

**Observatoire de linguistique Sens-Texte**

**OLST**

**Rapport d'activités**

**mai 2007 – avril 2008**

## **Table des matières**

1.	Membres réguliers et associés .....	3
1.1	Membres réguliers .....	3
1.2	Membres associés .....	3
2.	Projets et état d'avancement des travaux.....	3
2.1	Dictionnaires et bases de données lexicales.....	3
2.2	Analyses de données lexicales.....	4
2.3	Modélisation de données lexicales.....	5
2.4	Corpus.....	6
2.5	Traitement automatique de la langue.....	7
2.6	Autres.....	8
3.	Encadrement d'étudiants aux études supérieures .....	8
4.	Personnel.....	9
4.1	Étudiants .....	9
4.2	Stagiaires.....	11
5.	Autres activités .....	12
5.1	Séminaires, colloques .....	12
5.2	Professeurs, chercheurs invités.....	12
5.3	Collaborations.....	12
6	Publications.....	13
6.1	Livres ou chapitres de livres.....	13
6.2	Articles dans revues avec comité de lecture .....	14
6.3	Articles dans actes de congrès .....	15
6.4	Communications non publiées.....	17
6.5	Autres.....	18

## **1. Membres réguliers et associés**

### **1.1 Membres réguliers**

Jeanne Dancette (jusqu'en septembre 2007)  
Patrick Drouin  
Lidjia Iordanskaja  
Pascale Lefrançois (depuis mai 2007)  
Marie-Claude L'Homme (directrice)  
Elizabeth Marshman (depuis juillet 2007)  
Igor Mel'čuk  
Jasmina Milićević  
Etienne Nguyen (depuis décembre 2007)  
Alain Polguère (directeur adjoint)

### **1.2 Membres associés**

Jeanne Dancette (entre octobre 2007 et mars 2008)  
Sylvie Vandaele (jusqu'en mai 2008)

## **2. Projets et état d'avancement des travaux**

Cette section décrit les projets spécifiques sur lesquels ont travaillé les chercheurs de l'OLST durant la période s'étendant de mai 2007 à avril 2008.

### **2.1 Dictionnaires et bases de données lexicales**

*DiCo, base de données lexicale des dérivations sémantiques et collocations du français* (Responsables : Igor Mel'čuk et Alain Polguère). Le travail d'extension de la base DiCo s'est poursuivi. Nous prévoyons d'injecter durant l'été 2008 la base de locutions de Sara-Anne Leblanc (voir ci-dessous) à l'intérieur du DiCo. Les vocables de statut 1 seront aussi réinjectés en statut 0, une fois leur étiquetage sémantique validé. Un nouvel accès aux données du DiCo en ligne a été réalisé par Sébastien Cabot (programmeur) et Alain Polguère (responsable de projet) : le DiCoPop. Il s'agit d'une interface conviviale, qui permet notamment la génération automatique d'articles de dictionnaire de type LAF à partir des données du DiCo (tables SQL).

*Dictionnaire (anglais-français-espagnol) de la mondialisation du travail* (responsable : Jeanne Dancette). Les articles de nature encyclopédique sont rédigés dans les trois langues. Des tableaux de mots reliés par classe de relations sémantiques

permettent un rapprochement entre les termes (travail en cours). La rédaction de quelque 200 articles de fond est bien avancée. Environ 2000 autres fiches sont en cours de rédaction. Ces autres termes sont des termes reliés par une relation sémantique aux 200 termes correspondant aux notions de base. Ils font l'objet d'un traitement plus court (définition, équivalents et relations sémantiques). La recherche de la cohérence maximale entre mots reliés est facilitée par l'usage d'une base de données relationnelle.

***Dictionnaire fondamental de l'informatique et de l'Internet, DiCoInfo*** (responsable : Marie-Claude L'Homme). En 2007-2008, le travail suivant a été fait sur le dictionnaire : rédaction d'articles, mise en ligne d'articles dont la rédaction est très avancée (plus de 1000), enrichissement de la liste de termes reliés (plus de 20 000 liens sont décrits).

***Mise au point d'une base de connaissances bilingues en biologie cellulaire et moléculaire*** (responsable : Sylvie Vandaele). Il s'agit de la prolongation d'un projet déjà entrepris. Le travail a consisté à améliorer la structure relationnelle d'une base de connaissances spécialisées, à trier les données déjà accumulées et à corriger les entrées existantes (environ 500 entrées pour chacune des deux langues envisagées, l'anglais et le français). Le projet est terminé et la mise en ligne est en cours..

***Version multilingue du DiCoInfo*** (responsable : Marie-Claude L'Homme) : Des versions anglaise, coréenne et espagnole du DiCoInfo sont en cours de rédaction. En anglais, près de 600 acceptions ont été distinguées et plus de 4000 liens sémantiques répertoriés. En outre, les liens d'équivalence avec le DiCoInfo français sont établis pour environ 900 lexies. En coréen, (responsables : Hee Sook Bae et Marie-Claude L'Homme), environ 100 acceptions ont été distinguées et près de 200 liens sémantiques répertoriés. En outre, les liens d'équivalence avec le DiCoInfo français sont établis pour les 100 lexies décrites. En espagnol (responsables : Amparo Alcina et Marie-Claude L'Homme), la sélection des termes est amorcée par l'équipe de Castellón. Un séminaire de formation donné à Castellón a permis de commencer la rédaction de quelques articles.

## **2.2 Analyses de données lexicales**

***Analyse de collocations dans le domaine médical*** (responsable : Sylvie Vandaele). L'objectif du projet est de repérer et d'analyser les collocations importantes au sein des corpus de médecine. Outre l'étiquetage de corpus, on a utilisé un système d'extraction fondé sur un étiquetage syntaxique (Syntex; système mis au point par Didier Bourigault, Toulouse – MDu, MDa). Le projet est terminé et la correction des fiches avant mise en ligne est en cours. Un volet du projet sur les collocations a consisté à analyser le caractère métaphorique de l'expression de relations d'associations dans le domaine biomédical (avec Elizabeth Marshman).

***Étude comparative des marqueurs de relation causale dans un corpus de bitextes de vulgarisation*** (responsable : Elizabeth Marshman). Cette étude s'inscrit dans un axe de recherche qui consiste à repérer et à analyser des contextes contenant une relation causale. Nous visons à évaluer la variation interlinguistique dans les occurrences de divers éléments de contextes qui expriment des relations causales (les marqueurs de la relation, des expressions de certitude ou incertitude, etc.), afin d'évaluer l'impact de ces différences sur l'identification, l'analyse et l'utilisation de ces contextes riches en connaissances par des applications semi-automatiques. Nous comparons les résultats observés dans des bitextes à ceux obtenus dans une analyse de corpus comparables. Dans le cadre d'un projet pilote, un échantillon de contextes pour 12 verbes analysés antérieurement a été faite. Le projet continuera avec l'analyse d'autres données et éléments des contextes. Un élargissement vers d'autres relations est également prévu.

***Étude comparative des marqueurs de relation causale en corpus spécialisés*** (responsables Marie-Claude L'Homme et Elizabeth Marshman). Cette étude s'inscrit dans un axe de recherche qui consiste à repérer les contextes contenant une relation causale. Nous modélisons les marqueurs verbaux de la relation causale en tenant compte de leur structure actancielle et de la classe sémantique des actants. Nous cherchons à déterminer si l'appel aux classes sémantiques des arguments d'un marqueur ambigu peut permettre de le désambiguïser. Une liste d'environ 35 marqueurs français a été constituée à partir d'un corpus médical, d'un corpus d'informatique et d'un corpus de psychologie; des contextes sont extraits de corpus. L'analyse des contextes contenant un marqueur dans les corpus de médecine, d'informatique et de psychologie est terminée.

***Étude du lexique scientifique transdisciplinaire*** (responsable Patrick Drouin). Cette étude s'inscrit dans un axe de recherche portant sur le lexique commun à divers domaines scientifiques. Le lexique scientifique transdisciplinaire (LST) fait l'objet de descriptions lexicographiques en anglais et en français. L'objectif de ces descriptions est double : identifier les sens des unités lexicales recensés en contexte et décrire les collocations au sein desquelles ces unités interviennent.

### **2.3 Modélisation de données lexicales**

***Antonymie et terminologie*** (responsables : Marie-Claude L'Homme et Carles Tebé). L'objectif de ce projet consiste à mener une réflexion sur l'antonymie envisagée du point de vue de la terminologie. Des paires de termes définis comme étant contraires dans le DiCoInfo sont recherchés dans un corpus d'informatique afin d'identifier des marqueurs linguistiques de cette relation sémantique. Enfin, une modélisation plus systématique des paires de termes contraires sera réalisée dans le DiCoInfo.

***Conversion au format XML du DiCoInfo*** (responsables : Guy Lapalme et Marie-Claude L'Homme). Le DiCoInfo est désormais encodé et diffusé sous format XML. Un travail de conversion a permis d'exporter l'ensemble des descriptions du format Access

au format XML. Désormais, les terminologues éditent les articles au moyen de l'éditeur oXygen.

***Étude des adjectifs dans le domaine de l'environnement*** (responsable : Marie-Claude L'Homme) : Ce projet consistait à étendre le travail de description des adjectifs spécialisés dans le domaine de la médecine (réalisé dans le cadre d'une subvention CRSH) au domaine de l'environnement. Les adjectifs ont été extraits au moyen de TermoStat et la description d'environ 50 acceptions adjectivales a été réalisée. Les descriptions sont encodées dans une base de données Access.

***Modélisation lexicographique des locutions*** (responsables : Sara-Anne Leblanc et Alain Polguère). Continuation du travail décrit dans les rapports précédents, avec ajout d'une seconde base de données de structures syntaxiques de dépendance des constituants internes des locutions, en relation avec la base de données lexicale d'origine. Maintenant que beaucoup de données syntaxiques et grammaticales ont été prises en compte, nous prévoyons pour 2008-2009 développer l'aspect sémantique de la description des locutions, ce qui permettra notamment une meilleure représentation des variations observables avec ces unités lexicales.

***Notions fondamentales de lexicologie, lexicographie et terminologie*** (responsable : Alain Polguère et Ophélie Tremblay). Continuation du travail décrit dans les rapports précédents. Développement de l'ontologie des concepts de lexicologie/lexicographie se poursuit avec la collaboration d'Ophélie Tremblay — étudiante de doctorat du Département de didactique. Nous prévoyons pour fin 2008 une première mise en ligne de cette ontologie, appelée par euphémisme GTN (*Gros Tas de Notions*). Le GTN est utilisé comme base de connaissances dans le cadre du projet Lexitation (ontologisation des savoirs lexicographiques) d'Alain Polguère, financé par le CRSH.

***Système lexical du DiCo*** (responsable : Alain Polguère) : Ce travail est temporairement suspendu du fait du départ de l'étudiant de doctorat qui devait y participer.

## **2.4 Corpus**

***Constitution de corpus de thèses en français et en anglais*** (responsable : Patrick Drouin). Ces corpus seront à la base de travaux sur le lexique scientifique transdisciplinaire. Nous avons procédé à l'élaboration de deux corpus de thèse de 2 millions de mots. Les thèses recensées dans le corpus proviennent des domaines suivants : archéologie, chimie, droit, géographie, histoire, informatique, ingénierie, physique, psychologie. L'ensemble des documents ont été recueillis dans Internet et convertis dans un format normalisé.

***Constitution d'un corpus dans le domaine du développement durable*** (responsable : Marie-Claude L'Homme). Ce projet a mené à la constitution d'un corpus de textes de

langue française (env. 500 000 occurrences). Le corpus renferme des textes scientifiques et de vulgarisation.

***Constitution d'un corpus de bitextes de vulgarisation dans le domaine de la santé*** (responsable : Elizabeth Marshman). Ce projet vise la constitution d'un corpus de bitextes alignés anglais-français. Le corpus renferme des textes de vulgarisation tirés de sites Internet d'organismes canadiens dans le domaine de la santé. Le corpus compte actuellement environ 250 000 mots dans chaque langue, mais continue d'être alimenté.

***Constitution d'un corpus étalon annoté pour l'acquisition automatique de termes*** (responsable : Patrick Drouin). Ce projet a pour objectif l'élaboration d'un corpus étalon pouvant être utilisé pour évaluer les performances des logiciels d'acquisition automatique de termes. L'ensemble des documents qui composent le corpus relève du domaine de l'automobile et totalise approximativement 225 000 mots. La terminologie contenue dans le corpus est annotée manuellement, ce qui rend possible l'utilisation des mesures de précision et de rappel et la comparaison des résultats obtenus par les outils électroniques au jugement humain.

***Corpus Joseph Zobel*** (responsable : Patrick Drouin). Ce projet visait la mise en place d'un corpus couvrant la totalité de l'œuvre en prose de l'écrivain martiniquais Joseph Zobel. La lexicographie différentielle francophone a beaucoup progressé ces dernières années, entre autres grâce à des banques textuelles telles que Frantext, Suistext ou Québétext ; or, le français régional des Antilles est resté jusqu'à maintenant un parent pauvre de cette discipline. Le corpus a été numérisé, annoté en format XML et analysé afin d'en faire ressortir les spécificités lexicales.

***Enrichissement du corpus d'informatique*** (responsable : Marie-Claude L'Homme) : Ce projet avait pour objectif d'ajouter des textes d'informatique au corpus d'informatique français et anglais. Le corpus compte désormais environ 2 000 000 mots. Le corpus anglais, quant à lui, compte environ 1 000 000 mots.

***Corpus de biomédecine en vue de l'analyse de collocations*** (responsable : Sylvie Vandaele). Ce volet visait d'une part, à finaliser la préparation des corpus non annotés et annotés pour la mise en ligne et, d'autre part, à montrer un petit corpus (25 000 mots) très ciblé (traitant de notions complexes) permettant de compléter les projets de constitution de base de connaissance en biologie cellulaire et moléculaire, et d'analyse de collocations. Le corpus est terminé et la mise en ligne est en cours.

## **2.5 Traitement automatique de la langue**

***Lexique transdisciplinaire*** (responsable : Patrick Drouin). L'objectif du projet est de repérer automatiquement le lexique scientifique transdisciplinaire (LST) au sein d'un corpus de thèses et d'articles scientifiques reliés à des domaines variés. Afin d'atteindre cet objectif, diverses méthodes de traitement automatique de la langue mettant faisant

appel à des traitements linguistiques et statistiques sont mises de l'avant.

## 2.6 Autres

*Étude de ressources contenant des termes d'informatique* (responsable : Marie-Claude L'Homme) : Ce projet consiste à étudier le traitement d'un ensemble de termes d'informatique (extraits avec TermoStat) dans six ressources en ligne. Les ressources étudiées sont des dictionnaires terminologique, un dictionnaire de nature générale et une ontologie. Les éléments étudiés sont les suivants : présence des termes dans les ressources (tels quels ou dans une autre forme), sens répertoriés, liens sémantiques établis dans les articles.

## 3. Encadrement d'étudiants aux études supérieures

**Dominic Anctil** (3<sup>e</sup> cycle), *État des lieux sur l'enseignement du lexique au secondaire : description des problèmes lexicaux des élèves et des représentations des enseignants par rapport à l'enseignement du lexique et à l'erreur lexicale* (codirection : Pascale Lefrançois et Françoise Boch; en cours).

**Sahara Iveth Carreño Cruz** (3<sup>e</sup> cycle, traduction). *Méthodes de regroupement de variantes terminologiques en extraction automatique de termes : application à des corpus spécialisés anglais, espagnol et français* (co-direction : Marie-Claude L'Homme et Patrick Drouin; en cours).

**Mathilde Cloutier** (2<sup>e</sup> cycle, traduction) *Localisation du logiciel Crimson Editor* (Marie-Claude L'Homme; en cours).

**Angela Gira** (2<sup>e</sup> cycle, traduction). *Glossaire terminologique dans le domaine de l'environnement* (Marie-Claude L'Homme; en cours).

**Sylvie Girard** (2<sup>e</sup> cycle, M. Ed.), *Élaboration d'une séquence didactique pour l'apprentissage du vocabulaire en espagnol L2 ayant recours à la L1 des élèves* (Pascale Lefrançois ; en cours).

**Fadila Hadouche** (3<sup>e</sup> cycle, informatique). *Annotation automatique de rôles actanciels* (co-direction : Guy Lapalme et Marie-Claude L'Homme; en cours).

**Daniel Alberto Labonia** (3<sup>e</sup> cycle, linguistique). *Modélisation sémantique des ressources linguistiques en extraction automatique d'information* (Marie-Claude L'Homme; en cours).

**Annaïch Le Serrec** (2<sup>e</sup> cycle, traduction). *Étude sur l'équivalence de termes extraits automatiquement d'un corpus parallèle : contribution à l'extraction terminologique bilingue* (co-direction : Marie-Claude L'Homme et Patrick Drouin; 2008).



**Karine Pouliot** (2<sup>e</sup> cycle, M.A.), *Élaboration d'un vocabulaire d'apprentissage du français de la recherche* (codirection : Pascale Lefrançois et Monique Cormier; en cours).

**Caroline Poudrier** (2<sup>e</sup> cycle, linguistique). Titre à préciser (co-direction : Patrick Drouin et Alain Polguère; en cours).

**Virginia Raveau-O'Kelley** (2<sup>e</sup> cycle, traduction). *Localisation du site web Equip kids international* (Patrick Drouin; en cours).

**Paola Sardo** (2<sup>e</sup> cycle, traduction). *Description terminographique de termes juridiques* (Marie-Claude L'Homme; 2008).

**Charlotte Tellier** (2<sup>e</sup> cycle, traduction). *Description terminographique de verbes médicaux* (Marie-Claude L'Homme; en cours).

**Ophélie Tremblay** (3<sup>e</sup> cycle), Création d'une ontologie des connaissances métalexicales pour l'élaboration d'un module de cours en didactique du lexique destiné aux futurs maîtres au primaire en français langue maternelle (codirection : Pascale Lefrançois et Jacqueline Bourdeau; en cours).

**Wajdi Zaghouni** (2<sup>e</sup> cycle, linguistique) *L'extraction automatique des entités nommées, une étude comparative de l'arabe et du français* (Patrick Drouin; en cours).

**Hui Zhou** (2<sup>e</sup> cycle, traduction). *Analyse de localisation de sites Web consacrés aux Jeux Olympiques 2008* (Patrick Drouin; en cours).

## **4. Personnel**

### **4.1 Étudiants**

**Sandrine Astié** (baccalauréat en traduction) : constitution d'un corpus nécessaire pour compléter les travaux portant sur les collocations et la base de connaissance en biologie cellulaire et moléculaire.

**Samia Bouchaddakh** (doctorat en linguistique) : application des principes de la lexicologie explicative et combinatoire à la construction d'un dictionnaire bilingue informatisé français-arabe.

**Sylvie Boudreau** (doctorat en traduction) : mise au point des formats de conversion des bases de données en vue de l'affichage en ligne; mise au point de l'affichage des données en ligne.

**Laurence Brunet** (baccalauréat en traduction) : constitution et annotation d'un corpus étalon pour l'acquisition automatique de termes dans le domaine de l'automobile ;

constitution d'un corpus bilingue de thèses en vue du repérage automatique du vocabulaire scientifique transdisciplinaire; descriptions lexicographiques du vocabulaire scientifique transdisciplinaire..

**Stéphanie Caron** (baccalauréat en traduction): enrichissement du corpus d'informatique français

**Iveth Carreño Cruz** (doctorat en traduction) : version espagnole du DiCoInfo.

**Isabelle Carrière** (doctorat en traduction) : mise en ligne de corpus spécialisés; étude des adjectifs dans le domaine de l'environnement.

**Marianne Champagne** (maîtrise en traduction) : recherche documentaire et rédaction d'articles du dictionnaire de la mondialisation

**Rana Charabati** (maîtrise en traduction; diplôme en gestion internationale de projets; HEC) : gestion de la banque de données du dictionnaire de la mondialisation

**Sophie Comeau** (maîtrise en linguistique) : Relations de fonctions lexicales en enseignement du français langue seconde.

**Louis-Philippe Dargis** (baccalauréat en traduction) : version anglaise du DiCoInfo; enrichissement du corpus d'informatique anglais.

**Rita Hanitramalala** (doctorat en linguistique) : étude des adjectifs dans le domaine de l'environnement.

**Josh Holden** (doctorat en linguistique) : étude du lexique du déné suliné.

**Anne-Laure Jousse** (doctorat en linguistique) : modèle de structuration dynamique d'un lexique dans une perspective de transformation et de consultation.

**Mélanie Labelle** (maîtrise en traduction) : préparation et nettoyage des différents corpus à mettre en ligne.

**Thi Hoa Lê** (doctorat en didactique) : Enseignement des expressions figées métaphoriques aux apprenants vietnamiens.

**Sara-Anne Leblanc** (doctorat en linguistique) : modélisation lexicographique des unités lexicales phraséologiques (voir le projet *Modélisation lexicographique des locutions* à la section 2.3).

**Mireille Lépine** (maîtrise en traduction) : analyse des données obtenues automatiquement pour l'analyse des collocations.

**Amélie Paulus** (baccalauréat en traduction) : constitution d'un corpus bilingue d'articles scientifiques en vue du repérage automatique du vocabulaire scientifique transdisciplinaire.

**Caroline Poudrier** (maîtrise en linguistique) : évaluation de deux méthodes d'identification des liens lexicaux syntagmatiques (méthodes manuelle et statistique).

**Gaëlle Rabatel** (maîtrise en traduction) : étude de ressources électroniques contenant des termes d'informatique.

**Mariana Raffo** (maîtrise en traduction/doctorat à partir de janvier 2008) : analyse et nettoyage des données terminologiques de la base de connaissances de biologie cellulaire et moléculaire.

**Sophie Reid-Triantafyllos** (baccalauréat en traduction) : constitution et annotation d'un corpus étalon pour l'acquisition automatique de termes dans le domaine de l'automobile ; constitution d'un corpus bilingue de thèse en vue du repérage automatique du vocabulaire scientifique transdisciplinaire ; descriptions lexicographiques du vocabulaire scientifique transdisciplinaire.

**Melissa Samson-Fauteux** (maîtrise en linguistique) : structuration formelle des définitions analytiques (application à la lexicographie informatisée du français).

**Victoria Surtees** (DESS en traduction) : révision des articles en langue anglaise du dictionnaire de la mondialisation du travail; analyse de contextes contenant des marqueurs de relation cause-effet dans un corpus de psychologie.

**Corinne Thirion-Dupuy** (maîtrise en traduction) : site Web de l'OLST.

**Ophélie Tremblay** (doctorat en linguistique) : mise au point de descriptions dans le champ sémantique des sentiments, recherche sur l'interfaçage des modélisations formelles dans un contexte d'enseignement du lexique.

## 4.2 Stagiaires

**Hee Sook Bae**, Korterm, Corée (stagiaire post-doctoral) : DiCoInfo coréen.

**Andrei Philip**, Université autonome de Barcelone, Espagne (stagiaire de doctorat), modélisation de données lexicales en langues romanes.

**Aurélien Picton**, Université Toulouse-le-Mirail, France (stagiaire de doctorat) : étude de l'évolution des termes dans un corpus d'aéronautique.

**Vincent St-Amour**, Université de Montréal (stagiaire en informatique, baccalauréat, collaboration avec le groupe RALI) : version XML du DiCoInfo.

**Carles Tebé**, Université Pompeu Fabra, Barcelone, Espagne (stagiaire post-doctoral) : antonymie en terminologie.

## **5. Autres activités**

### **5.1 Séminaires, colloques**

*Séminaires OLST-RALI* : Organisation de séminaires hebdomadaires en collaboration avec le RALI (Recherche appliquée en linguistique informatique, Département d'informatique et de recherche opérationnelle).

### **5.2 Professeurs, chercheurs invités**

**Jacquelinet Christian** (Agence de la biomédecine, France), novembre 2007.

### **5.3 Collaborations**

Convention d'échanges qui nous liait à l'Université Stendhal (Grenoble) (renouvellement), sous forme d'une convention tripartite incluant maintenant l'Université de Cologne.

Équipe de recherche en syntaxe et sémantique (Didier Bourigault) : utilisation et évaluation de Syntex, un analyseur syntaxique.

Équipe Tecnolettra / Université Jaume I (Amparo Alcina) : version espagnole du DiCoInfo.

Équipe TexMex, Irisa / Université de Rennes (Vincent Claveau et Pascale Sébillot) : Détection automatique de relations paradigmatiques.

Guy Lapalme, RALI (Recherche appliquée en linguistique informatique), Université de Montréal : conversion au format XML du DiCoInfo.

Laboratoire LICEF de la Télé-Université à Montréal (Jacqueline Bourdeau) : Projet *Lexitation* d'ontologisation des savoirs lexicographiques en vue de leur utilisation dans les disciplines de linguistique appliquée.

Marie-Josée Hamel et Jasmina Milicevic (Dalhousie University) : Projet de dictionnaire électronique de reformulation pour apprenants du français langue seconde.

## 6 Publications

### 6.1 Livres ou chapitres de livres

- Dancette, J. (à paraître). Questions sociolinguistiques et terminologiques de la mondialisation du travail. *Actes. Lexicographie et terminologie : histoire de mots. Colloque en l'honneur d'Henri Béjoint*. Université Lyon-2, France.
- Jousse, A.-L., A. Polguère & O. Tremblay. (2008). Du dictionnaire au site lexical pour l'enseignement/apprentissage du vocabulaire. In F. Grossmann & S. Plane (dir.). *Lexique et production verbale. Vers une meilleure intégration des apprentissages lexicaux*, Éducation et didactiques, Villeneuve d'Ascq : Presses Universitaires du Septentrionp., pp. 141-157.
- Iordanskaja, L. & I. Mel'čuk (à paraître). Connotation in Linguistic Semantics. In T. Berger, K. Gutschmidt, S. Kempgen & P. Kosta (dir.) : *The Slavic Languages: An International Handbook of their History, their Structure and their Investigation*, Berlin—New York : de Gruyter.
- L'Homme, M.C. (2008, à paraître). Large Terminological Databases. *International Encyclopedia of Lexicography*.
- L'Homme, M.C. & S. Vandaele (éd.) (2007). *Lexicographie et terminologie : compatibilité des modèles et des méthodes*, Ottawa : Presses de l'Université d'Ottawa.
- L'Homme, M.C. & S. Vandaele (2007). Introduction. Lexicographie et terminologie : disciplines sœurs ou pratiques distinctes ? In L'Homme, M.C. & S. Vandaele (éd.). *Lexicographie et terminologie : compatibilité des modèles et des méthodes*, Ottawa : Presses de l'Université d'Ottawa, pp. 1-26.
- Lefrançois, P. (2008). À l'école des utilisateurs de dictionnaires : une analyse de quelques dictionnaires scolaires québécois. In M.C. Cormier & J.-C. Boulanger (dir.), *Les dictionnaires de la langue française au Québec – De la Nouvelle-France à aujourd'hui*. Montréal : Presses de l'Université de Montréal, pp. 309-340.
- Marshman, E. (soumis) “Challenges of Lexical Knowledge Patterns for Bilingual Terminology Work.” In L. Bowker (ed), *University of Ottawa School of Translation and Interpretation 35th Anniversary Volume*.
- Mel'čuk, I. (2007). Lexical Functions. In H. Burger, D. Dobrovolskij, P. Kühn & N. Norrick (dir.) : *Phraseology. An International Handbook of Contemporary Research*, Berlin—New York : de Gruyter, pp. 119-131.
- Mel'čuk, I. (2008). Semantic Transition Rules (of the Semantic Module of a Meaning-Text Linguistic Model). In K. Gerdes, T. Reuther & L. Wanner (dir.) : *Meaning-Text Theory 2007 [WSA, Sonderband 69]*, Munich—Vienne: WSA.

- Mel'čuk, I. (à paraître). Zero Affixes and Nominal Cases in Daghestanian Languages. In O. Fëdorova (dir.) : *Festschrift S. Kodzasov*, Moscou.
- Mel'čuk, I. & A. Polguère. (2007). *Lexique actif du français. L'apprentissage du vocabulaire fondé sur 20 000 dérivations sémantiques et collocations du français*, Champs linguistiques, Bruxelles : De Boeck & Larcier, Bruxelles, 525 pages.
- Mel'čuk, I. & A. Polguère. (à paraître). *Syntactic Dependency in Linguistic Description*, Language Companion Series, Amsterdam/Philadelphia: John Benjamins.
- Polguère, A. (2007). Soleil insoutenable et chaleur de plomb : le statut linguistique des greffes collocationnelles. In M.-C. L'Homme & S. Vandaele (dir.) : *Lexicographie et terminologie : compatibilité des modèles et des méthodes*, Ottawa : Les Presses de l'Université d'Ottawa, pp. 247-291.
- Vandaele, S. & M. Raffo (à paraître) Signification lexicales et notionnelles dans les domaines de spécialité. *Actes. Lexicographie et terminologie : histoire de mots. Colloque en l'honneur d'Henri Béjoint*. Université Lyon-2, France.

## 6.2 Articles dans revues avec comité de lecture

- Charron, A. & O. Tremblay. (2007). Pour enseigner les expressions comme un poisson dans l'eau. *Vivre le primaire*, 20 (4).
- Dancette, J. (2007). Semantic Relations in Retailing. *Terminology* 13(2), pp. 201-223.
- Dancette, J. (2007) La mondialisation du travail : des pratiques sociales à la terminologie et de la terminologie à l'usage. *Revue internationale sur le travail et la société* 5(2), pp. 64-83, 2007.  
[http://www.uqtr.ca/revue\\_travail/Articles/2007Vol5Num2pp64-83Dancette.pdf](http://www.uqtr.ca/revue_travail/Articles/2007Vol5Num2pp64-83Dancette.pdf).
- Drouin, P. (2007). Identification automatique du lexique scientifique transdisciplinaire. *Revue française de linguistique appliquée* 12(2), pp. 45-64.
- L'Homme, M.C. (2008, à paraître). "DiCoInfo. Une méthodologie pour une nouvelle génération de dictionnaires spécialisés", *Traduire*.
- L'Homme, M.C. (2008). "Ressources ontologiques et ressources linguistiques : une analyse comparative dans le domaine de l'informatique", *Revue française de linguistique appliquée* 13(1), pp. 97-118.
- L'Homme, M.C. (2007). "Defining concepts: starting from concepts themselves or from empirical data". A comment of Ekaterina Mhaana's article entitled "Processual concept relations by methods of terminology", *Terminology Science and Research* 18 (<http://lipas.uwasa.fi/hut/svenska/iitf/tsr2005/vol18/>).

- Lefrançois, P. (2008). Faites-en voir de toutes les couleurs à vos élèves! *Vivre le primaire* 21(2), pp. 27-28.
- Marshman, E. (à paraître) “Expressions of uncertainty in candidate knowledge-rich contexts: A comparison in English and French specialized texts.” *Terminology* 14(1), pp. 124-151.
- Marshman, E. (2007) “Towards strategies for processing relationships between multiple relation participants in knowledge patterns: An analysis in English and French.” *Terminology* 13(1), pp.1–34.
- Marshman, E., M.C. L’Homme & V. Surtees (à paraître). “Portability of lexical markers of the cause-effect relation”, *Corpora* 13(2).
- Mel’čuk, I. (à paraître). Phraséologie dans la langue. *Repères & Applications (VI), XXIV<sup>e</sup> Journées Pédagogiques sur l’Enseignement du Français en Espagne*, Barcelone, 3-5 septembre 2007.
- Mel’čuk, I. & A. Polguère. (à paraître). Prédicats et quasi-prédicats sémantiques dans une perspective lexicographique. *LIDIL*, numéro « Syntaxe et sémantique des prédicats, dirigé par Z. Guenthéva & Eva Novakova.
- Mel’čuk, I. & L. Wanner (à paraître). Morphological Mismatches in Machine Translation. *Machine Translation*.
- Polguère, A. (à paraître) Lexical systems: Graph models of natural language lexicons. *Language Resources and Evaluation*, numéro spécial « Multilingual Language Resources and Interoperability », sous la direction de A. Witt & G. Sérasset, Springer Verlag.
- Vandaele, S., M. Raffo & S. Boudreau S. (à paraître) Les défis de la traduction spécialisée : mise en œuvre d'un site de référence en biomédecine, TTR.

### **6.3 Articles dans actes de congrès**

- Bae, H.S. & M.C. L’Homme (2007, à paraître). Converting a Monolingual Lexical Database into a Multilingual Specialized Dictionary. In *Proceedings of the Conference on Multilingualism and Applied Comparative Linguistics*, Bruxelles.
- Drouin, P. & S. Reid Triantafyllos (à paraître). Automatic extraction of transdisciplinary scientific collocations in General Scientific Language. In *Proceedings of Language for Special Purposes 2007*.
- Drouin, P. & A. Le Serrec (2007). Référentiel terminologique et ontologie :

- complémentarité et interdépendance, In *Actes de l'Atelier Ontologies et Textes.. Terminologie et intelligence informatique (TIA) 2007*, Sophia-Antipolis, France, pp. 29-38.
- L'Homme, M.C. & A. Polguère (à paraître). Mettre en bons termes les dictionnaires de langue de spécialité et les dictionnaires de langue générale. *Actes. Lexicographie et terminologie : histoire de mots. Colloque en l'honneur d'Henri Béjoint*. Université Lyon-2, France.
- Leblanc, S.-A. (2007). Description of French Idioms in a Lexical Database. In *Proceedings of the Third International Conference on the Meaning Text Theory (MTT'2007)*, Klagenfurt.
- Marshman, E. (à paraître) Using lexical knowledge patterns for terminology work in English and French: Some important differences. In *Proceedings of the 16th European Symposium on Language for Special Purposes*, Hamburg, Germany, 27–31 August 2007.
- Marshman, E. & M.C. L'Homme (2008). “Portabilité des marqueurs de la relation causale : étude sur deux corpus spécialisés”, In *Corpus et dictionnaires de langues de spécialité*, Université Lumière Lyon-2, Lyon, 28-29 septembre 2006.
- Marshman, E. & S. Vandaele (à paraître) Metaphorical conceptualization of associations in medical texts: An analysis in English and French. In *Proceedings of the 16th European Symposium on Language for Special Purposes*, Hamburg, Germany, 27–31 August 2007.
- Polguère, A. (2007). Is ‘Meaning-Text Generative Lexicon’ a Swearword? In *Proceedings of the Fourth International Workshop on Generative Approaches to the Lexicon (GL'2007)*, Paris.
- Polguère, A. (2007). Lessons from the *Lexique actif du français*. In *Proceedings of the Third International Conference on the Meaning Text Theory (MTT'2007)*, Klagenfurt.
- Polguère, A. (2008). La lexicographie explicative et combinatoire à l'épreuve de l'informatisation. In *Lexicographie et informatique. Bilan et perspectives*, Pré-actes du Colloque international à l'occasion du 50<sup>e</sup> anniversaire du lancement du projet du Trésor de la Langue Française, Nancy, pp. 119-128.
- Thibault, A. & P. Drouin (à paraître). Le lexique de Joseph Zobel, auteur antillais : extraction semi-automatique des particularismes lexicaux. In *Actes du XXVe Congrès International de linguistique et de philologie romane (CILPR 2007)*.



## 6.4 Communications non publiées

- Dancette, J. & R. Charabati Une terminologie multilingue exploitant les relations sémantiques. Colloque international Terminologie : approches transdisciplinaires, Gatineau, 2-4 mai 2007.
- L'Homme, M.C. (2007). Collecting, Encoding and Organizing Collocates in a Terminological Database. *XVth European Symposium on Language for Special Purposes (LSP)*, University of Hamburg (Germany).
- L'Homme, M.C. (2007). La petite histoire des actants dans le DiCoInfo (Dictionnaire fondamental de l'informatique et de l'Internet). *Terminologie, approches transdisciplinaires*, Université du Québec en Outaouais, Gatineau (Québec).
- Leblanc, S.-A. (2008). L'alexicalisation des termes interprétatifs, essai de caractérisation signifiante de *kitsch* et *trash*. *Colloque du kitsch au trash: excès de mauvais goût et éloge du rebut*, Université de Montréal.
- Lefrançois, P. (2008). *Le « bon usage » du dictionnaire en classe dans le contexte du renouveau pédagogique*. Participation à une table ronde tenue à la Faculté des sciences de l'éducation de l'Université de Montréal, Montréal.
- Lefrançois, P. (2007). *Are reference books the key to solving syntactic and lexical problems at university level?* Présenté à la 15e Conférence européenne sur la lecture, Berlin (Allemagne).
- Lefrançois, P. & O. Tremblay (2008). *Le dictionnaire, matériau de construction d'activités sur la langue*. Atelier offert à la librairie Alire (Longueuil) et à la librairie Monet (Montréal).
- Lefrançois, P. & O. Tremblay (2007). *L'enseignant de français au primaire, un curieux professionnel*. Présenté au congrès de l'Association québécoise des professeurs de français (AQPF), Saint-Hyacinthe.
- Vandaele, S. & M. Raffo (2007). Indices de conceptualisation et syntagmes terminologiques nominaux à composante prédicative en biologie cellulaire. *Terminologie : Approches transdisciplinaires*, Université du Québec en Outaouais, Gatineau, du 4 au 6 mai 2007.
- Vandaele, S. & M. Raffo (2007). Revisiter l'équivalence : modes de conceptualisation et réseaux lexicaux. XX<sup>e</sup> congrès de l'Association canadienne de traductologie, *La formation en traduction : pédagogie, docimologie et technologie*, Université de Saskatoon du 26 au 28 mai 2007.

## 6.5 Autres

- Blanco, X., Marc Van Campenhoudt & M.C. L'Homme (éd.) (2008, à paraître). *Lexique, dictionnaires et connaissance dans une société multilingue*, numéro spécial de la revue *Cahiers de linguistique*.
- Carreño Cruz, I., P. Drouin, M.C. L'Homme & C. Crow (2007). *An Overview of Existing Linguistic Methods for the Recognition of Semantic Relations in Textual Corpora*, Observatoire de linguistique Sens-Texte (OLST).
- Drouin, P. & F. Doll (2007). *List of Main Terms for the Subject Area of IEDs*. Montreal: Observatoire de linguistique Sens-Texte (OLST).
- Drouin, P. & M.C. L'Homme (2007). *SACOT Project: Final Report*. Montreal: Observatoire de linguistique Sens-Texte (OLST).
- L'Homme, M.C. (dir.) (2007). DicoInfo. Dictionnaire fondamental de l'informatique et de l'Internet. Version Web (<http://olst.ling.umontreal.ca/dicoinfo>).
- Lefrançois, P. et M. Cormier (coordinatrices). *Semaine des dictionnaires 2008* (29 mars au 4 avril 2008).
- Lefrançois, P., O. Tremblay & D. Anctil (2008). *Trousse d'activités pédagogiques autour du dictionnaire*. En ligne : <http://diffusion-didactique.scedu.umontreal.ca/documents/document.php?id=14>.