

## Cours & Atelier

**Bruno Gaume**

IRIT-UPS, Toulouse F-31062 Cedex 4, France

[gaume@irit.fr](mailto:gaume@irit.fr)

**Résumé du cours :** *Traitements computationnels et sémantiques du lexique, application à l'acquisition du langage*

### 1) Les systèmes de représentation pour les structures et les relations lexicales

La modélisation de la dynamique d'acquisition du sens des mots en L1 (par le jeune enfant pour sa langue maternelle), ou L2 (par l'apprenant pour une langue seconde), s'inscrit nécessairement dans un système de représentation qualitatif et quantitatif des principales structures et relations lexicales : polysémie, synonymie, hyperonymie/hyponymie, méronymie, métaphore, champ sémantique, ...) permettant de comparer les modèles à l'expérimentation. En effet, la variabilité des réponses des sujets rend difficile l'évaluation des modèles face à l'expérimentation (pour un même dispositif expérimental les réponses des différents sujets sont souvent multiples et variées, et les comparer aux réponses du modèle à évaluer nécessite souvent de qualifier la relation lexicale ou/et de mesurer la distance inter-items entre les diverses réponses : sujets/sujets et sujets/modèle). Dans ce cadre nous présenterons deux types de représentations : les représentations à bases de vecteurs et celles à base de réseaux.

### 2) La collecte des informations lexicales

Une fois le système de représentation défini formellement il faut alors collecter les informations lexicales afin de les représenter dans le système choisi. Dans ce cadre nous présenterons deux types de collectes : celles à base de corpus et celles à base de dictionnaires.

### 3) La modélisations de quelques structures et relations lexicales

Nous présenterons ensuite trois approches et quelques unes de leurs applications pour la modélisation lexicale :

- Prox <http://prox.irit.fr/>
- LSA <http://lsa.colorado.edu/>
- Hyperlex <http://www.up.univ-mrs.fr/~veronis/demos/index.html?http://www.up.univ-mrs.fr/~veronis/demos/hyperlex.html>

### 4) Application à l'acquisition du langage

Nous aurons alors tout en main pour une présentation de quelques applications à la modélisation de la dynamique d'acquisition du lexique avec une focalisation sur le traitement opéré par un Prox, un modèle de la proximité sémantique inter-items à partir de données issues d'expérimentations.

**Résumé de l'Atelier :** *Lexique & Sémantique (Prox : <http://prox.irit.fr/>)*

L'atelier commencera par un exposé de 30 minutes sur la structure des réseaux lexicaux qui sont des réseaux petits mondes hiérarchiques et une rapide présentation de l'analyseur syntaxique de corpus SYNTAX <http://www.univ-tlse2.fr/erss/textes/pagespersos/bourigault/index.html> ainsi que la base lexicale distributionnelle du français « LES VOISINS DE LE MONDE » <http://www.univ-tlse2.fr/erss/voisinsdelemonde/framesetMain.html>.

Nous étudierons ensuite avec Prox la structure de plusieurs réseaux lexicaux (verbes, noms, adjectifs) issus de dictionnaires de synonymes ou bien de dictionnaires de langues : le Trésor de la Langue Française informatisé (TLFi) pour ensuite faire quelques manipulations de démétaphorisations automatisées : « Alice **déshabille** l'arbre » : effeuiller, élaguer, déraciner, tailler, ... pour enfin analyser, avec Prox, les réponses d'enfants décrivant des actions sur films vidéo : « elle casse le livre »...

## Références :

- Bourigault D., Fabre C., Frérot C., Jacques M.-P. & Ozdowska S. (2005), Syntex, analyseur syntaxique de corpus, in *Actes des 12èmes journées sur le Traitement Automatique des Langues Naturelles*, Dourdan, France.  
<http://www.univ-tlse2.fr/erss/textes/pagespersos/bourigault/taln05-bourigault-frerot.pdf>
- Denhière G., Lemaire B., A (2004) Computational Model of Children's Semantic Memory, in Eds. *Proceedings 26th Annual Meeting of the Cognitive Science Society*, pages pp. 297-302, Chicago  
[http://cogprints.org/3777/01/cogsci04\\_1.pdf](http://cogprints.org/3777/01/cogsci04_1.pdf)
- Duvignau, K., Gardes-Tamine, J., Gaume, B. (2004) Proximité sémantique et métaphore verbale chez l'enfant. *Le langage et l'homme*, EME, Belgique.
- Gaume B., (2004) Balades Aléatoires dans les Petits Mondes Lexicaux, In *I3 Information Interaction Intelligence* vol.4 - n°2 - 2004, CEPADUES édition.  
<http://www.revue-i3.org/volume04/numero02/index.htm>
- Véronis, J. (2003). Hyperlex : cartographie lexicale pour la recherche d'informations. *Actes de la Conférence Traitement Automatique des Langues (TALN'2003)* (pp. 265-274). Batz-sur-mer (France): ATALA.  
<http://www.up.univ-mrs.fr/veronis/pdf/2003-taln.pdf>