

Stage de master « éditeur DV-LS »

Contexte

La recherche proposée se situe en traitement automatique de la langue (TAL) appliqué aux langues des signes (LS), ou TALS. En particulier, elle s'inscrit dans un objectif de génération automatique de la LS par des signeurs virtuels.

Chacune langue à part entière, les LS sont des langues orales et non écrites, de modalité gestuelle et non phonatoire. Cette différence de canal et ses conséquences fortes sur la structuration des énoncés empêchent de plaquer les modèles TAL traditionnels sur les problèmes du TALS. Par exemple, l'utilisation productive de l'espace tridimensionnel de signation permet des constructions iconiques et concises non trivialement assimilables à des notions connues en français écrit.

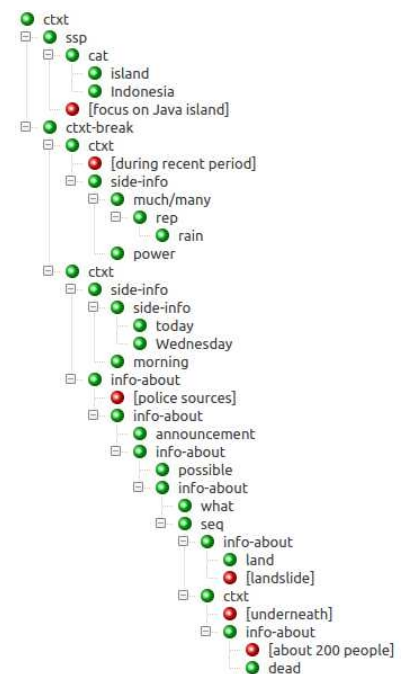
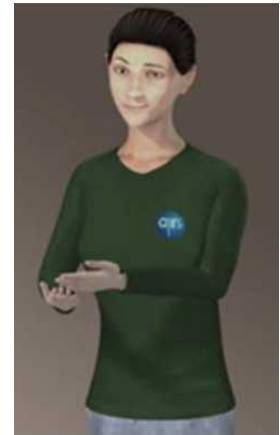
Le modèle AZee [1 ; 2 ; 3] développé au LIMSI permet, à partir d'une représentation formelle arborescente, la synthèse directe d'énoncés en LS avec toutes les articulations et synchronisations requises. Formellement, les arbres AZee sont comparables à des arbres syntaxiques en cela qu'ils représentent une structure profonde et compositionnelle des messages signés, mais leurs nœuds représentent des fonctions sémantiques interprétables plutôt que des catégories syntaxiques apparaissant comme arbitraires pour les LS. AZee se trouve donc proche d'une représentation sémantique des énoncés.

Par ailleurs, les traducteurs voulant appréhender le sens d'un texte et le restituer en LS sans l'influence du français pratiquent la « déverbalisation » au moyen de schémas représentant les entités du discours et les liens entre elles [4 ; 5 ; cf. annexe]. Certains sourds natifs de la LS pratiquent aussi ce type de schémas (DV) pour éviter le français qui linéarise et se distance de la pensée visuelle qui leur est naturelle. Aucun standard ne pèse pour l'instant sur la construction des DV, mais vu leur nature éminemment sémantique et spatial, nous postulons la possibilité d'un passage automatique de la DV vers AZee.

Objectifs

L'objectif du stage est de développer l'architecture et l'interface (par exemple pour navigateur web) d'un éditeur graphique pour composer des DV, en procédant en deux temps.

La première étape revient à fournir une version graphique des arbres AZee, c'est-à-dire des dessins dans le plan dont les primitives ne sont ni plus ni moins que des correspondances avec les nœuds élémentaires utilisés dans les arbres (expressions) en AZee. À la manière du "Qt designer", on doit pouvoir agencer des icônes et combiner ou imbriquer des contenants équivalents à des expressions AZee.



La seconde étape visera à l'abstraction du modèle AZee en permettant de spécifier des éléments graphiques correspondant à des structures de plus haut niveau, c'est-à-dire à des sous-arbres constitués de plusieurs nœuds. En effet, des structures-type apparaissent régulièrement en AZee, portant des fonctions sémantiques identifiées que les auteurs de DV mobilisent en tant que telles, et font apparaître directement sur les schémas.

Le défi de conception est que l'éditeur doit permettre un ensemble évolutif de nœuds et structures. Cet ensemble devra donc être spécifié à part et chargé à l'exécution.

Retombées sur le métier et la LS

Le but à terme est d'intégrer le travail à la conception de logiciels de TAO (traduction assistée par ordinateur) pour la LS. En effet, là où les traducteurs de texte à texte voient leur tâche facilitée par un grand nombre d'outils de traitement (édition, exploration et recherche multi-documents, concordanciers, mémoires de traduction...), la langue des signes ne dispose d'aucun équivalent. Pourtant, la reconnaissance récente et croissante de la place des LS pour l'accessibilité dans la société crée une demande croissante à laquelle il faut répondre en équipant ces langues d'outils aujourd'hui inexistantes.

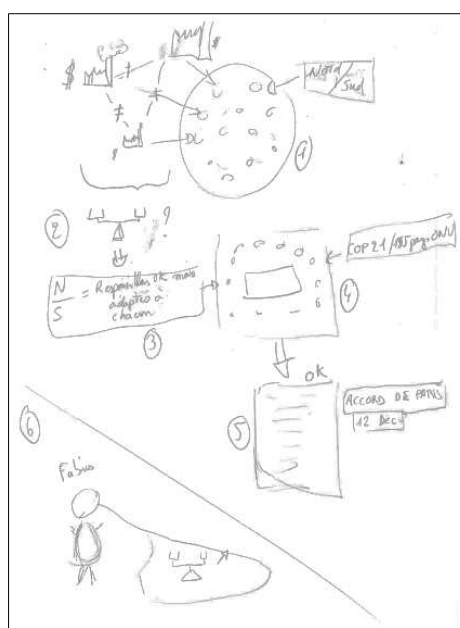
Aussi, grâce au défi posé par ces langues particulières, le projet devrait faire progresser le TAL en général (hors LS) sur la piste de la traduction par le sens.

Annexe : exemple de DV

Texte original en français

« Le meilleur équilibre possible. » Ainsi le président de la COP 21, Laurent Fabius, a-t-il qualifié l'« accord de Paris » – c'est son nom officiel – adopté, samedi 12 décembre, par les 195 pays membres de la Convention-cadre des Nations unies sur les changements climatiques. Il s'agit, de fait, d'un compromis porté par le principe de la « justice climatique ». C'est-à-dire par la reconnaissance que les pays du Nord et ceux du Sud ont des « responsabilités communes mais différenciées » dans le changement climatique, et que leurs « capacités respectives » à y faire face sont inégales.

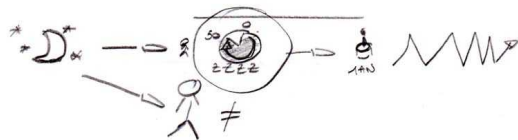
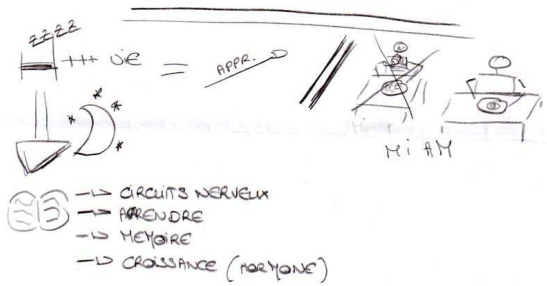
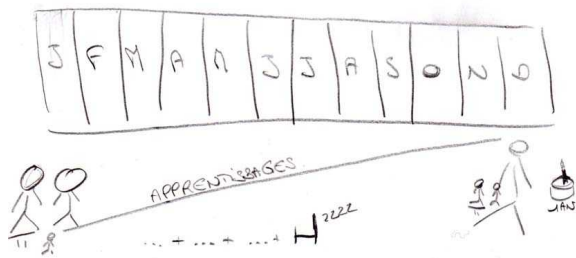
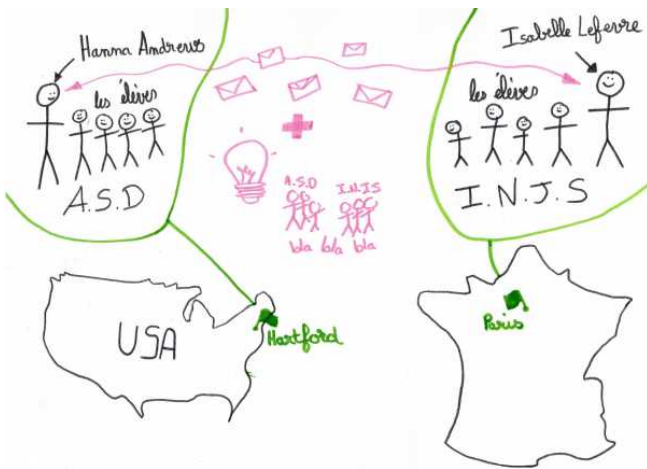
Schéma de déverbalisation



Lecture :

1. Dans le monde, les pays du Nord et du Sud produisent des effets sur le climat, mais n'ont pas tous les mêmes ressources économiques.
2. Comment faire pour un juste équilibre entre tous ?
3. Justice climatique = une responsabilité partagée mais un financement adapté à chaque pays
4. COP 21 (195 pays de l'ONU) s'est tenue
5. Accord de Paris le 12 décembre
6. M. Fabius le qualifie d'équilibré

Autres exemples :



Références

- [1] M. Filhol, M. N. Hadjadj, A. Choisier (2014), *Non-manual features: the right to indifference*, Representation and Processing of Sign Languages: beyond the manual channel, Language resource and evaluation conference (LREC), Iceland.
- [2] M. Filhol (2014), *Grammaire réursive non linéaire pour les langues des signes*, Traitement automatique de la langue naturelle (TALN), Marseille, France.
- [3] M. Filhol, M. N. Hadjadj, B. Testu (2015), *A rule triggering system for automatic text-to-Sign translation*, in Universal access in the information society (UAIS), vol. 15/4, C. Stephanidis (Ed.), special issue on "Recent Advances in Sign Language Translation and Avatar Technology", R. Wolfe, E. Efthimiou, J. Glauert, T. Hanke, J. McDonald, J. Schnepf (guest eds.), Springer.
- [4] P. Guitteny (2007), *Langue des signes et schémas*, dans Traitement automatique de la langue, TAL vol. 48:3, pp. 205–229.