

**Observatoire de linguistique Sens-Texte**

**OLST**

**Rapport d'activités**

**mai 2011 – avril 2012**

## Table des matières

1. Membres réguliers et associés .....	4
Membres réguliers .....	4
Membres associés .....	4
Membres adjoints.....	4
2. Projets et état d'avancement des travaux.....	4
Applications documentaires.....	4
Dictionnaires et bases de données lexicales .....	5
Analyses de données lexicales.....	5
Modélisation de données lexicales .....	7
Accès aux données lexicales.....	7
Corpus.....	8
Traitement automatique de la langue.....	8
Paraphrase.....	9
Didactique du lexique .....	9
Traductique .....	10
3. Encadrement d'étudiants aux études supérieures .....	10
4. Personnel .....	13
Étudiants .....	13
Stagiaires.....	13
5. Autres activités .....	14
5.1 Séminaires, colloques .....	14
5.2 Professeurs, chercheurs invités .....	14
5.3 Collaborations.....	14
6. Publications .....	15
6.1 Livres ou chapitres de livres.....	15

6.2	Articles dans revues avec comité de lecture .....	16
6.3	Articles dans actes de congrès .....	18
6.4	Communications non publiées .....	21
6.5	Autres .....	23

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

### **Membres réguliers**

Amparo Alcina (TecnoLeTTra - Universitat Jaume I) (depuis le 1<sup>er</sup> juin 2011)  
Dominic Anctil (depuis le 1<sup>er</sup> février 2012)  
Lyne Da Sylva  
Patrick Drouin (directeur adjoint)  
Dominic Forest (depuis mai 2011)  
Lidjia Iordanskaja  
Pascale Lefrançois  
Marie-Claude L'Homme (directrice)  
Sébastien Marengo (CATIFQ — Université de Sherbrooke) (depuis le 1<sup>er</sup> juin 2011)  
Elizabeth Marshman  
Igor Mel'čuk  
Jasmina Milićević (Département de français, Université Dalhousie)

### **Membres associés**

Dina El Kassas (Département de Français — Université de Minya – Égypte)  
(depuis 1<sup>er</sup> juin 2011)  
Aline Francoeur (Département de langues, linguistique et traduction, Université Laval)  
Alain Polguère (Université de Lorraine, CNRS, ATILF)  
Ophélie Tremblay (Département de didactique des langues, Faculté des sciences de l'éducation – UQAM)

### **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 2011 à avril 2012.

### **Applications documentaires**

*Condensés de textes pour une assimilation rapide de l'information biomédicale* (responsable : Lyne Da Sylva; co-chercheurs : Michel Gagnon, École Polytechnique de Montréal et Alain Polguère, OLST). Subventionné par la Chaire Bell en recherche interdisciplinaire sur les technologies émergentes, Programme de Financement de démarrage des recherches, ce projet vise à développer un prototype de condensés de

textes basés sur trois opérations de base : la segmentation automatique de texte, la sélection de phrases-clés pour chaque segment et la compression automatique de ces phrases basée sur une analyse syntaxique par dépendances. Les travaux cette année ont porté sur la mise en place finale du prototype, avec la chaîne complète des traitements développés indépendamment au préalable.

## **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 français, anglais et espagnols ont été extraits à l'aide de TermoStat et font actuellement l'objet d'une sélection. Entre 150 et 200 fiches sont en ligne dans les trois langues décrites. Au début de 2012, des termes appartenant au domaine des énergies renouvelables ont été ajoutés au DiCoEnviro.

***JuriDiCo*** (responsable : Janine Pimentel). Ce projet vise à construire une base de données terminologique portant sur le domaine du droit. Des corpus comportant des jugements anglais et portugais de la Cour Suprême du Canada et du Supremo Tribunal de Justiça du Portugal ont été assemblés. La base de données est élaborée selon les principes mis au point dans le cadre du projet FrameNet et selon les principes mis au point dans le cadre du DiCoInfo. Quelques adaptations mineures ont été apportées à la base de données DiCoInfo, notamment l'insertion des rubriques portant sur les frames. Actuellement, le JuriDiCo inclut seulement des verbes spécialisés. Les verbes ont été extraits à l'aide de TermoStat et ont fait l'objet d'une sélection. Quelque cent fiches sont actuellement rédigées dans chacune des langues. Le JuriDiCo permet aux utilisateurs de faire des recherches par termes ou par frames.

***Version multilingue du DiCoInfo*** (responsable : Marie-Claude L'Homme). Des versions anglaise et espagnole du DiCoInfo sont en cours de rédaction. En anglais, plus de 800 acceptions ont été distinguées et près de 5000 liens sémantiques répertoriés. En outre, les liens d'équivalence avec le DiCoInfo français sont établis pour environ 1000 lexies. En espagnol, quelques articles sont en ligne.

## **Analyses de données lexicales**

***Annotation syntaxico-sémantique de contextes spécialisés*** (responsables : 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 500

verbes du DiCoInfo et environ 200 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 a débuté au cours de l'été 2010 et quelque 300 verbes anglais sont annotés. Quelques noms de nature prédicative sont également en cours d'annotation.

***Équivalence en terminologie*** (responsables : 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 et espagnols sont annotés : les propriétés syntactico-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, le français et l'espagnol.

***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. Le LST se construit autour d'environ 3000 entrées (environ 1500 dans chaque langue). La première ébauche des entrées en anglais et en français est terminée et l'établissement d'équivalences entre les deux langues est débuté.

***Extraction et stockage de relations terminologiques à partir de bitextes*** (responsable : Elizabeth Marshman). Ce projet (subventionné en 2010-2011 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. Des tests auprès d'étudiants dans des programmes en traduction pour évaluer leurs besoins et préférences en matière du contenu et de l'interface de consultation des données sont en cours.

***Observatoire de néologie du Québec (OBNEQ)*** (responsable : Aline Francoeur). L'OBNEQ a été créé en 2005 par le professeur Pierre Auger de l'Université Laval, en collaboration avec l'Office québécois de la langue française (OQLF), dans le cadre du Réseau des observatoires de néologie des langues romanes (NEOROM). Le réseau

NEOROM est coordonné par la professeure Maria Teresa Cabré, de l'Université Pompeu Fabra à Barcelone. Les travaux menés à l'OBNEQ ont pour but d'étudier le phénomène de la néologie en langue générale et en langues de spécialité, tant du point de vue théorique que pratique.

**RHWCD – DiCoInfo** (responsable : Marie-Claude L'Homme en collaboration avec Ilan Kernerman). Ce projet vise à étudier les possibilités d'enrichissements 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. La rédaction des articles faisant défaut à la nomenclature du RHWCD est en cours (plus de 300 articles ont été rédigés jusqu'à maintenant).

### **Modélisation de données lexicales**

**Notions fondamentales de lexicologie, lexicographie et terminologie** (responsables : Alain Polguère et Ophélie Tremblay). Il s'agit de la construction d'une ontologie, appelée par euphémisme GTN (*Gros Tas de Notions*). Le GTN a été utilisé comme base de connaissances dans le cadre du projet Lexitation (ontologisation des savoirs lexicographiques) d'Alain Polguère, financé par le CRSH et achevé en 2010. Le GTN, dans sa composante décrivant les notions lexicales, a servi de base pour la recherche de thèse de doctorat d'Ophélie Tremblay, soutenue en octobre 2009. Il a aussi été exploité et développé, dans son secteur syntaxique, dans le cadre de la recherche de maîtrise d'Évelyne Arbour-Barbaud, dont le mémoire a été soumis en mai 2012 (réunion du jury début juin 2012). Avec le départ d'Alain Polguère pour l'Université de Lorraine, le projet est transféré en France et temporairement dormant.

### **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. Ce projet a entraîné une révision en profondeur des collocations déjà encodées dans le DiCoInfo.

**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 des unités terminologiques trouvées dans des relations lexicales spécifiques (hyperonymie, synonymie, antonymie, analogie, etc.), ainsi que la navigation et l'exploration visuelles de ces réseaux. Une seconde version incluant quelques fonctionnalités d'aide à la compilation des fiches est développée dans le langage de programmation *Prolog*.

***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 présentées dans les résultats de l'interface de recherche 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. Une évaluation du module de navigation est également en cours.

## **Corpus**

***Enrichissement des corpus spécialisés*** (responsable : Marie-Claude L'Homme). Le corpus d'informatique anglais a été enrichi de plusieurs centaines de milliers de mots pour alimenter le DiCoInfo anglais. En outre, deux corpus juridiques sont désormais accessibles : un corpus anglais et un corpus portugais. Enfin, le corpus d'informatique espagnol a été enrichi de nouveaux textes.

***Texto4Science / Text4Science*** (responsable Patrick Drouin). Le projet *Texto4Science* a pour objet de compiler un corpus de messages textes (textos) canadiens en anglais et en français. La première version entièrement annotée du corpus français (*Texto4Science*, 7300 textos) est complétée et la cueillette du corpus anglais (*Text4Science*) est lancée depuis le mois de décembre 2011.

## **Traitement automatique de la langue**

***Acquisition automatique de la terminologie : TermoStat*** (responsable : Patrick Drouin). Le logiciel d'acquisition automatique de terminologie TermoStat est désormais offert en version multilingue (anglais, espagnol, français et portugais) grâce à une interface Web améliorée. Cette nouvelle version permet aux utilisateurs de gérer leurs corpus et assure une sécurité accrue des données personnelles. La version 3 offre deux nouvelles options de navigation lexicale dans les résultats de l'acquisition des termes. Une version TermoStat a été intégrée au projet SensUnique financé par l'Agence Nationale de la Recherche de France.

***Annotation automatique des actants*** (responsables : 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 et une évaluation ont été réalisées pour la première tâche. Les résultats sont prometteurs.



***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 font l'objet d'une annotation en XML. L'objectif est de mettre un 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. La dernière année a été consacrée à une révision de l'annotation et à une validation semi-automatique de l'annotation obtenue à l'aide de scripts élaborés à cette fin.

## **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 des résumés, traduction automatique, etc.).

## **Didactique du lexique**

***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.

***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. La liste établie dans le cadre de ce projet a été exploitée dans le cadre de l'élaboration de la nomenclature dite *d'amorçage* du projet RELIEF de construction d'un *Réseau Lexical du Français* (RLF) mené par Alain Polguère à l'ATILF CNRS (Nancy).

## Traductique

*Attitudes des langagiers professionnels envers des technologies langagières* (responsable : Elizabeth Marshman): Une enquête en ligne a été menée entre février et avril 2012 auprès de plus de 250 langagiers pour évaluer leurs réactions aux effets des technologies langagières sur la nature et l'environnement de leur travail. Le projet vise à aider des concepteurs, des employeurs, des associations de langagiers et des enseignants à mieux comprendre les effets de l'intégration des technologies par le marché du travail sur les perceptions des professionnels de leur travail et leur rôle, ainsi qu'à mieux adapter les logiciels et la façon dont ils sont présentés aux professionnels afin de favoriser une intégration satisfaisante.

### 3. Encadrement d'étudiants aux études supérieures

**Palou Allard** (3<sup>e</sup> cycle, traductologie). *Gestion de la terminologie dans des environnements de traduction* (Elizabeth Marshman, 2012)

**Évelyne Arbour Barbaud** (2<sup>e</sup> cycle, linguistique). *Identification et modélisation lexicographique des dépendances syntaxiques régies du français* (Alain Polguère ; mémoire déposé en mai 2012 pour évaluation par le jury début juin).

**Marie-Pier Bernard** (2<sup>e</sup> cycle, traduction). *Traduction commentée d'un texte intitulé « Breast Cancer » paru dans le manuel Berek & Novak's Gynecology, édition de 2007* (Aline Francoeur; décembre 2011).

**Gabriel Bernier-Colborne** (2<sup>e</sup> cycle, traduction). *Élaboration d'un corpus étalon pour l'évaluation d'un extracteur de termes* (co-direction Patrick Drouin et Marie-Claude L'Homme, en cours).

**Marie-Claude Demers** (2<sup>e</sup> cycle, traduction). *Travail de recherche en lexicographie bilingue : enrichissement de la nomenclature d'un dictionnaire général à partir de celle d'un dictionnaire spécialisé* (Marie-Claude L'Homme, Janvier 2012).

**Philippe Desjardins** (2<sup>e</sup> cycle, traduction). *Traduction commentée d'un article intitulé « North American approach to the restoration of Sphagnum dominated peatlands »* (Aline Francoeur; en cours).

**Widade El Cherif** (2<sup>e</sup> cycle, linguistique). *Vers une analyse sémantique fine des noms d'agent en français* (Jasmina Milićević, mémoire accepté en décembre 2011).

**Mélanie Gagnon** (2<sup>e</sup> cycle, traduction). *Recherche terminologique thématique bilingue (anglais-français) sur les énergies renouvelables* (Marie-Claude L'Homme, Janvier 2012).

**Nizar Ghazzawi** (3<sup>e</sup> cycle, traduction). *Onto-terminologie de la maintenance informatique : étude terminologique et ontologique des verbes et des déverbaux* (Marie-Claude L'Homme; en cours).

**Émilie Giroux**. (2<sup>e</sup> cycle, terminologie). *Étude terminologique du vocabulaire de la Marine* (Aline Francoeur; en cours).

**Janelle Giroux**. (2<sup>e</sup> cycle, linguistique). *Étude des pratiques d'écriture en contexte de clavardage sportif : partisans et journalistes commentent les matchs des Canadiens de Montréal* (Aline Francoeur; en cours).

**Claudia Grenon** (2<sup>e</sup> cycle, traduction). *Traduction commentée d'un article intitulé « Economic recessions, banking reforms and the future of capitalism »* (Aline Francoeur; en cours).

**Laure Guitard** (3<sup>e</sup> cycle, 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).

**Francis Hébert** (3<sup>e</sup> cycle, sciences de l'information). *Analyse du sujet et environnement documentaire : une approche intégrée de l'organisation des connaissances* (co-direction : Dominic Forest et Lyne Da Sylva; en cours).

**Nicolas Imreh** (2<sup>e</sup> cycle, traduction). *Conception, mise sur pied et utilisation d'un corpus bilingue de termes du domaine de l'assurance collective au Québec et au Canada : repérage des différences qualitatives existant entre les termes en L1 et L2 de textes de nature contractuelle* (Aline Francoeur; en cours).

**Tetiana Kalinichenko** (2<sup>e</sup> cycle, traduction). *L'explicitation comme manifestation de l'asymétrie linguistique de la traduction anglais-français dans les textes environnementaux* (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** (3<sup>e</sup> cycle, traduction). *Équivalence terminologique : étude comparative dans des corpus de genres textuels différents* (Marie-Claude L'Homme; thèse déposée).

**Particia Lee Men Chin** (3<sup>e</sup> cycle, linguistique). *Stratégies d'apprentissage des langues secondes dans un environnement informatisé : une méta analyse qualitative de l'utilisation du courrier électronique* (Jasmina Milićević, co-supervision avec Marie-Josée Hamel de l'université d'Ottawa, thèse acceptée en août 2011).

**François Louis** (3e cycle, linguistique). *Breton Vannetais - Idiomatique, Oralité et Standard*. (Jasmina Milićević, co-supervision avec Francis Favereau de l'Université de Rennes 2, en cours).

**Anna Malkowska** (3<sup>e</sup> cycle, linguistique). *Les apports slaves en français : perspectives lexicologiques et lexicographiques* (codirection : Aline Francoeur et Benoît Leblanc; en cours).

**Cheryl McBride** (3<sup>e</sup> cycle traductologie). *L'enseignement de la traductique à distance* (Elizabeth Marshman, en cours).

**Danièle Morisset** (2<sup>e</sup> cycle, terminologie). *Étude terminologique : les hydroliennes* (Aline Francoeur; en cours).

**Karen Oostveen** (2<sup>e</sup> cycle, traduction). *Traduction commentée d'un chapitre intitulé « Beyond the simple snapshot » tiré de Digital Photography: The Missing Manual, de Chris Grover et Barbara Brundage* (O'Reilly Media, 2006) (Aline Francoeur; décembre 2011).

**Janine Pimentel** (3<sup>e</sup> cycle, traduction). *Criteria for the Validation of Specialized Verb Equivalents: Applications in Bilingual Terminography* (Marie-Claude L'Homme et Rute Costa, Université nouvelle de Lisbonne; mai 2012).

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

**Judith Rémillard** (3<sup>e</sup> cycle traduction). *Intégration de la traduction automatique dans l'environnement de traductologie* (Elizabeth Marshman et Marta Gomez; en cours).

**Atousa Reyhanipoursani** (3<sup>e</sup> cycle, linguistique). *Les anthroponymes et les toponymes : des noms propres à l'origine de noms communs. Étude lexicologique et lexicographique* (Aline Francoeur; en cours).

**Louis Thivierge** (2<sup>e</sup> cycle, terminologie). *Psychogénèse du terme* (Aline Francoeur; en cours).

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

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

## **4. Personnel**

### **Étudiants**

**Suzie Baudry** (Bacc. en traduction). Rédaction de fiches dans le DiCoEnviro.

**Gabriel Bernier-Colborne** (maîtrise en traduction). Annotation d'un corpus étalon dans le domaine de l'automobile.

**Geneviève Camirand** (Bacc. en traduction). Projet RHWCD - DiCoInfo.

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

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

**Caroline Gagné** (Bacc. en traduction). Projet RHWCD - DiCoInfo.

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

**Julie Gariépy** (maîtrise en traductologie). Annotations de contextes pour le projet d'extraction et stockage de relations terminologiques à partir de bitextes.

**Marie-France Gignac** (Bacc. en traduction). Révision des entrées du lexique scientifique transdisciplinaire et établissement des équivalences.

**Charissa Harms** (maîtrise en traductologie). Annotations de contextes pour le projet d'extraction et stockage de relations terminologiques à partir de bitextes.

**Tetiana Kalinichenko** (maîtrise en traduction). Annotations contextuelles du DiCoEnviro français.

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

### **Stagiaires**

**Thibault Bughin**, Université Catholique de Louvain (Master en linguistique) : Détection automatique des rendez-vous dans un corpus de textos.

**María de la Nava Maroto García** (stage postdoctoral). Relations argumentales et circonstanciées entre verbes et noms dans le domaine de la céramique industrielle.

## **5. Autres activités**

### **5.1 Séminaires, colloques**

*Colloque scientifique dans le cadre du congrès de l'ACFAS : « Bibliothèques numériques : au carrefour des disciplines et des institutions »*. 80e congrès de l'ACFAS, Palais des congrès, Montréal, 8-10 mai 2012. (Responsable : Lyne Da Sylva ; coresponsable : Audrey Laplante, EBSI). Ce colloque a réuni divers acteurs concernés par le développement de bibliothèques numériques; une séance en particulier était dédiée aux aspects technologiques, incluant les technologies de traitement automatique de la langue utiles pour extraire des informations des documents textuels.

Organisation de l'atelier *TEXTOS : dimensions culturelles, linguistiques et pragmatiques* dans le cadre du colloque annuel de l'ACFAS.

*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**

**Rute Costa** (CLUNL, Universidade Nova de Lisboa)

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

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

### **5.3 Collaborations**

Lynne Bowker et Elizabeth Marshman (Université d'Ottawa) : Textto4Science.

Pascaline Dury (Centre de recherches en traduction en terminologie, Université Lyon-2, France) : Travaux sur l'identification automatique de la variation diachronique en terminologie.

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

Christian Guilbault (Simon Fraser University) : Textto4Science.

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.

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) : Projet d'évaluation utilisateur du DiCoInfo.

Aurélie Picton (Département de traitement informatique multilingue, Université de Genève) : Travaux sur l'identification automatique de la variation diachronique en terminologie.

## **6 Publications**

### **6.1 Livres ou chapitres de livres**

Alcina, A. 2011. Los traductores automáticos en la red. In *El español en el mundo. Anuario del Instituto Cervantes 2010-11*. Instituto Cervantes, Madrid.

Alcina, A. 2011. Teaching and Learning Terminology. New strategies and methods. In Alcina, A. (ed.). *Teaching and Learning Terminology. New strategies and methods*. John Benjamins, Amsterdam

Alcina, A. (ed.) 2011. *Teaching and Learning Terminology. New strategies and methods*. John Benjamins, Amsterdam.

Forest, D. & L. Da Sylva (eds.). 2011 Exploration de textes et recherche d'information, In *Canadian Journal of Information and Library Science*. 35:3, 222 p.

Francoeur, A., J. Cl. Boulanger, P. Auger, & J.-P. Roy (eds.) 2011. Une archéologie de la genèse et de l'histoire de la terminologie. Témoignages tirés de textes fondateurs. In *Langues et linguistique* 33, 254 p.

L'Homme, M.C. 2013, à paraître. Encoding of Terminological Data in Electronic Form", In Chapelle (ed.). *The Encyclopedia of Applied Linguistics*, C.A. Oxford, UK: Wiley-Blackwell.

- L'Homme, M.C., B. Robichaud & P. Leroyer. 2012, à paraître. Encoding collocations in DiCoInfo: From formal to user-friendly representations. In Granger, S. and M. Paquot (eds.). *Electronic Lexicography*, Oxford: Oxford University Press.
- Marengo, S. 2011. *Les adjectifs jamais attributs. Syntaxe et sémantique des adjectifs constructeurs de la référence*, Louvain-la-Neuve, De Boeck/Duculot.
- Marshman, E. 2012, accepté. La traductique. In Delisle, J. & M. Fiola (eds.). *La traduction raisonnée*, 3e édition. Presses de l'Université d'Ottawa.
- Mel'čuk, I. 2011. Phrasèmes dans le dictionnaire. In Anscombe, J.C. & S. Mejri (eds.). *Le figement linguistiques: la parole entravée*, Paris: Honoré Champion, 41-61.
- Mel'čuk, I. 2011. Word Order in Russian. In Boguslavskij, I., L. Iomdin & L. Krysin, (eds.). *Slovo i jazyk (Sbornik statej k vos'midesjatiletiju akademika Ju.D. Apresjana)*, Mel'čuk, I. (2011) Moskva: Jazyki slavjanskix kul'tur, 499-525.
- Mel'čuk, I. 2012. Bi-nominative Sentences in Russian. In Makarova, V. (ed.). *Russian Language Studies in North America*, London: Anthem, 85-106.
- Mel'čuk, I & J. Milićević. 2013, à paraître. *Introduction à la linguistique*. Paris: Hermann.
- Milićević, J. 2013, à paraître. Impersonal Constructions in Serbian, Meaning-Text Style. In Kor Chahine, I. & Ch. Zaremba (eds.). *Selected Papers from the 6th Annual Meeting of the Slavic Linguistic Society (Aix-en-Provence, septembre 2011)*. Aix-en-Provence: Presses de l'Université de Provence.
- Pastor, V. & A. Alcina. 2011. Acceso a la información terminológica en Internet. Técnicas para traductores. In Maruenda-Bataller, S. and B. Clavel-Arroitia (eds.). *Multiple Voices in Academic and Professional Discourse: Current Issues in Specialised Language Research, Teaching and New Technologies*. Cambridge Scholars Publishing, 243-256.
- Tremblay, O. & Mottet, M. 2012. Représentations du français oral standard et de la langue d'enseignement chez de futurs maîtres au primaire. In Bergeron, R. et G. Plessis-Bélair (eds.). *Représentations, analyses et descriptions du français oral, de son utilisation et de son enseignement au primaire, au secondaire et à l'université. Montréal*. Éditions Peisaj, 81-96.

## 6.2 Articles dans revues avec comité de lecture

- Ancil, D. 2013, à paraître. Portrait des erreurs lexicales d'élèves de 3e secondaire en production écrite et propositions de pistes didactiques. In *Pratiques*, 155-156, *Lexique et écriture*.



- Beck, D. & I. Mel'čuk. 2011. Morphological Phrasemes and Totonacan Verbal Morphology. In *Linguistics*, 49(1), 175-228.
- Bowker, L. & E. Marshman (2009, paru en 2012). Towards a model of active and situated learning in the teaching of computer-aided translation: Introducing the CERTT project. In *Journal of Translation Studies* 13 (1&2).
- Dury, P. & P. Drouin. 2011. When terms disappear from a specialized lexicon: A semi-automatic investigation into 'necrology'. In *ICAME Journal* 35, 19-33.
- Flórez, S. & A. Alcina. 2011. Catálogo de software libre para la traducción. *Tradumàtica*, 9, 57-73.
- Forest, D., & L. Da Sylva. 2011. Fouille de texte et recherche d'information / Text mining and information retrieval (Introduction). In *Revue canadienne des sciences de l'information et des bibliothèques*, Numéro spécial « Fouille de texte et recherche d'information ». 35(3), 217-228.
- Flórez, S. & A. Alcina. 2011. Free / Open-Source Software for the Translation Classroom. A Catalogue of Available Tools. In *The Interpreter and Translator Trainer*. 5(2), 325-357.
- Francœur, A. 2011. L'écrivain français et le dictionnaire dans son œuvre : objectivation, symbolisation, symbiose constante. In Glinoyer, A. & C. Paquette (eds.). *Mémoires du livre / Studies in Book Culture*. 2(2). [<http://www.erudit.org/revue/memoires/2011/v2/n2/1001763ar.html>]
- Iordanskaja, L. & I. Mel'čuk. 2011. Verbes parenthétiques illocutifs en russe. In Lee, S.H. (ed.). *Du lexique au dictionnaire en passant par la grammaire. Hommage à Chai-song Hong*, Paris: Garnier, 177-196. [= *Cahiers de lexicologie*, 2011-1, n°98.]
- L'Homme, M.C. 2013, à paraître. Adding syntactico-semantic information to specialized dictionaries: an application of the FrameNet methodology, In Heid, U. (ed.). *Lexicographica*.
- L'Homme, M.C. & G. Bernier-Colborne. 2012, à paraître. Terms as labels for concepts, terms as lexical units. In *Applied Ontology*. Numéro spécial : "Ontologies and Terminologies: Continuum or Dichotomy".
- Marshman, E., C. Harms & J. Gariépy. 2012, à paraître. Helping language professionals relate to terms: Terminological relations and termbases. In *JoSTrans* 18.

- Mel'čuk, I. 2012. Zavisimosti-2011: Otnošenje zavisimosti v jazyke i v lingvistike [= Relation de dépendance en langue et en linguistique]. In *Voprosy jazykoznanija*, 1, 3-26.
- Mel'čuk, I. & J. Milićević. 2011. "Budala jedna!"-Type Construction in Contemporary Serbian. *Journal of Slavic Linguistics*, 1, 19(1), 85-118.
- Mel'čuk, I & J. Milićević. 2011. The Budalo jedna!-Type Construction in Contemporary Serbian. In *Journal of Slavic Linguistics*. 19(1), 85-118.
- Pastor, V. & A. Alcina. 2010. Search Techniques in Electronic Dictionaries: a Classification for Translators. In *International Journal of Lexicography*. 23(3), 307-354.
- Pimentel, J., M.C. L'Homme & M.È. Laneville. 2012. General and Specialized Lexical Resources: a Study on the Potential of Combining Efforts to Enrich Formal Lexicons. *International Journal of Lexicography* 25(2), 152-190.

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

- Ancil, D. & O. Tremblay. 2013, à paraître. Au coeur du verbe : la structure sémantico-syntaxique. Quelques pistes didactiques. In *Colloque « Le complexe du verbe »*, IUFM de Lyon, France.
- Bernier-Colborne, G. 2012. Defining a Gold Standard for the Evaluation of Term Extractors. In *Actes de l'atelier ColabTKR du congrès LREC 2012*, Istanbul, Turquie.
- Bernier-Colborne, G. 2012. Application d'un algorithme de traduction statistique à la normalisation de textos. In *Actes de la 14e Rencontre des étudiants chercheurs en informatique pour le traitement automatique des langues (RECITAL 2012)*.
- Demers, M.C., I. Kernerman & M.C. L'Homme. 2012, à paraître. Lexicographic interchange between a specialized and a general language dictionary. In *Euralex 2012*, Oslo, Norway.
- Drouin, P., & G. Bernier-Colborne. 2011. Acquisition automatique de termes et lexique scientifique transdisciplinaire. In *Actes de la conférence Terminologie & Ontologie : Théories et applications (TOTh 2011)*, Annecy.
- Drouin, P. & F. Doll. 2011. Potentiel terminologique, quel sens prendre ? In *8<sup>e</sup> Journées scientifiques du Réseau LTT*, octobre 2009, Lisbonne, pp. 441-554.

- Drouin, P. & G. Bernier-Colborne. 2011, à paraître. Acquisition automatique de termes et lexique scientifique transdisciplinaire, In *Terminologie & Ontologie : Théories et applications* (TOTh 2011), mai 2011, Annecy.
- Estellés, A. & A. Alcina. 2012. Assessing the use of Ceramic Tile Standards for information extraction in terminological ontology building, In *Terminology and Knowledge Engineering* (TKE 2012), Madrid.
- Estellés, A. & A. Alcina. 2012, Prospects for Building a Ceramic Tiles Ontology based on ISO Standards information. In *Terminology & Ontology: Theories and applications* (TOTh 2012).
- Hadouche, F., S. Desgroseillers, J. Pimentel, M.C. L'Homme & G. Lapalme. 2011. Identification des participants de lexies prédicatives : évaluation en performance et en temps d'un système d'annotation automatique. In *Terminologie et intelligence artificielle* (TIA 2011), Institut national des langues orientales INALCO, Paris.
- Hadouche, F., G. Lapalme & M.C. L'Homme. 2011. Attribution de rôles sémantiques aux actants des lexies verbales. In *Traitement automatique des langues TALN*, Avignon.
- Iordanskaja, L. & I. Mel'čuk. 2011. Illocutive Parenthetical Verbs in Russian. In Boguslavsky, I. and L. Wanner (eds.). *Proceedings of the 5<sup>th</sup> International Conference on Meaning-Text Theory*, Barcelona, 120-133.
- Jousse, A.L., M.C. L'Homme, P. Leroyer & B. Robichaud. 2011. Organizing collocates in a dictionary of computing and the Internet according to user needs. In Boguslavsky, I. and L. Wanner (eds.). *Fifth Meaning-Text Theory Conference* (MTT 2011). Barcelone: Université Pompeu Fabra.
- Laroche, A., P. Drouin & G. Bernier-Colborne. 2011. Étude de l'influence de la taille du corpus de référence sur l'extraction terminologique automatique contrastive. *Proceedings of the 9th International Conference on Terminology and Artificial Intelligence* (TIA). Paris, 66-72
- L'Homme, M.C. 2012, à paraître. Le verbe terminologique : un portrait de travaux récents. In *Actes du Congrès mondial de linguistique française* (CMLF 2012), Université Lyon-2, Lyon.
- L'Homme, M.C. & J. Pimentel. 2012. Capturing syntactico-semantic regularities among terms: An application of the FrameNet methodology to terminology. In *Language Resources and Evaluation, LREC 2012*. Istanbul, Turkey.

- Le Serrec, A. 2012, à paraître. Étude comparative de l'équivalence terminologique en corpus parallèle et en corpus comparable : application au domaine du changement climatique. In *Actes du colloque international « La notion d'unité en sciences du langage : aspects lexicologiques, terminologiques et traductologiques »* (9es Journées scientifiques du réseau Lexicologie, Terminologie, Traduction). Université Paris 13 (Villetaneuse, France).
- Marengo, S. 2012, à paraître. Unités linguistiques et parties du discours : pour une grammaire à l'écoute de la lexicologie, In *Actes du colloque international « La notion d'unité en sciences du langage : aspects lexicologiques, terminologiques et traductologiques »* (9es Journées scientifiques du réseau Lexicologie, Terminologie, Traduction). Université Paris 13 (Villetaneuse, France).
- Marengo, S. 2012, à paraître. Complémentation sémantique et syntaxique : quatre cas de figure pour l'adjectif en français. In Dan Van Raemdonck et al. (ed.), *Actes du colloque international « ComplémentationS »*, Bruxelles, Peter Lang.
- Marsac, F. & S. Marengo. 2012, à paraître. L'accord du participe passé des verbes de perception régissant une construction infinitive : reconsidération en vue d'une représentation formelle en Théorie Sens-Texte, In *Actes du colloque international « Le participe passé entre accords et désaccords »*. 16-18 juin 2011, Université d'Opole (Pologne).
- Maroto, N., M.-C. L'Homme & A. Alcina. 2012. Semantic Relations Established by Specialized Processes Expressed by Nouns and Verbs: Identification in a Corpus by means of Syntactico-semantic Annotation. In *Language Resources and Evaluation (LREC 2012)*, Istanbul.
- Mel'čuk, I. 2011. Dependency in Language-2011. In Gerdes, K., E. Hajičová, L. Wanner (eds.). *International Conference on Dependency Linguistics: Depling 2011*, Barcelona, 1-16.
- Milićević, J. & A. Tsedryk. 2011. Assessing and Improving the Paraphrasing Competence of FLS Learners, in Boguslavskij, I. & Wanner, L. (eds). *Proceedings of the 5th International Conference on the Meaning-Text Theory (MTT2012)*, pp. 176-185.
- Ortego, T. 2011. La compilation de DiCoEnviro en espagnol. In *IIIe Congrès International de Linguistique de Corpus (CILC)*, Universidad Politécnica de Valencia.
- Pimentel, J. 2012. à paraître. Description de verbes juridiques au moyen de la sémantique des cadres. In *Terminologie & Ontologie : Théories et applications* (Toth 2011), Annecy 2011, 26-27 mai 2011.

Pimentel, J. 2012. Identifying equivalents of specialized verbs in a bilingual corpus of judgments: A frame based methodology. In *Language Resources and Evaluation* (LREC 2012). Istanbul, Turkey.

Robichaud, B. 2011. A Graph Visualization Tool for Terminology Discovery and Assessment. In Boguslavskij, I. & Wanner, L., eds, *Proceedings of the 5th International Conference on the Meaning-Text Theory* (MTT2012), pp. 252-261.

Robichaud, B. 2012. Logic Based Methods for Terminological Assessment, In *Language Resources and Evaluation* (LREC 2012). Istanbul, Turkey.

#### **6.4 Communications non publiées**

Bernier-Colborne, G. 2012. Évaluer les extracteurs de termes : l'élaboration d'un corpus étalon. *Séminaires OLST-RALI-Mprime*, Université de Montréal, Montréal.

Da Sylva, L. 2011. Extracting a vocabulary for specialized indexing: comparison among three corpora. In *Corpus Linguistics Conference*, University of Liverpool, Liverpool.

Da Sylva, L. Bibliothèques numériques et traitement automatique de la langue, *Séminaires OLST-RALI-Mprime*, Université de Montréal, Montréal.

Francœur, A. 2011. Policies and Choices in Dictionary Adaptation: How Noël Chomel's Dictionnaire économique (1718) and Pierre Joseph Macquer's Dictionnaire de chymie (1766) were translated into English. In *18th Biennial Conference of the Dictionary Society of North America*, McGill University, Montréal.

Francœur, A. 2011. Les méthodes de travail enseignées à l'université. In *X<sup>e</sup> Symposium du Conseil fédéral de terminologie*, Gatineau.

Francoeur, A. et G. Tremblay. 2012. Terminographie "orientée objet" : le *Dictionnaire descriptif et visuel d'objets / Descriptive and Visual Dictionary of Objects* de Parcs Canada. *Séminaire OLST-RALI*, Département de linguistique et de traduction, Université de Montréal.

L'Homme, M.C. 2011. Collocations dans le DiCoInfo : des représentations formelles aux présentations conviviales. *Séminaires du CIRAL*, Université Laval, Québec.

L'Homme, M.C. 2011. Collocations dans le DiCoInfo avant et après. *Séminaires OLST-RALI-Mprime*, Université de Montréal, Montréal.

- Le Serrec, A. 2011. Étude comparative de l'équivalence terminologique en corpus parallèle et en corpus comparable : application au domaine du changement climatique. *Séminaires OLST-RALI-Mprime*, Université de Montréal, Montréal.
- Le Serrec, A. 2012 Étude comparative de l'équivalence terminologique en corpus parallèle et en corpus comparable : application au domaine du changement climatique. *Colloque étudiant : Courants actuels en terminologie et technologies de la traduction*. Université d'Ottawa. Ottawa.
- Lefrançois, P. et M.-É. de Villers. 2011. Le vocabulaire des élèves québécois du secondaire – du portrait diagnostique aux pistes d'intervention didactique. In *79e congrès de l'ACFAS*, Sherbrooke, 11 mai 2011.
- Lefrançois, P. et M.-É. de Villers. 2011. Le vocabulaire des élèves québécois de 3e secondaire – un portrait documenté. In *Congrès de l'Association québécoise des professeurs de français (AQPF)*, Trois-Rivières, 3 novembre 2011.
- Maroto, N. 2011. Conceptual relations established by processes: the case of firing in industrial ceramics. In *Séminaires OLST-RALI-Mprime*, Université de Montréal, Montréal.
- Maroto, N. 2011. Conceptual relations established by processes: the case of cocción (firing) in industrial ceramics. In *Séminaires OLST-RALI-Mprime*, Université de Montréal, Montréal.
- Marshman, E. & C. Harms. 2011. Translation of lexical markers of terminological relations: Observations in English-French bitexts. In *18th European Symposium on Language for Special Purposes*, Perm, Russia, 22-26 August 2011.
- Marshman, E. 2012. Canadian Translation Students and Language Technologies: Observations from the LinguisTech Project. In *Congrès de l'Association canadienne de traductologie*. Waterloo, Ontario, 31 mai 2012.
- Marshman, E. 2012. Powering translation and empowering translators: A dual role for language technologies? In *4<sup>th</sup> IATIS Conference*, Belfast, Northern Ireland, 22-29 juillet 2012.
- Marshman, E. 2012. C is for terminology. In *Journées d'études sur l'avenir de la terminologie*. Comité mixte de terminologie au Canada. Gatineau, Québec, 7 mars 2012.
- Pastor, V. & A. Alcina. 2012, Clasificación de técnicas de búsqueda en diccionarios electrónicos. In *Congreso de Lexicografía*, Madrid.

Pimentel, J. 2011. Méthodologie de description de verbes juridiques basée sur la sémantique des cadres. In *Séminaires OLST-RALI-Mprime*, Université de Montréal, Montréal.

Pimentel, J. 2011. Criteria for the validation of specialized verb equivalents: Applications in bilingual terminography. In *Colloque étudiant : Courants actuels en terminologie et technologies de la traduction*. Université d'Ottawa. Ottawa.

Robichaud, B. 2011. Le DiCoInfo Visuel. In *Séminaires OLST-RALI-Mprime*, Université de Montréal, Montréal.

## 6.5 Autres

Anctil, D. et al. 2012. Projet « Ouvrir le dictionnaire ».

Site Web : <http://www.ccdmd.qc.ca/fr/dictionnaire>.

L'Homme, M.C. 2012. *DiCoEnviro. Le dictionnaire fondamental de l'environnement*.

Site Web : [http://olst.ling.umontreal.ca/cgi-bin/dicoenviro/search\\_enviro.cgi](http://olst.ling.umontreal.ca/cgi-bin/dicoenviro/search_enviro.cgi)

Lefrançois, P. et M.-É. de Villers. 2012. *Enquête sur les connaissances lexicales des jeunes Québécois francophones*. Rapport de recherche (non publié) soumis au ministère de l'Éducation, du Loisir et du Sport.

Pimentel, J. 2011. [Compte rendu de l'ouvrage : Philippe Gréciano & John Humbley (dir.) 2011. *Langue et droit : terminologie et traduction*. Vol. XVI, 2011-1 de la *Revue Française de Linguistique Appliquée*. 136 p.]. In *Hermes – Journal of Language and Communication in Business*. 47, pp. 115-119

Pimentel, J. & B. Robichaud. 2012. JuriDiCo : Base de données terminologique portant sur le domaine du droit.

Site Web : <http://olst.ling.umontreal.ca/cgi-bin/juridico/search.cgi>

Robichaud, B. 2011. *Le DiCoInfo Visuel : navigation visuelle dans le DiCoInfo*.

Site Web : <http://olst.ling.umontreal.ca/dicoinfo/visuel.php>

Tremblay, O. 2012. La chanson pour travailler l'oral et l'écrit en classe de français. In *Québec Français* 165, 51-53.