

## ***Formes et fonctions de l'énumération en discours : de l'oral spontané à l'écrit planifié***

**Journées d'étude internationales S'caladis  
organisées par Josette Rebeyrolle et Gilles Corminboeuf**

**Université de Toulouse 2 Jean Jaurès, 19-20 novembre 2015**

### ***Mise en forme et indices linguistiques de surface pour l'extraction de connaissances***

Jean-Philippe Fauconnier (Institut de Recherche en Informatique de Toulouse)

Ce travail s'inscrit dans le cadre de la construction et de la population de ressources sémantiques à partir de textes de type encyclopédique ou scientifique (Aussenac-Gilles et Jacques 2008). L'originalité de notre travail réside dans l'extraction de relations paradigmatiques lexicales (Murphy 2003) exprimées au-delà de la linéarité du texte (Kamel et al. 2012). Pour cela, nous nous appuyons sur la sémantique véhiculée par les indices typographiques et dispositionnels qui ont pour fonction de suppléer des formulations strictement linguistiques qui seraient plus difficilement exploitables.

L'étude que nous proposons ici concerne les structures énumératives verticales (Luc 2001; Ho-Dac et al. 2012; Rebeyrolle et Péry-Woodley 2014) qui, bien qu'affichant des discontinuités entre leurs différents composants, présentent un tout sur le plan sémantique. Ces structures textuelles sont souvent révélatrices de relations hiérarchiques (Sumida et al. 2008). Notre méthode montre que l'exploitation de la mise en forme, en combinaison avec un faisceau d'indices linguistiques de surface, facilite l'identification de ces relations. Deux points sont abordés : la caractérisation de la nature de la relation et l'extraction de ses arguments.

#### **Références**

- Aussenac-Gilles, N. et Jacques, M.-P. (2008). Designing and evaluating patterns for relation acquisition from texts with caméléon. *Terminology*, 14:45–73.
- Ho-Dac, L., Fabre, C., Péry-Woodley, M.-P., Rebeyrolle, J. et Tanguy, L. (2012). An empirical approach to the signalling of enumerative structures. *Discours. Revue de linguistique, psycholinguistique et informatique*, (10).
- Kamel, M., Mojahid, M. et Rothenburger, B. (2012). "Quand rédiger c'est décrire" - mise en forme matérielle des textes et construction d'ontologies à partir de textes. In *23es Journées Francophones d'Ingénierie des Connaissances*, pages 133 – 148.
- Luc, C. (2001). Une typologie des énumérations basée sur les structures rhétoriques et architecturales du texte. In *Actes de la 8e Conférence sur le Traitement Automatique des Langues Naturelles (TALN 2001)*, pages 263–272.
- Murphy, M. L. (2003). *Semantic relations and the lexicon : Antonymy, synonymy and other paradigms*. Cambridge University Press.
- Rebeyrolle, J. et Péry-Woodley, M.-P. (2014). Énumération et structuration discursive. In *SHS Web of Conferences*, volume 8, pages 3183–3196. EDP Sciences.
- Sumida, A., Yoshinaga, N. et Torisawa, K. (2008). Boosting precision and recall of hyponymy relation acquisition from hierarchical layouts in wikipedia. In *Proceedings of the 6th International Conference on Language Resources and Evaluation*.