peuvent elles-mêmes provenir d'une production obtenue par une entrée fonctionnelle. Par exemple, on trouve dans le corpus la phrase : *Dix ans après l'évacuation musclée de l'église Saint-Bernard, le 23 août 1996 à Paris, les sans papiers et leurs soutiens ne veulent pas être "dans la commémoration" mais dans le "combat", ...*

Celle-ci est traduite avec l'imbrication suivante, en parfaite cohérence avec les fonctions sémantiques des règles utilisées :

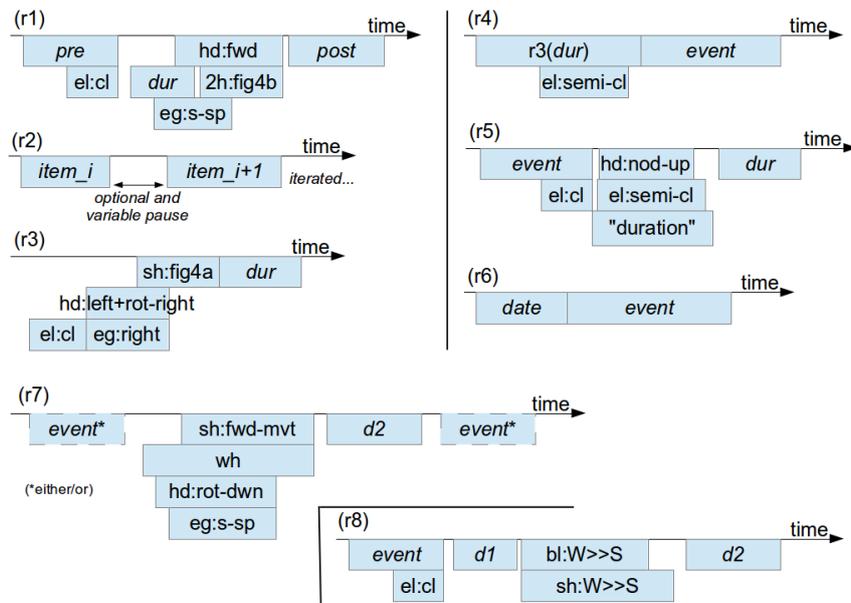


FIGURE 5 – Règles de production

```

r2(item_1 = r6(date = "23/08/1996",
              event = "évacuation musclée de l'église"),
    item_2 = r3(dur = "10 ans"),
    item_3 = ...)

```

Cette observation est importante car elle vérifie la propriété pour les règles d'être mutuellement récursives, condition sans laquelle le langage généré par la grammaire serait fini, donc la couverture du modèle nécessairement incomplète. Cette propriété est donc partagée avec les grammaires génératives, mais il reste une différence de taille avec celles-ci : l'ordre des feuilles ne représente absolument rien de la séquence finale du discours. En effet, celle-ci n'est pas vue comme une séquence, les règles pouvant mettre leur contenu en parallèle. Les sous-arbres ne sont finalement ordonnés que pour être nommés, et s'il s'agit de formes à signer, leur position relative temporelle dépend complètement de ce que dicte la règle.

Précision des règles Certaines fonctions peuvent paraître pour le moins étonnantes. Pour cause, notre méthodologie étudie la langue des signes à part et sans regard sur le texte, comme une langue nouvelle méritant ses formalismes propres. La distinction entre les périodes de plus et moins de 10 jours est un exemple de ce qu'aucun paradigme TAL syntaxico-sémantique n'aurait envisagé, et pourtant bien vérifié en corpus. Aussi bien dans le cas des durées d'événements que des durées de séparation entre événements consécutifs, on observe ce même seuil de durée. Notre approche rigoureusement basée sur les données nous semble éviter les biais de la "pensée textuelle" (linéarité, organisation syntaxique de mots prélistés...) et garantir que les règles de production trouvées soient bien les briques d'une "pensée signe".

Niveaux linguistiques des règles La vision traditionnelle en TAL, et jusqu'alors transposée en TALS, considère tout énoncé comme une séquence d'éléments lexicaux déterminés a priori, éventuellement modifiés par les règles syntaxiques gouvernant l'ordre de la séquence, et séparés par une période de transition plus ou moins courte. Cette vision a longtemps été portée par le courant générativiste et considérée comme universelle. Ici, les niveaux syntaxique, sémantique, etc. ne sont pas a priori délimités, et les opérations linguistiques – et donc les règles correspondantes – ne nécessitent pas d'être catégorisées en niveaux avant d'être décrites. Quels que soient leurs niveaux, un lien direct est créé entre ces fonctions et des articulations observables, assimilables en partie à un niveau phonétique, tout en imbriquant des structures de niveaux supérieurs. Notre approche décroïssonne donc l'ensemble des fonctions linguistiques décrites, remettant ainsi en question l'universalité du modèle à couches (niveaux empilés). La validation de ce modèle multi-linéaire et sans dominance syntaxique, en passant à l'échelle d'une couverture plus large de la langue, permettra de nourrir la réflexion sur les universaux du langage.

4 Conclusion et perspective en traduction

Nous avons présenté une approche et une méthodologie pour élaborer des grammaires formelles pour les langues des signes, basées sur corpus et sans hypothèse linguistique quant aux niveaux et fonctions modélisées. Cette approche permet d'écrire les éléments de forme invariants, quitte à les paramétrer, pour des fonctions linguistiques identifiées, chacune donnant lieu à une règle de production directement animable par un signeur virtuel par exemple.

Une perspective intéressante d'utilisation du modèle existe selon nous en traduction automatique de texte vers LS. L'imbrication des règles représentant l'énoncé à signer est donnée en entrée du système de génération peut servir directement de représentation fonctionnelle du discours à traduire, puisque les règles contenues dans celle-ci portent, par construction, chacune une fonction interprétée (lexicale, sémantique, à but rhétorique, etc.). Avec une vision similaire à celle employée côté cible (liens forme–fonction), nous proposons de rechercher les formes du côté source ayant la même fonction. En d'autres termes et avec un point de vue TAL, pour chaque règle de production connue en langue des signes, il s'agit de repérer dans le texte les formes (en français écrit cette fois) pouvant être interprétées de la manière donnée par la fonction de la règle. Lorsqu'une forme textuelle est ainsi repérée, la règle peut alors être déclenchée pour une utilisation dans la traduction à proposer en sortie. Par exemple, la règle "liste ouverte", dont la fonction est l'énumération non-exhaustive d'items non mutuellement exclusifs, peut être déclenchée dès lors que l'on observe dans le texte une suite d'éléments séparés par des virgules et de même type syntaxique, et introduite par "comme" ou "par exemple".

Le problème global, illustré sur la figure 6, devient donc un ensemble de sous-problèmes TAL de recherche d'information, chacun prévu pour déclencher une règle en utilisant tous les moyens possibles (apprentissage, règles à base de motifs...), et ses arguments autant que possible. Les règles étant déclenchées séparément, elles doivent ensuite être combinées (nuage gris) pour former la sortie du système de traduction (cadre bleu), servant du même coup d'entrée du système de synthèse (cadre jaune). Nous démarrons aujourd'hui ce travail exploratoire et espérons en présenter des résultats prochainement.

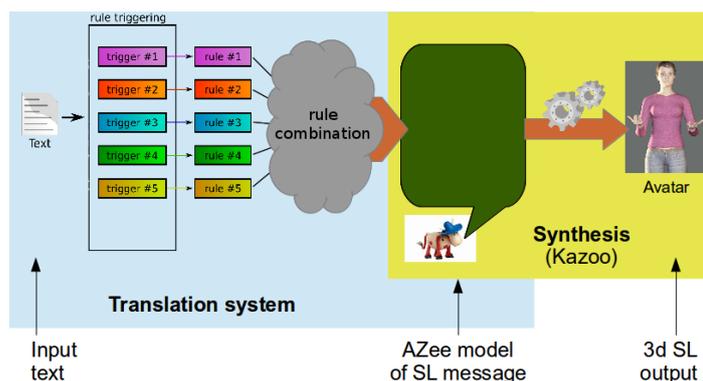


FIGURE 6 – Architecture pour une traduction automatique à base de déclenchement de règles

Références

- CUXAC C. (2000). *Langue des signes française, les voies de l'iconicité*, volume 15–16. Paris Ophrys.
- FILHOL M., HADJADJ M. N. & CHOISIER A. (2014). Non-manual features : the right to indifference. In *Language resource and evaluation conference (LREC), 6th workshop on the representation and processing of Sign Language : beyond the manual channel*, Reykjavik, Islande.
- FILHOL M., HADJADJ M. N. & TESTU B. (2013). A rule triggering system for automatic text-to-sign-translation. In *Sign Language translation and avatar technology (SLTAT)*, Chicago, USA.
- HANKE T. (2002). Visicast deliverable d5-1 : interface definitions. ViSiCAST project report.
- MARSHALL I. & SÁFÁR E. (2004). Sign language generation in an ale hpsg. In *Proceedings of HPSG04*.
- MOODY B. (1997). *La langue des signes, dictionnaire bilingue LSF/français*. Paris : IVT.
- STOKOE W. (1960). Sign language structure : an outline of the visual communication system of the american deaf. *Studies in Linguistics, Occasional Papers*, 8.