

Observatoire de linguistique Sens-Texte

OLST

Rapport d'activités

mai 2010 – avril 2011

Table des matières

1. Membres	3
Membres réguliers	3
Membres associés	3
Membres adjoints	3
2. Projets et état d'avancement des travaux	4
Dictionnaires et bases de données lexicales	4
Analyses de données lexicales	4
Modélisation de données lexicales	5
Accès aux données lexicales	6
Corpus	6
Traitement automatique de la langue	6
Paraphrase	7
Didactique du lexique	7
3. Encadrement d'étudiants aux études supérieures	8
4. Personnel	10
Étudiants	10
Stagiaires	11
5. Autres activités	12
5.1 Séminaires, colloques	12
5.2 Professeurs, chercheurs invités	12
5.3 Collaborations	13
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	20

1. Membres réguliers et associés

Membres réguliers

Lyne Da Sylva

Patrick Drouin (directeur adjoint)

Lidjia Iordanskaja

Pascale Lefrançois

Marie-Claude L'Homme (directrice)

Elizabeth Marshman

Igor Mel'čuk

Jasmina Milićević

Alain Polguère (jusqu'au 1^{er} septembre 2010)

Membres associés

Amparo Alcina, TecnoLeTTra, Universitat Jaume I.

Aline Francoeur, Département de langues, linguistique et traduction, Université Laval
(depuis le 2 février 2011)

Sébastien Marengo, CATIFQ, Université de Sherbrooke
(depuis le 28 mars 2011)

Alain Polguère, Nancy-Université & ATILF CRNS
(depuis le 1^{er} septembre 2010)

Ophélie Tremblay, Département de didactique des langues, UQAM,
(depuis le 15 février 2010)

Membres adjoints

Benoît Robichaud

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 2010 à avril 2011.

Dictionnaires et bases de données lexicales

DiCoEnviro (responsable : Marie-Claude L'Homme). Ce projet vise à construire une base de données terminologique portant sur le domaine de l'environnement. La base de données est élaborée selon les principes mis au point dans le cadre du DiCoInfo. Des corpus comportant des textes anglais, français et espagnols et portant sur la thématique du réchauffement climatique ont été assemblés. Les termes ont été extraits à l'aide de TermoStat et font actuellement l'objet d'une sélection. Quelque cent fiches sont actuellement rédigées dans chacune des langues et une grande partie d'entre elles sont en ligne.

Version multilingue du DiCoInfo (responsable : Marie-Claude L'Homme) : Des versions anglaise 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 espagnol (responsables : Amparo Alcina et Marie-Claude L'Homme), la sélection des termes est amorcée par l'équipe de Castellón (un financement a été obtenu auprès de l'Université Jaume I pour ces travaux). Un séminaire de formation donné à Castellón a permis de commencer la rédaction de quelques articles. Quelques articles sont en ligne depuis mars 2010.

Analyses de données lexicales

Annotation syntaxico-sémantique de contextes spécialisés (responsable : Marie-Claude L'Homme et Guy Lapalme). Ce projet (financé par le Conseil de recherches en sciences humaines, CRSH) vise à fournir, pour des termes du DiCoInfo et du DiCoEnviro, une annotation syntaxico-sémantique. L'annotation porte sur les termes de nature prédicative ainsi que sur leurs actants et circonstants. L'équipe a mis au point un protocole d'annotation réalisé dans une structure XML. Jusqu'à présent, plus de 200 verbes du DiCoInfo et environ 50 verbes du DiCoEnviro sont annotés. Les premières annotations ont porté sur les verbes français. Toutefois, l'annotation des contextes verbaux anglais a commencé et celle des verbes espagnols débutera au cours de l'été 2010. Quelques noms de nature prédicative sont également en cours d'annotation

Élaboration d'une base de données terminologiques à partir de bitextes (responsable : Elizabeth Marshman). Ce projet (subventionné par le Conseil des recherches en sciences humaines du Canada) vise l'élaboration d'une base de données qui donne accès aux couples de termes et leurs équivalents dans le domaine de la médecine,

accompagnés d'occurrences de ces termes tirés d'un corpus de bitextes vulgarisés et semi-spécialisés. La base de données met en évidence des relations terminologiques (par exemple, générique-spécifique, partie-tout, cause-effet) dans lesquelles les termes participent dans le corpus, ainsi que les marqueurs lexicaux de ces relations dans les deux langues. La base de données pourra servir non seulement à l'élaboration d'outils d'extraction semi-automatique de ces relations à partir des marqueurs, mais aussi comme aide à la compréhension des termes pour la formation des traducteurs et terminologues.

Équivalence en terminologie (responsable : Marie-Claude L'Homme et Guy Lapalme). Ce projet (financé par le Conseil de recherches en sciences humaines, CRSH) vise à élaborer une nouvelle typologie de l'équivalence terminologique, typologie qui tiendra compte des termes de nature prédicative. Des contextes contenant des termes anglais et français sont annotés : les propriétés syntaxico-sémantiques de termes de nature prédicative ainsi que de leurs actants sont annotés. Cette annotation servira de toile de fond pour analyser les phénomènes d'équivalence entre l'anglais et le français. L'espagnol sera également pris en compte dans une phase ultérieure du projet.

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. La première étape avance à grands pas et les sens ont fait l'objet de descriptions lexicographiques sous XML. Nous en sommes à l'étape de l'établissement de l'équivalence entre les langues. Les collocations constituent la prochaine phase du projet

RHWCD – DiCoInfo (responsable : Marie-Claude L'Homme en collaboration avec Ilan Kernerman). Ce projet vise à étudier les possibilités d'enrichissement croisés d'un dictionnaire général, à savoir le Random House Webster's Collegiate Dictionary, et d'un dictionnaire spécialisé, le DiCoInfo. Des comparaisons automatiques et manuelles des nomenclatures ont été effectuées. Une analyse des articles manquants de part et d'autre est en cours.

Modélisation de données 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.

Accès aux données lexicales

Accès aux collocations multilingues spécialisées (responsables : Marie-Claude L'Homme et Benoît Robichaud en collaboration avec Patrick Leroyer). Ce projet consiste à permettre d'accéder aux traductions des collocations répertoriées dans le DiCoInfo et le DiCoEnviro. Cet accès est réalisé en repérant les fonctions lexicales (FL) associées aux collocations décrites dans chacune des langues.

DiCoInfo visuel (responsable : Benoît Robichaud). Ce projet vise le développement d'une interface « visuelle » au DiCoInfo qui permet à la fois d'afficher sous la forme de réseaux graphiques l'organisation d'unités terminologiques trouvées dans des relations lexicales spécifiques (hyponymie, synonymie, antonymie, analogie, etc.), ainsi que la navigation et l'exploration visuelles de ces réseaux.

Navigation dans les liens de combinatoire du DiCoInfo (responsable : Marie-Claude L'Homme en collaboration avec Anne-Laure Jousse, Patrick Leroyer et Benoît Robichaud). Ce projet consiste à proposer une nouvelle présentation des liens de combinatoire dans le DiCoInfo français. Plutôt que de présenter l'ensemble des liens sous forme de liste ordonnée, on propose un regroupement des liens collocationnels en fonction de classes sémantiques plus générales. Ces classes sont proposées dans chaque entrée si elles contiennent des correspondances dans l'entrée en question. Un prototype a déjà été conçu et un travail d'uniformisation des liens lexicaux dans le dictionnaire est en cours.

Corpus

Constitution d'un corpus de bitextes 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 et semi-spécialisé 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.

Traitement automatique de la langue

Annotation automatique des actants (responsable : Guy Lapalme et Marie-Claude L'Homme). Ce projet (financé par le Conseil de recherches en sciences humaines, CRSH) vise à alléger le travail d'annotation syntaxico-sémantique des termes français du DiConfo et du DiCoEnviro (voir ce projet décrit plus haut). La méthode comporte les étapes suivantes : 1) identification des participants; 2) distinction des actants et des circonstants; 3) attribution automatique de rôles sémantiques. Des expérimentations ont été réalisées pour la première tâche. Les résultats sont prometteurs. Une extension des modèles d'annotation à l'anglais et à l'espagnol doit commencer au cours de l'été 2010.

Constitution et annotation d'un corpus étalon pour l'acquisition automatique de termes dans le domaine de l'automobile (responsable : Patrick Drouin). L'ensemble des termes d'un corpus du domaine de l'automobile fait l'objet d'une annotation en XML. L'objectif est de mettre en place un corpus étalon permettant l'évaluation automatique de l'acquisition automatique de termes, mais avant tout l'élaboration de critères de sélection des termes et une description des diverses relations qu'ils entretiennent en corpus. Le corpus anglais d'environ 200 000 mots est annoté et un corpus français comparable est en cours d'élaboration

Paraphrase

Extraction de règles de paraphrasage à partir de corpus de textes parallèles bilingues (responsable : Jasmina Milićević). Subventionné par le *Research Development Fund for the Humanities and Social Sciences* de Dalhousie University jusqu'en juin 2011, ce projet vise à repérer les liens paraphrastiques dans un corpus de textes littéraires bilingues (les paires de langues impliquées sont russe-français, français-anglais et espagnol-français.) La description des liens se fait sous la forme de règles de paraphrasage utilisables pour diverses tâches de TAL (reformulation, création de résumés, traduction automatique, etc.).

Didactique du lexique

Élaboration et validation d'un instrument d'évaluation du vocabulaire d'élèves québécois de 3e secondaire (responsable : Pascale Lefrançois, en collaboration avec Marie-Éva de Villers). Ce projet consiste à développer un questionnaire évaluant différents aspects de la connaissance lexicale des élèves québécois francophones de 3e secondaire, à partir d'une variété de tâches fonctionnelles et métalinguistiques. Le questionnaire a été validé par des experts, puis par des passations expérimentales auprès d'élèves de trois régions du Québec. Une fois le questionnaire validé, il pourra être administré dans une enquête visant à faire le portrait des connaissances lexicales des jeunes Québécois.

Portrait du vocabulaire d'élèves québécois de 3e secondaire (responsable : Pascale Lefrançois, en collaboration avec Marie-Éva de Villers). Dans sa première phase, ce projet consistait à développer un questionnaire évaluant différents aspects de la connaissance lexicale des élèves québécois francophones de 3e secondaire, à partir d'une variété de tâches fonctionnelles et métalinguistiques. Le questionnaire a été validé par des experts, puis par des passations expérimentales auprès d'élèves de trois régions du Québec. Le questionnaire validé est administré à 2000 élèves représentatifs de la population scolaire québécoise afin de faire le portrait de leurs connaissances lexicales.

Constitution de listes de mots pour l'apprentissage de l'orthographe et du lexique au primaire et au début du secondaire (responsable : Pascale Lefrançois, en collaboration avec Ophélie Tremblay). À la demande du MELS, ce projet consistait à sélectionner, à

partir d'un corpus tiré d'ouvrages de littérature de jeunesse et d'autres listes, les mots devant être appris par les élèves du primaire et du premier cycle du secondaire. Pour chaque mot, un exemple d'utilisation a été fourni, en plus d'une indication de la polysémie et de la présence de cooccurrents.

3. Encadrement d'étudiants aux études supérieures

Marta Allard (3e cycle, traduction, Université d'Ottawa). *Terminology Management for Translators in Translation Environments*. (codirection : Lynne Bowker et Elizabeth Marshman; en cours).

Dominic Anctil (3^e cycle, didactique), *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; thèse soutenue le 17 mai 2011).

Évelyne Arbour-Barbaud (2^e cycle, linguistique) : *Dépendances syntaxiques du français: problèmes d'identification et d'encodage lexicographique* (Alain Polguère; en cours).

Gabriel Bernier-Colborne (2e cycle, traduction) : *Constitution et annotation d'un corpus étalon pour l'acquisition automatique de termes dans le domaine de l'automobile* (codirection : Patrick Drouin et Marie-Claude L'Homme; en cours).

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

Marie-Claude Demers (2e cycle, traduction). *Méthode d'enrichissement de la nomenclature d'un dictionnaire général : ajout de termes d'informatique* (Marie-Claude L'Homme; en cours).

Stephanie Doyle (2e cycle, linguistique). *Apprentissage et enseignement des catégories flexionnelles verbales du français et de l'anglais : étude comparative*. (Jasmina Milićević; accepté en janvier 2011)

Widade El Cherif (2e cycle, linguistique). *Dérivation des noms d'agent en français*. (Jasmina Milićević; en cours).

Mélanie Gagnon (2^e cycle, traduction). *Recherche terminologique dans le domaine de l'environnement* (Marie-Claude L'Homme, en cours).

Laure Guitard (3e cycle, École de bibliothéconomie et des sciences de l'information). *Développement d'un modèle pour la définition de critères régissant l'indexation-sujet de documents d'archives.* (Lyne Da Sylva, en cours).

Fadila Hadouche (3^e cycle, informatique). *Annotation automatique de rôles actanciels en corpus spécialisé* (codirection : Guy Lapalme et Marie-Claude L'Homme; thèse soutenue en mars 2011).

Joshua Holden (3^e cycle). *Description du lexique fondamental du déné suliné en relation avec la morphologie du verbe déné* (Igor Mel'cuk; thèse soutenue en janvier 2011)

Anne-Laure Jousse (3e cycle, linguistique). *Modèle de structuration des relations lexicales fondé sur le formalisme des fonctions lexicales* (cotutelle : Sylvain Kahane – Paris 10 & Paris 7 (pour l'inscription de la thèse) – et Alain Polguère; thèse soutenue en juillet 2010).

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

Audrey Laroche (3e cycle). *Identification automatique de la polysémie lexicale : application à la création semi-automatisée d'un dictionnaire d'analogies et à la désambiguïsation automatique* (codirection : Patrick Drouin et Philippe Langlais; en cours).

Annaïch Le Serrec (3^e cycle, traduction). *Équivalence terminologique : étude comparative dans des corpus de genres textuels différents* (Marie-Claude L'Homme; en cours).

Sara-Anne Leblanc (3e cycle, linguistique). *Combinatoire des locutions françaises dans une perspective lexicographique : à l'interface entre lexique et grammaire* (codirection : Alain Polguère et Sylvain Kahane – Paris 10; en cours).

Patricia Lee Men Chin (3e cycle, linguistique appliquée). *Étude du rôle des logiciels d'ALAO dans le développement des stratégies d'apprentissage : une méta-analyse.* (codirection : Marie-Josée Hamel, Université d'Ottawa, et Jasmina Milićević; en cours.)

François Louis (3e cycle, linguistique). *Le vocabulaire d'un bretonnant de naissance de Pluméliau, région de Pontivy* (codirection : Francis Favreau, Université de Rennes-2 et Jasmina Milićević; en cours.)

Cheryl McBride (3e cycle, traduction, Université d'Ottawa). *Teaching Translation Technologies by Distance Education* (Elizabeth Marshman; en cours).

Janine Pimentel (3^e cycle, traduction). *Criteria for the Validation of Specialized Verb Equivalents: Applications in Bilingual Terminography* (codirection : Marie-Claude L'Homme et Rute Costa, Université nouvelle de Lisbonne; en cours).

Caroline Poudrier (2^e cycle, linguistique). *Évaluation de deux méthodes d'identification des liens lexicaux syntagmatiques : méthodes manuelle et statistique* (codirection : Alain Polguère et Patrick Drouin).

Karine Pouliot (2^e cycle, didactique). *Élaboration d'un vocabulaire du lexique transdisciplinaire scientifique : principaux termes et collocations* (codirection : Pascale Lefrançois et Monique Cormier; en cours).

Mélissa Samson-Fauteux (2^e cycle, linguistique). *Représentations du sens dans la théorie linguistique Sens-Texte* (codirection : Igor Mel'čuk et Jasmina Milićević; accepté en novembre 2010)

Alexandra Tsedryk (3^e cycle, linguistique). *Compétence paraphrastique en français langue seconde*. (Jasmina Milićević; en cours.)

David Wilton (2^e cycle, linguistique). *Description lexicographique des lexies dénotant des animaux dans un dictionnaire explicatif et combinatoire*. (Jasmina Milićević; en cours.)

Ibtissam Yacoub (2^e cycle, traduction). *Recherche terminologique dans le domaine de l'environnement* (Marie-Claude L'Homme; en cours).

4. Personnel

Étudiants

Suzie Beaudry (baccalauréat en traduction). Rédaction de fiches dans le DiCoEnviro.

Gabriel Bernier Colborne (maîtrise en traduction). Constitution et annotation d'un corpus étalon pour l'acquisition automatique de termes dans le domaine de l'automobile.

Anne Bertrand (maîtrise en linguistique). Analyse des résultats du projet Texto4Science.

Marie-Claude Demers (maîtrise en traduction). Rédaction de fiches dans le DiCoInfo anglais.

Suzanne DesGroseilliers (baccalauréat en traduction). Rédaction de fiches dans le DiCoEnviro.

Mélanie Gagnon (maîtrise en traduction). Rédaction de fiches dans le DiCoEnviro.

Marie-Hélène Girard (maîtrise en traduction). Rédaction de fiches dans le DiCoInfo.

Charissa Harms (maîtrise en traductologie, Université d'Ottawa). Enrichissement du corpus et analyse des données pour le projet des marqueurs.

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; navigation dans les liens lexicaux du DiCoInfo.

Marie-Ève Laneville (maîtrise en traduction) : Classes d'hyponymes dans le DiCoInfo

Audrey Laroche (doctorat en linguistique). Développement de TermoStat Web.

Annaïch Le Serrec (doctorat en traduction). Annotations contextuelles et DiCoEnviro.

Amélie Paulus (baccalauréat en traduction). Annotation de textos pour le projet Texto4Science.

Janine Pimentel (doctorat en traduction). Annotations contextuelles du DiCoInfo anglais.

Eugénie Rompré Brodeur (baccalauréat en traduction). Annotation de textos pour le projet Texto4Science.

Stagiaires

Sébastien Marengo, Université de Strasbourg 2 (stagiaire post-doctoral). Étude des structures régies par les gouverneurs syntaxiques adjectivaux en français.

Nava Maroto García, Université Complutense, Madrid (stagiaire post-doctorale). Étude des relations « activité – argument » et « activité – circonstant ») dans le domaine de la céramique industrielle.

Aurélien Picton, Université Toulouse-le-Mirail (stagiaire post-doctorale). Étude des aspects diachroniques d'un corpus du domaine de la criminologie.

Teresa Ortego Antón, Université de Valladolid (stagiaire doctorale). Participation à la rédaction du DiCoEnviro espagnol.

5. Autres activités

5.1 Séminaires, colloques

17^e Conférence sur le Traitement automatique des langues naturelles (TALN 2010), juillet 2010.

Journées des thésards de l'OLST. Observatoire de linguistique Sens-Texte (OLST), juin 2011.

Journées lexicographie / terminologie de l'Observatoire de linguistique Sens-Texte (OLST), juin 2010.

Préparation de la Conférence Meaning-Text Theory (MTT 2011), Barcelone, Espagne, Septembre 2011.

Préparation d'une Table ronde « Différentes perspectives sur le traitement automatique des documents: l'indexation » dans le cadre du Congrès des milieux documentaires du Québec, Montréal, 4 novembre 2010.

Préparation du 6^e colloque international DEFT (Défi Fouille de Texte), Université de Montréal, Montréal, 23 juillet 2010. Membre du comité d'organisation.

Préparation du Colloque « Approches sémiotiques au design », dans le cadre du *78^e congrès de l'ACFAS*, Université de Montréal, Montréal, 10-11 mai 2010. Coresponsable du colloque : Lyne Da Sylva (responsable : Yves Marcoux, EBSI).

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

Pascaline Dury (Centre de recherche en terminologie et en traduction, Université Lyon-2, France)

Ilan Kernerman (K Dictionaries, Tel Aviv, Israël).

Patrick Leroyer (Centlex, Université d'Aarhus, Danemark)

Carlos Subirats Rüggeberg (Universitat Autònoma de Barcelona et International Computer Science Institute (ICSI) de Berkeley).

5.3 Collaborations

Convention d'échanges qui nous liait à l'Université Stendhal (Grenoble) : renouvellement sous la forme d'une convention tripartite, maintenant avec l'Université de Cologne; renouvelée pour une période de 3 ans.

Pascaline Dury (Centre de recherches en traduction et terminologie, Université Lyon-2, France) : Description des particularités diachroniques d'un corpus du domaine de l'écologie et établissement d'une méthodologie de repérage semi-automatique des nécrologismes.

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

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

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

K Dictionaries (Tel Aviv, Israël) : Enrichissement de ressources lexicographiques et terminologiques, comparaison des nomenclatures du DiCoInfo et du Random House Webster's Collegiate Dictionary.

Guy Lapalme (Recherche appliquée en linguistique informatique - RALI, Université de Montréal) : conversion au format XML du DiCoInfo et prises en compte de données multilingues.

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.

Patrick Leroyer (Centlex, Université d'Aarhus, Danemark) : Projets d'accès aux traductions multilingues spécialisées et de navigation dans les liens lexicaux du DiCoInfo.

6 Publications

6.1 Livres ou chapitres de livres

L'Homme, M.C. (2011, à paraître). Large Terminological Databases. *International Encyclopedia of Lexicography*.

L'Homme, M.C. (2010). Designing Terminological Dictionaries for Learners based on Lexical Semantics: The representation of actants. In Fuentes-Olivera, P. (ed.).

- Specialised Dictionaries for Learners*, Berlin/New York: De Gruyter, pp. 141-153.
- Marshman, E. (soumis) Challenges of Lexical Knowledge Patterns for Bilingual Terminology Work. In Bowker, L. (ed). *University of Ottawa School of Translation and Interpretation 35th Anniversary Volume*.
- Marshman, E. Expressions of uncertainty in candidate knowledge-rich contexts: A comparison in English and French specialized texts (2010). In Auger, A. & C. Barrière, (eds.) *Probing Semantic Relations: Exploration and Identification in Specialized Texts*. (John Benjamins, 2010). p. 121–148 (paru dans *Terminology* 14(1).)
- Marshman E. & L. Bowker. (accepté). Attitudes of key stakeholders in translator training: Can CERTT help turn conflict into collaboration? In Borodo, M. and S. Hubscher-Davidson, (eds.) *Global Trends in Translator and Interpreter Training: Mediation and Culture*. Continuum.
- Marshman, E. & L. Bowker. (soumis). Translation Technology Training for the 21st Century: Observations from the Collection of Electronic Resources in Translation Technologies (CERTT) Project. In Melby, A.K. & M. Savova, (eds.) *Bringing Together Government, Education, and Industry in T&I*. John Benjamins.
- Mel'čuk, I. (2010) The Government Pattern in the Explanatory Combinatorial Dictionary. In: G.-M. de Schryver, ed., *A Way with Words: Recent Advances in Lexical Theory and Analysis. A Festschrift for Patrick Hanks*, Kampala (Uganda): Menha Publishers, 89 -118.
- Mel'čuk, I. (2011). Phrasèmes dans le dictionnaire. In J.-C. Anscombre & S. Mejri (eds.) *Le figement linguistiques: la parole entravée*, Paris: Honoré Champion, 41-61.
- Mel'čuk, I & J. Milićević. (sous presse). *Introduction à la linguistique*. Paris: Hermann; pp. 750.

6.2 Articles dans revues avec comité de lecture

- Beck, D. & Mel'čuk. (2011). I. Morphological Phrasemes and Totonacan Verbal Morphology. *Linguistics* 49(1), pp. 175-228.
- Drouin, P. (2010). Extracting a bilingual transdisciplinary scientific lexicon. *Proceedings of eLexicography*, p. 43-54
- Dury, P. & P. Drouin (2011). When terms disappear from a specialized lexicon: A semi-automatic investigation into 'necrology'. *ICAME Journal*, 35, p. 19-33.

- Hadouche, F., M.C. L'Homme & G. Lapalme (2010). Automatic Annotation of Actants. In Granger, S. & M. Paquot (eds.). *eLexicography in the 21st Century: New challenges, new applications. Les Cahiers du Cental*.
- Iordanskaja, L. & I. Mel'čuk (2010). La coordination des adjectifs modificateurs en russe et en français: la conjonction russe I et la conjonction française ET. *Méta* 55(1), pp. 168-185.
- L'Homme, M.C. (2010). Designing Specialized Dictionaries with Natural Language Processing. Examples of applications in the fields of computing and climate change. In Granger, S. & M. Paquot (eds.). *eLexicography in the 21st Century: New challenges, new applications. Les Cahiers du Cental*.
- Le Serrec, A., M.C. L'Homme, P. Drouin & O. Kraif (2010). Automating the compilation of specialized dictionaries: use and analysis of term extraction and lexical alignment. *Terminology* 16(1), pp. 77-106.
- Mel'čuk, I & J. Milićević (sous presse). Budalo jedna!-Type Construction in Contemporary Serbian. *Journal of Slavic Linguistics*; pp. 25.
- Paquin, A. & P. Drouin (2010). « Néologie du terrorisme : médiation et intégration », *Nologica : revue international de néologie*, 4, p. 105-118.
- Pimentel, J., M.C. L'Homme & M.È. Laneville (soumis). General and Specialized Lexical Resources: a Study on the Potential of Combining Efforts to Enrich Formal Lexicons, *International Journal of Lexicography*.

6.3 Articles dans actes de congrès

- Arbour-Barbaud, É. (2010). Carnaval : une base de données de régimes généralisés. In Neveu, F. et al. (éds), *CMLF 2010 – 2e Congrès Mondial de Linguistique Française*, <http://www.linguistiquefrancaise.org/>, 861-874.
- Da Sylva, L. (2010) Extraction semi-automatique d'un vocabulaire savant de base pour l'indexation automatique. *Actes du congrès Traitement automatique des langues naturelles* (TALN 2010), Université de Montréal, 22 juillet 2010. [http://www.iro.umontreal.ca/~felipe/TALN2010/Xml/Papers/all/taln2010_submission_86.pdf]
- Da Sylva, L. (2010). Integrating knowledge from different sources for automatic back-of-the-book indexing. *Actes du Congrès 2010 de l'Association Canadienne des Sciences de l'Information*, Université Concordia, Montréal, 3 juin 2010. [http://www.cais-acsi.ca/proceedings/2010/CAIS061_DaSylva_Final.pdf]

- Drouin, P. (2010). From a Bilingual Transdisciplinary Scientific Lexicon to Bilingual Transdisciplinary Scientific Collocations. *Proceedings of the XIV Euralex International Congress*, 10 pages.
- Drouin, P. & F. Doll (2010). Corpus Genres and Contrastive Automatic Term Extraction. *Terminology and Knowledge Engineering (TKE 2010)*, 12-13 August, Dublin, 21 pages.
- Drouin, P. & F. Doll (2011). Potentiel terminologique, quel sens prendre ? *Actes des 8es Journées scientifiques du Réseau LTT*, p. 441-554
- Drouin, P. & S. Reid Triantafyllos (2010). Automatic extraction of transdisciplinary scientific collocations in General Scientific Language. *Proceedings of Language for Special Purposes (LSP 2007)*, 13 pages.
- Dury, P. & P. Drouin (2010), L'obsolescence des termes en langues de spécialité, une étude semi-automatique de la « nécrologie » en corpus informatisés, appliquée au domaine de l'écologie. *Proceedings of the XVII European Symposium On Languages For Specific Purposes*, 15 pages.
- Hadouche, F., G. Lapalme & M.C. L'Homme (2010). Identification des actants et circonstants par apprentissage machine. *Traitement automatique des langues TALN 2010*, 19 au 23 juillet, Université de Montréal, Montréal.
- L'Homme, M.C., P. Leroyer & B. Robichaud (2010). Advanced Encoding for Multilingual Access in a Terminological Database – A Matter of Balance. *Terminology and Knowledge Engineering (TKE 2010)*, 12-13 August, Dublin.
- Laroche, A. (2010). Attribution d'auteur au moyen de modèles de langue et de modèles stylométriques. In *Actes de la 12e Rencontre des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RÉCITAL 2010)*, 19-22 juillet 2010, Montréal, 10 p.
- Laroche, A. & P. Langlais (2010). Revisiting Context-based Projection Methods for Term-Translation Spotting in Comparable Corpora. In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING 2010)*, 23-27 août 2010, Beijing, pp. 617-625.
- Milićević, J. (2010). Extraction de paraphrases sémantiques et lexico-syntaxiques de corpus parallèles bilingues. 2010. In Langlais, P. et al. (éd.). *Actes de la 17e conférence Traitement Automatique des Langues Naturelles (TALN)*, Montréal, Canada, pp. 10.
[\[http://www.iro.umontreal.ca/~felipe/TALN2010/Xml/Papers/all/taln2010_submission_18.pdf\]](http://www.iro.umontreal.ca/~felipe/TALN2010/Xml/Papers/all/taln2010_submission_18.pdf)

- Milićević, J. & A. Polguère. (2010) Ambivalence sémantique des noms de communication langagière en français. 2010. In Neveu, F. et al. (éds.). *Second Congrès Mondial de Linguistique Française (CMLF 2010)*. 978-2-7598-0534-1, Paris 2010, Institut de linguistique française; 1029-1050 [DOI 10.1051/cmlf/2010102].
- Ortego, T. (à paraître). La compilation de DiCoEnviro en espagnol. In *IIIe Congrès International de Linguistique de Corpus (CILC)*, Universidad Politécnic de Valencia: 7 avril 2011.
- Pimentel, J. & M.C. L'Homme (2011). Annotation syntaxico-sémantique de contextes spécialisés : application à la terminographie bilingue. In Van Campenhoudt, M., T. Lino et R. Costa (dir.) *Passeurs de mots, Passeurs d'espoir. Actes des 8èmes Journées du Réseau LTT – Lexicologie, Terminologie et Traduction*, Lisbonne, 15-17 octobre 2009, Éditions des archives contemporaines, Paris, pp. 651-670.
- Thibault, A. & P. Drouin (2010). Le lexique de Joseph Zobel, auteur antillais : extraction semi-automatique des particularismes lexicaux, *Actes du XXVe Congrès International de linguistique et de philologie romane (CILPR 2007)*, 8 pages.

6.4 Communications non publiées

- Carreño, I. (2010). Étude comparative trilingue de la variation terminologique. *Semaine LexTerm (Lexicographie et Terminologie)*, Observatoire de linguistique Sens-Texte, Université de Montréal, 31 mai au 7 juin 2010.
- Carreño, I. & A. Le Serrec. (2010). Équivalence en corpus spécialisés parallèles et comparables. In *78^e Congrès de l'Acfas*, Montréal, 10-14 mai 2010.
- Da Sylva, L. (2011). Bibliothèques numériques et traitement automatique de la langue, *Séminaire RALI-OLST-MITACS*, Département d'informatique et de recherche opérationnelle, Université de Montréal, 9 février 2011.
- Da Sylva, L. (2011). Séance « Méthodologies de recherche en linguistique ». Dans le *séminaire de doctorat SCI7002 Méthodes de recherche*, École de bibliothéconomie et des sciences de l'information, 23 novembre 2010.
- Da Sylva, L. (2010). Introduction et conclusion-synthèse lors de la Table ronde « Différentes perspectives sur le traitement automatique des documents: l'indexation », dans le cadre du *Congrès des milieux documentaires du Québec*, Montréal, 4 novembre 2010.

- Drouin, P. (2011). Approches linguistiques et statistiques à la représentation du potentiel terminologique : À la recherche des termes ou la quête initiatique du Graal, *Centre international de recherche en aménagement linguistique* (CIRAL), Université Laval, conférencier invité.
- Drouin, P. (2011). Texto4Science. *Centre de recherche sur les technologies émergentes*, Université de Montréal, conférencier invité.
- Drouin, P. (2010). Texto4Science. *IXe Symposium du Conseil fédéral de terminologie*, Gatineau, Canada, conférencier principal.
- Drouin, P. (2010). Collocations extraction. Université Jaume 1, Castellon, Espagne, conférencier invité.
- Drouin, P. (2010). The endless quest for termhood: Exploring statistical and linguistic techniques. Universidad Politécnica de Valencia, Valencia, Espagne, conférencier invité.
- Drouin, P. (2010). TermoStat et navigation dans le réseau lexical. *Semaine LexTerm (Lexicographie et Terminologie)*, Observatoire de linguistique Sens-Texte, Université de Montréal, 31 mai au 7 juin 2010.
- Drouin, P. & A. Bertrand (2011). Texto4Science. *Séminaires RALI-OLST-MITACS*, Département d'informatique et de recherche opérationnelle, Université de Montréal, 6 avril 2011.
- L'Homme, M.C. (2010). Le DiCoInfo et le DiCoEnviro : deux ressources lexicales pour la terminologie. *Semaine LexTerm (Lexicographie et Terminologie)*, Observatoire de linguistique Sens-Texte, Université de Montréal, 31 mai au 7 juin 2010.
- L'Homme, M.C. (2010). Terms as labels for concepts, terms as lexical units: comparing two approaches to the terminological unit. Conférence invitée dans le cadre du Symposium *Approaches to the Lexicon*, Copenhagen Business School, Copenhagen, Denmark, 8 au 10 décembre 2010.
- Labonia, D. (2010). Dictionnaires spécialisés pour le traitement de la paraphrase. *Semaine LexTerm (Lexicographie et Terminologie)*, Observatoire de linguistique Sens-Texte, Université de Montréal, 31 mai au 7 juin 2010.
- Laroche, A. (2010). Attribution d'auteur au moyen de modèles de langue et de modèles stylométriques. *Séminaires RALI-OLST MITACS*, Université de Montréal, 13 octobre 2010.

- Laroche, A. (2010). Identification automatique de la polysémie à partir des profils de cooccurrence large des signifiants lexicaux. *Semaine LexTerm (Lexicographie et Terminologie)*, Observatoire de linguistique Sens-Texte, Université de Montréal, 31 mai au 7 juin 2010.
- Le Serrec, A. & J. Pimentel (2010). Équivalence en terminologie : repérage et validation en corpus parallèle et en corpus comparable. *Semaine LexTerm (Lexicographie et Terminologie)*, Observatoire de linguistique Sens-Texte, Université de Montréal, 31 mai au 7 juin 2010.
- Léard, J.-M. & S. Marengo (2011). Pour une organisation sémantique de la grammaire, *Séminaires OLST-RALI-MITACS*, Université de Montréal, 23 mars 2011.
- Léard, J.-M. & S. Marengo (2011). Pour une organisation sémantique de la grammaire, *Séminaires du CATIFQ*, Université de Sherbrooke, 10 mars 2011.
- Leroyer, P., M.C. L'Homme & B. Robichaud (2010). Accès performant aux collocations pour la traduction spécialisée : exploitation de la sémantique des liens lexicaux dans une base terminologique multilingue. *Semaine LexTerm (Lexicographie et Terminologie)*, Observatoire de linguistique Sens-Texte, Université de Montréal, 31 mai au 7 juin 2010.
- Marengo, S. (2010). Complémentation sémantique et syntaxique : le cas des adjectifs en français, *Colloque international « ComplémentationS »*, Université de Saint-Jacques-de-Compostelle, 21 octobre 2010.
- Marengo, S. (2010). Dépendances sémantiques et syntaxiques : quatre cas de figure pour l'adjectif en français. *Séminaires OLST-RALI-MITACS*, Université de Montréal, 22 septembre 2010.
- Marengo, S. (2010). Les compléments de l'adjectif en français et leurs propriétés syntaxiques : une approche lexicographique. *78e Congrès de l'Acfas*, Université de Montréal, 13 mai 2010.
- Marshman, E., J. Quirion & L. Bowker (2011). Student comfort levels and the teaching of translation technologies. In *Monterey Forum 2011*, Monterey, Californie, É.U, 8-9 avril 2011.
- Marshman, E. & C. Harms (accepté). Translation of lexical markers of terminological relations: Observations in English-French bitexts. *18th European Symposium on Language for Special Purposes*, Perm, Russia, 22-26 August 2011.
- Milićević, J. & A. Tsedryk (2010). Paraphrasez, s'il vous plaît. Tester la compétence paraphrastique d'apprenants de FLS. In *34th Annual Meeting of the Atlantic Provinces Linguistics Association*. Charlottetown, 12-13 November 2010.

- Ortego, T. (2010). Traitement du lexique spécialisé dans les dictionnaires généraux anglais-espagnol : la perspective du traducteur. Séminaire RALI-OLST. 24 novembre 2010.
- Pimentel, J., M-E. Laneville & M.C. L'Homme (2010). Sens généraux et sens spécialisés : étude comparative des ressources lexicales FrameNet et DiCoInfo. In *78^e Congrès de l'Acfas*, Montréal, 10-14 mai 2010.
- Robichaud, B. (2010). Mashup des ressources lexicographiques et terminologiques de l'OLST. *Semaine LexTerm (Lexicographie et Terminologie)*, Observatoire de linguistique Sens-Texte, Université de Montréal, 31 mai au 7 juin 2010.
- Robichaud, B. (2011). Le DiCoInfo Visuel. *Séminaire RALI-OLST-MITACS*, Département d'informatique et de recherche opérationnelle, Université de Montréal, 27 avril 2011.
- Tremblay, O. (2010). Exploitation du lexique Actif du français pour l'enseignement du vocabulaire. *Semaine LexTerm (Lexicographie et Terminologie)*, Observatoire de linguistique Sens-Texte, Université de Montréal, 31 mai au 7 juin 2010.

6.5 Autres

- Da Sylva, L. (2011). Compte rendu de lecture : *Concevoir l'index d'un livre : histoire, actualité, perspectives*, de Jacques Maniez et Dominique Maniez, Paris : ADBS Éditions, 2009. *Documentation et bibliothèques*, 57(1):57-59, 2011.
- Desgroseillers, S. (2011). *Guide de l'utilisateur. Un guide pour la création de fiches terminologiques dans le DicoInfo et le DicoEnviro*. Observatoire de linguistique Sens-Texte (OLST).
- Pimentel, J. (2010). Compte rendu de lecture : *The Specialised Lexicographical Approach: A Step further in Dictionary-making*, de Nuria Edo Marzá. In *Terminology* 16(2), pp. 230-234
- Pimentel, J. (2010). Compte rendu de lecture : *Topics in Language Resources for Translation and Localisation*, de E. Rodrigo (ed.) In *Meta* 55(4), pp. 852-853.
- Robichaud, B. (2011) *Le DiCoInfo Visuel : navigation visuelle dans le DiCoInfo*. Observatoire de linguistique Sens-Texte (OLST), Université de Montréal.
Site Web : <http://olst.ling.umontreal.ca/dicoinfo/visuel.php>