

Modélisation de l'espace discursif pour l'analyse de la langue des signes

Boris Lenseigne et Patrice Dalle
IRIT-TCI - Université Toulouse 3
118 route de Narbonne 31062 Toulouse cedex 4
lenseign@irit.fr
dalle@irit.fr

Mots-clefs : Espace de signation, monologue, dialogue, analyse de corpus vidéo

Keywords: Signing space, monolog, dialog, video corpora analysis

Résumé Cet article présente un modèle de la structure des énoncés en langue de signes (LS) qui s'articule autour de la notion d'espace de signation. Nous présentons dans un premier temps ce modèle tel qu'il a été conçu pour l'analyse de séquences d'images dans un contexte réduit. Puis nous faisons le bilan des éléments manquants pour étendre ce contexte. Dans une seconde partie, nous présentons une extension de ce modèle aux situations de dialogue et continuons la discussion sur les éléments qui doivent y apparaître dans ce second cas. Cette discussion ouvre la voie à l'introduction de modèles additionnels pour pouvoir analyser la LS dans son ensemble.

Abstract We propose a model for the structure of sign language utterances (SL) that is based on the signing space notion. We first present this model as it was designed for the analysis of reduced-context utterances. After that, we discuss about the elements that lack to generalize that model. In the second part, we present the extension of the model to the representations of dialogs and we continue the discussion about the elements that are missing in the second case. This discussion opens a wide file for the study of additionnal models that will help the complete analysis of the SL.

1 Introduction

Les recherches linguistiques sur la LS ont connu ces dernières années un essor important qui a abouti à des descriptions de cette langue suffisamment précises et complètes pour pouvoir envisager leur intégration dans des systèmes informatiques. Notre recherche s'appuie fortement sur ces résultats et vise à proposer des modèles informatiques permettant de rendre compte de la structure des énoncés en LS et de la relier à la façon dont ceux-ci sont réalisés. Ces travaux se situent dans le cadre de la conception de systèmes d'analyse d'images dédiés et constituent une approche originale de cette langue qui trouve des applications dans de nombreux autres domaines (animation d'avatars, études linguistiques, langages gestuels, ...).

En effet, dans la grande majorité des systèmes existants, qu'il s'agisse d'interpréter la LS ou de faire signer un personnage de synthèse, l'énoncé est considéré comme une succession de signes isolés, éventuellement coarticulés. En interprétation, les connaissances sur la structure de l'énoncé concernent alors l'ordre dans lequel les signes sont produits. Il peut s'agir de connaissances statistiques (Hienz *et al.*, 1999) (Liang & Ouhyoung, 1996) ou de contraintes sur la structure des énoncés (Starnier & Pentland, 1995). La structure spatiale des énoncés reste, en revanche, rarement prise en compte : les systèmes de reconnaissance se limitent le plus souvent à l'interprétation des verbes directionnels (Sagawa *et al.*, 1997). L'interprétation d'énoncés spatialisés ne reste possible qu'à l'aide d'hypothèses fortes sur leur structure (Braffort, 1996). En synthèse, la structure spatiale des énoncés apparaît de façon implicite dans (Lebourque, 1998) par l'introduction de la notion de cible pour un mouvement. Elle est en revanche explicitement représentée dans le système proposé dans (Huenerfauth, 2004). Pour l'heure toutefois, aucun modèle formel de l'espace de signation n'a été proposé et la question de son exploitation pour le traitement automatique de la LS a été peu abordée.

Nous présenterons, ici, un tel modèle élaboré initialement dans un cas mono-locuteur. Nous en commenterons les limites actuelles et présenterons son extension aux situations de dialogue.

2 Modélisation de l'espace discursif dans le cas mono-locuteur

2.1 Modèle symbolique de l'espace de signation

Ce modèle est basé sur la théorie de l'iconicité présentée dans (Cuxac, 2000). Il représente l'agencement des différentes entités de l'énoncé et les relations sémantiques qui les relient. Le type de chaque entité est défini à partir du type des relations susceptibles de la relier aux autres.

Une première version du modèle a été proposée dans (Lenseigne, 2004) pour l'interprétation d'énoncés de requêtes. Il s'agissait donc de ne représenter qu'un sous-ensemble de la LS. Nous abordons l'analyse de l'énoncé par le biais de la construction de l'espace de signation. On accède alors au sens de l'énoncé en considérant les entités évoquées et leur fonction dans l'énoncé à partir de leur agencement dans cet espace. Dans le contexte de requêtes, on distingue quatre types de fonctions : les *localisations temporelles absolues* ou *relatives* au temps courant de l'énoncé, les *localisations spatiales* des différentes entités dans l'espace de signation et la fonction *d'actance* qui peut, le cas échéant, impliquer plusieurs entités. A ce niveau, ne sont représentés que les éléments du discours susceptibles d'être référencés par le signeur et d'être impliqués dans les relations décrites ci-dessus. Ces éléments sont instanciés via une modélisation du comportement qui permet de déduire le type de l'entité à partir des attitudes, mouvements, gestes et expressions du locuteur. D'autres niveaux syntaxiques et lexicaux seront nécessaires pour construire le sens de l'énoncé.

2.2 Représentation interne et externe du modèle

Du point de vue informatique, l'espace de signation (classe `EspaceSignation`¹) est représenté comme un volume cubique divisé régulièrement en `Emplacements`. Chaque `Emplacement` peut contenir zéro, une ou plusieurs `Entités`. Réciproquement, une `Entité` appartient à

¹Les termes écrits dans une police "machine à écrire" sont des éléments du modèle

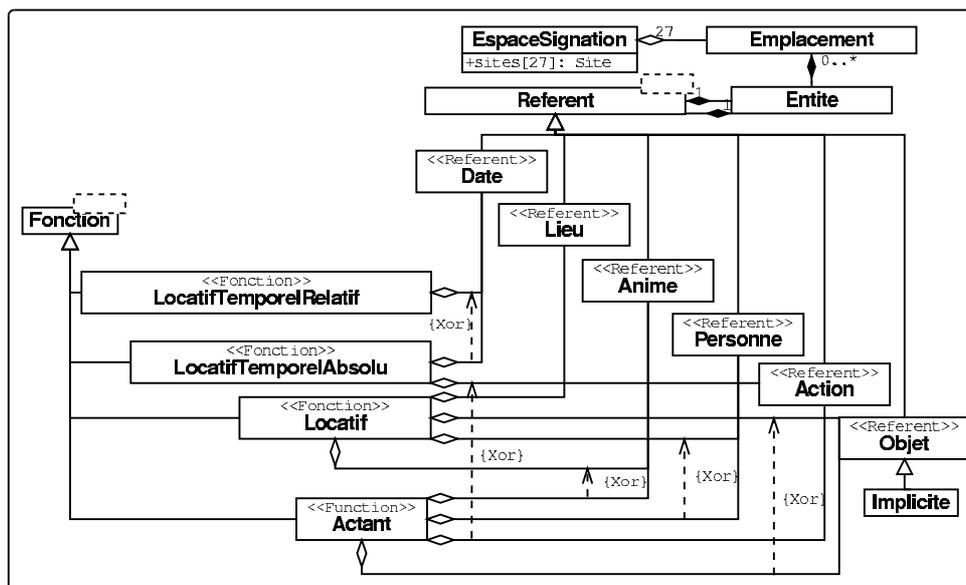


Figure 1: Diagramme de classes UML de la représentation symbolique de l'espace de signation.

un ou plusieurs Emplacements et possède un type qui est un Référent. Les différents types d'Entité permettent de contraindre les relations susceptibles de les relier, c'est-à-dire les différentes Fonctions que l'entité pourra prendre dans l'énoncé (locatif, actant, ...). L'architecture générale du modèle est donnée sous forme de diagramme de classes UML (fig. 1). Cette structure de données est en outre dotée de mécanismes permettant d'assurer la cohérence de la représentation lors de la création d'une Entité de type donné. Un outil interactif a été réalisé pour construire et visualiser l'utilisation de l'espace de signation (fig. 2a). Notre objectif est d'automatiser cette construction par analyse d'image.

2.3 Discussion à propos du modèle

Ce modèle ne permet, actuellement, de représenter que les relations temporelles, spatiales et d'actance. Ceci est insuffisant pour représenter l'ensemble des références potentielles

Ainsi les notions qui viennent qualifier une entité n'apparaissent pas initialement dans l'espace de signation. Pourtant, elles peuvent être réifiées et devenir alors référençables. D'autre part ce modèle n'inclut ni les relations de composition, c'est-à-dire le fait de pouvoir référencer l'élément d'un ensemble ni celles de généralisation/spécialisation. Enfin il ne permet pas les références aux localisations spatiales dans la mesure où ces dernières sont représentées implicitement dans l'agencement des Entités dans l'espace de signation.

3 Extension aux situations de dialogue

La LS est presque toujours utilisée en situation de dialogue; les locuteurs partagent alors le même espace discursif et les entités évoquées par un locuteur peuvent être référencées indifféremment par chacun des locuteurs. Cette version du modèle permet de prendre en compte ces interactions.

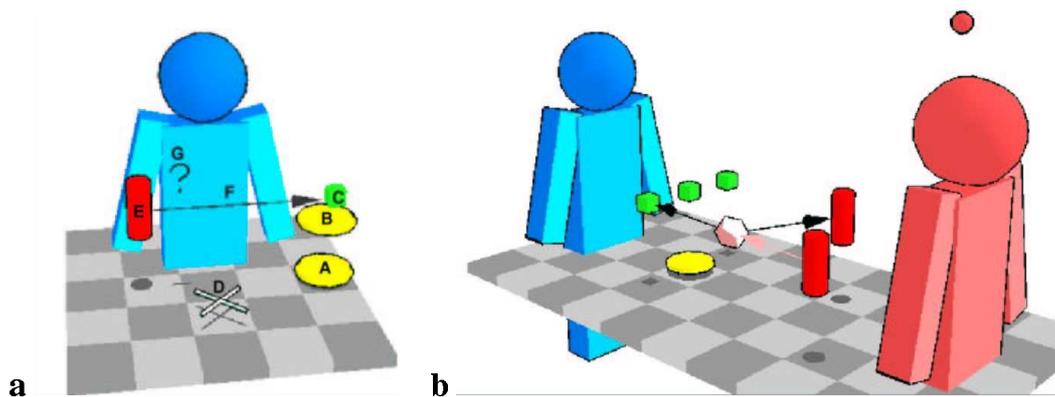


Figure 2: **(a)** Exemple de construction de l'espace de signation visualisé sous forme d'une scène 3D pour la phrase (traduite signe à mot): à Toulouse (A), au cinéma Utopia (B), le film qui passe (C), jeudi 26 février à 9h30 (D), la personne (E), qui a fait le film (F), qui-est-ce ? (G). **(b)** Exemple de construction de l'espace de signation en situation de dialogue : (locuteur à gauche) Sur le site, il y a un espace pour le public, un pour l'administration et un pour les salariés . (locuteur à droite) Les salariés peuvent consulter leur espace et ça ne regarde pas le public (extrait du corpus TALS : CE1.mov).

3.1 Modélisation de l'espace discursif en situation de dialogue

Les mécanismes qui régissent la construction de l'espace de signation en situation de dialogue sont identiques à ceux utilisés dans les monologues à la seule différence qu'une entité peut être évoquée par l'un ou l'autre des locuteurs. La représentation interne d'une Entité intègre cette information. En revanche la vérification de la cohérence de la construction de l'espace de signation se fait de façon globale. Nous avons également étendu les possibilités de l'outil interactif pour faire apparaître les deux locuteurs et leur prise de parole respectives. (fig. 2b).

3.2 Discussion

L'utilisation de la LS en situation de dialogue fait apparaître de nouveaux éléments qui ne sont actuellement pas pris en compte dans le modèle. Ces éléments concernent le contenu du discours lorsque les échanges sont au niveau "méta-linguistique", c'est-à-dire lorsque les locuteurs parlent de ce qui est dit, voire de la façon dont c'est dit, ou les éléments de "régulation" (gestes phatiques,...). Du point de vue de la réalisation de l'énoncé, le modèle ne prend pas en compte les mécanismes d'"alias" permettant la reprise, par le second locuteur, d'une entité évoquée par le premier, afin d'éviter une mauvaise interprétation d'un geste déictique.

4 Conclusion

Le modèle que nous avons proposé fournit une description, à un niveau très général, des énoncés en LS, qui peut être exploitée dans de nombreux domaines d'application : en interprétation en génération d'énoncés par exemple.

Son utilisation pour la transcription de corpus vidéo a mis en avant la nécessité de compléter cette représentation pour prendre en compte de nouveaux types de références. Des modèles

supplémentaires doivent être élaborés afin de pouvoir prendre en compte le lexique et la réalisation de l'énoncé. Sur le plan syntaxique, nous avons déjà proposé, pour l'analyse automatique de vidéos, un modèle de comportement permettant de décrire la construction de l'espace de signation en termes de séquences de gestes (Lenseigne, 2004). Pour l'introduction du lexique, nous utilisons, en attendant les résultats de recherches en cours sur les formes graphiques de la LS², le formalisme SignWriting³.

Dans sa version actuelle, le modèle proposé n'en constitue pas moins un outils très intéressant qui ouvre de nombreux questionnements tant sur la linguistique de la LS que sur sa modélisation informatique. Son exploitation en tant qu'outil interactif s'est en outre avérée prometteuse tant pour analyser la LS que pour l'expliquer, voire l'enseigner.

Références

BRAFFORT A. (1996). *Reconnaissance et compréhension de geste, application à la langues des signes*. PhD thesis, Université Paris XI, UFR Sciences, LIMSI.

CUXAC C. (2000). *La langue des Signes française. Les voies de l'iconicité*. ISBN 2-7080-0952-4. Paris: Faits de langue, Ophrys.

HIENZ H., BAUER B. & KRAISS K. (1999). Hmm-based continuous sign language recognition using stochastic grammars. In S. BERLIN, Ed., *Lecture Notes in Artificial Intelligence : Procs 3rd Gesture Workshop'99 on Gesture and Sign-Language in Human-Computer Interaction*, p. 165–184, Gif-sur-Yvette, France: A. Braffort, R. Gherbi, S. Gibet, J. Richardson, D. Teil.

HUENERFAUTH M. (2004). Spatial representation of classifier predicates for machine translation into american sign language. In *Workshop on Representation and Processing of Sign Language, 4th International Conference on Language Ressources and Evaluation (LREC 2004)*, p. 24–31, Lisbon Portugal.

LEBOURQUE T. (1998). *Spécification et génération des gestes naturels; Application à la LSF*. PhD thesis, Université Paris XI-SUD, UFR Sciences, LIMSI.

LENSEIGNE B. (2004). *Intégration de connaissances linguistiques dans un système de vision. Application à l'étude de la langue des Signes*. PhD thesis, Université Paul Sabatier, Toulouse 3.

LIANG R. & OUHYOUNG M. (1996). A sign language recognition system using hidden markov model and context sensitive search. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, p. 59–66, Hongkong.

SAGAWA H., TAKEUCHI M. & OHKI M. (1997). Description and recognition methods for sign language based on gesture components. In *in Proceedings of IUI*, Orlando, Florida.

STARNER T. & PENTLAND A. (1995). *Real-Time American Sign Language Recognition From Video Using Hidden Markov Models*. Rapport interne TR-375, M.I.T Media Laboratory Perceptual Computing Section.

²<http://lsscript.limsi.fr/>

³<http://www.signwriting.org>