

Observatoire de linguistique Sens-Texte

OLST

Rapport d'activités

Mai 2020 – avril 2021

Table des matières

| | |
|---|-----------|
| 1. Membres réguliers et associés | 4 |
| Membres réguliers..... | 4 |
| Membres associés | 4 |
| Membre adjoint..... | 4 |
| 2. Projets et état d'avancement des travaux | 5 |
| 2.1. Applications documentaires..... | 5 |
| 2.2. Dictionnaires et bases de données lexicales | 5 |
| 2.3. Analyses de données lexicales..... | 5 |
| 2.4. Corpus..... | 7 |
| 2.5. Traitement automatique de la langue..... | 7 |
| 2.6. Didactique du lexique | 8 |
| 2.7. Représentation et visualisation de données | 10 |
| 3. Encadrement d'étudiants aux études supérieures | 10 |
| 4. Personnel..... | 13 |
| 4.1. Étudiants..... | 13 |
| 4.2. Stagiaires..... | 15 |
| 4.3. Professionnels de recherche | 15 |
| 5. Autres activités | 15 |
| 5.1 Séminaires et colloques | 15 |
| 5.2 Professeurs, chercheurs invités et séjours de recherche | 15 |
| 5.3 Collaborations | 15 |

| | | |
|----------|---|-----------|
| 6 | Publications et communications..... | 17 |
| 6.1 | Livres et chapitres de livres..... | 17 |
| 6.2 | Articles dans des revues avec comité de lecture | 18 |
| 6.3 | Articles dans des actes de congrès..... | 20 |
| 6.4 | Communications non publiées | 20 |
| 6.5 | Articles dans des revues professionnelles..... | 21 |
| 6.6 | Autres..... | 21 |

1. Membres réguliers et associés

Membres réguliers

Dominic Anctil

Lyne Da Sylva

Patrick Drouin (directeur)

Lidjia Iordanskaja

Anna Joan Casademont (Département Sciences humaines, Lettres et
Communication, TÉLUQ)

François Lareau (directeur adjoint)

Pascale Lefrançois

Marie-Claude L'Homme

Sébastien Marengo

Elizabeth Marshman (École de traduction et d'interprétation, Université d'Ottawa)

Igor Mel'čuk

Jasmina Milićević (Department of French, Dalhousie University)

Antonio San Martín Pizarro (Département des langues modernes et de traduction, UQTR)

Ophélie Tremblay (Département de didactique des langues – UQAM)

Antoine Venant

Fabienne Venant (Département de mathématiques – UQAM)

Membres associés

Aurélien Picton (TIM, Faculté de traduction et d'interprétation, Université de Genève)

Alain Polguère (Université de Lorraine, CNRS, ATILF)

Membre adjoint

Benoît Robichaud

2. Projets et état d'avancement des travaux

Cette section décrit les projets spécifiques sur lesquels les chercheurs de l'OLST ont travaillé durant la période s'étendant de mai 2020 à avril 2021.

2.1. Applications documentaires

Entités et identité dans les jeux de données du web sémantique : une exploration sémiotique des enjeux pour les milieux documentaires (responsable : Lyne Da Sylva). Subventionné par le CRSH, le projet de recherche porte sur le web sémantique et ses données ouvertes liées (« Linked Open Data » ou LOD). L'objectif est d'étudier les descriptions du web sémantique à la lumière des besoins et des attentes des milieux documentaires, et vise à formuler des recommandations sur les utilisations propices de la technologie. Une attention particulière sera portée à la synonymie, soit l'identification d'une seule entité à l'aide d'identifiants distincts. Ce projet n'a guère avancé cette année, dû aux responsabilités administratives de la chercheuse.

2.2. 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 et fournit une description détaillée des propriétés lexico-sémantiques des termes. En 2020-2021, des termes chinois et espagnols ont été ajoutés. Les versions italiennes et portugaises sont également enrichies dans le cadre de collaborations avec l'Université de Vérone et l'Université de Brasilia.

Lexitrans (responsable : Patrick Drouin). 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). Au cours de l'année, l'ensemble des entrées françaises et anglaises ont été enrichies avec des contextes. Toutes les fiches en français ont été revues afin de raffiner les acceptions ainsi que la très grande majorité des fiches anglaises.

2.3. Analyses de données lexicales

Analyse/Étude de la variation dialectale en langues de spécialité (responsables : Aurélie Picton et Patrick Drouin). La prise en compte de la variation en langues de spécialité s'est largement développée, mais peu de travaux traitent de plusieurs types de variations simultanément : « multi-variation ». Les travaux dans le cadre de ce projet se sont concentrés cette année sur l'exploration de techniques d'identification et de visualisation de la variation dans le domaine de l'aide humanitaire, en particulier à l'aide du langage

de programmation R. Cet axe rejoint également cette année le projet « Patients partenaires » des HUG (Hôpitaux Universitaires Genevois), avec le développement d'un projet de thèse sur l'analyse des termes et de leur circulation en diabétologie, chez les spécialistes et les patients dits « experts » (patients atteints de la même maladie chronique et formés à l'accompagnement d'autres patients).

Annotation syntactico-sémantique de contextes spécialisés (responsable : Marie-Claude L'Homme). Ce projet (initialement financé par le Conseil de recherches en sciences humaines, CRSH) vise à fournir, pour des termes du DiCoInfo et du DiCoEnviro, une annotation syntactico-sémantique. L'annotation portait initialement sur les termes de nature prédicative ainsi que sur leurs actants et circonstants. Elle est faite désormais pour des termes non prédicatifs et des termes quasi-prédicatifs. L'équipe a mis au point un protocole d'annotation réalisé dans une structure XML. Jusqu'à présent, environ 30 000 contextes sont annotés en français, en chinois, en portugais, en espagnol et en anglais. Le travail d'annotation est utilisé dans la découverte de « frames » dans les deux ressources terminologiques.

Découverte de frames sémantiques dans le domaine de l'environnement (responsable : Marie-Claude L'Homme). Ce projet (initialement financé par le Conseil de recherches en sciences humaines, CRSH) vise à découvrir des cadres sémantiques valables pour le domaine de l'environnement. On procède d'abord par une comparaison avec le contenu de FrameNet pour tenter d'établir des correspondances. Les unités lexicales anglaises apparaissant à la fois dans le DiCoEnviro et dans FrameNet sont identifiées automatiquement. Lorsque aucune correspondance n'est possible, de nouveaux frames sont définis. Nous dégageons également des relations entre frames pour construire de petits scénarios propres au domaine de l'environnement. Les résultats de ce travail peuvent être visualisés dans une interface appelée « A Framed Version of DiCoEnviro ». En 2020-2021, des termes chinois et portugais ont été ajoutés au *Framed DiCoEnviro*.

Développement d'une méthodologie d'élaboration de définitions terminologiques : analyse de corpus et variation contextuelle (responsable : Antonio San Martín). Ce projet (financé par le FRQSC, Soutien à la recherche pour la relève professorale) est motivé par un problème habituel de l'élaboration des définitions terminologiques : le rôle du contexte (tout facteur affectant l'interprétation d'un signe) est souvent ignoré. En outre, les effets du contexte en terminologie sont peu connus. En particulier, un phénomène central, mais peu étudié, est celui de la variation contextuelle interdomaine, lequel se produit lorsqu'un concept n'active pas toujours les mêmes caractéristiques lors de son usage dans différents domaines de connaissances. Le projet vise le développement d'une méthodologie d'élaboration de définitions terminologiques basée sur l'analyse de corpus qui inclut des règles pour la détection, la caractérisation et la gestion de la variation contextuelle interdomaine.

Élaboration d'une liste de vocabulaire transdisciplinaire pour le primaire (responsables : Ophélie Tremblay et Patrick Drouin). Ce projet a consisté à élaborer une liste de vocabulaire transdisciplinaire à partir d'un corpus de manuels scolaires recommandés par le Ministère de l'Éducation et de l'Enseignement Supérieur du Québec. Le corpus de manuels a été analysé pour générer une première version de la liste lexicale. La liste contient 1389 noms, 716 verbes, 381 adjectifs et 146 adverbes (2632 unités lexicales). Des activités didactiques sont également en cours d'élaboration.

2.4. Corpus

Enrichissement des corpus spécialisés (responsables : Marie-Claude L'Homme). Le corpus d'environnement espagnol a été enrichi de textes portant sur les pêcheries durables.

Néonum : infrastructure de veille pour la néologie en français québécois (responsables Patrick Drouin, François Lareau et Philippe Langlais). Ce projet (financé par le Conseil de recherches en sciences humaines, CRSH) a pour but de mettre en place une première infrastructure exploitable pour l'identification rapide et systématique de néologismes en français québécois. Le nettoyage du corpus obtenu du projet CO.SHS s'est poursuivi jusqu'à l'automne 2020.

2.5. Traitement automatique de la langue

Acquisition automatique de la terminologie : TermoStat (responsable : Patrick Drouin, Benoît Robichaud). Le logiciel d'acquisition automatique de terminologie TermoStat est désormais offert en version multilingue (anglais, espagnol, français, italien et portugais) grâce à une interface Web. Le travail débuté en vue d'une refonte de l'interface Web s'est poursuivi. Il s'inscrit dans un mouvement qui vise à séparer le code du moteur d'extraction des termes de l'interface du logiciel afin d'en faciliter la gestion et de mettre en place des API permettant d'exploiter TermoStat sans devoir interagir avec l'interface Web. Une version 4.0 du logiciel est en période de test, mais elle reste à compléter.

Extraction de relations sémantiques en français sous forme de word sketches : EcoLexicon Semantic Sketch Grammar – French (ESSG-fr) (responsable : Antonio San Martín). Ce projet, subventionné par le Fonds institutionnel de recherche de l'UQTR, a pour objectif de créer une grammaire de sketches française pour Sketch Engine basée sur des patrons de connaissances. L'ESSG-fr permettra d'extraire des relations sémantiques en français à partir de tout corpus hébergé dans Sketch Engine. L'extraction des paires hyponymiques est déjà disponible.

Génération automatique de définitions lexicographiques : Define (responsable : François Lareau). Il s'agit d'un projet financé par le CRSH (Développement Savoir) qui vise à adapter GenDR à la génération automatique de définitions lexicographiques dans le Réseau Lexical du français (ATILF). Le but à plus long terme est d'intégrer toutes les

données du RF-fr dans GenDR et d'utiliser le générateur comme outil de validation des données lexicographiques.

Génération automatique de texte multilingue : GenDR (responsable : François Lareau). GenDR est un réalisateur profond multilingue qui modélise l'interface sémantique-syntaxe pour la réalisation linguistique en génération de texte. Cette année, Linna He a déposé un mémoire sur la construction d'un VerbNet du mandarin pour GenDR. Deux autres mémoires sont en cours, sur le traitement des locutions et des relatives.

Néonum : infrastructure de veille pour la néologie en français québécois (responsables : Patrick Drouin, François Lareau et Philippe Langlais). Dans le cadre du projet Néonum, des premiers tests de dépistage des néologismes ont été effectués afin de valider l'utilité des corpus mis en place et d'orienter les travaux de la prochaine année. Des résultats prometteurs laissent déjà penser que l'infrastructure pourra rencontrer nos objectifs de recherche. L'année a été principalement consacrée à l'élaboration d'algorithmes pour l'identification des néologismes sémantiques en langue générale. Une analyse d'un corpus de langue spécialisée (la revue *Méta*) a été faite l'objet d'une analyse afin d'y observer la néologie lexicale sur une période de 60 ans.

Jeux de données artificiels pour évaluer l'apprentissage de structures latentes (responsable : Antoine Venant). Supposons, par exemple, qu'on souhaite entraîner un système à produire des représentations sémantiques. On peut vouloir produire d'abord une structure syntaxique arborescente, pour ensuite exploiter cette structure et produire une représentation sémantique. Toutefois, guider l'apprentissage explicitement à l'aide d'annotations syntaxiques est très coûteux, et séparer l'apprentissage de la structure syntaxique de l'utilisation qui en est faite ensuite peut mener à propager certaines erreurs en cascade, en raison de biais théoriques ou pratiques. Une solution attrayante pour éviter ces problèmes est d'apprendre simultanément à produire une structure arborescente (latente) et à l'utiliser, sans autre forme de guidage que la tâche finale à résoudre (la production de représentation sémantique dans notre exemple). Mais lorsque cela échoue, il est difficile de déterminer si la cause en est algorithmique ou théorique. Ce nouveau projet, financé par le CRSNG (Découverte) vise à créer et utiliser des jeux de données artificiels pour évaluer plus précisément ces techniques d'apprentissage de structure arborescentes latentes. Cela devrait permettre de contrôler : 1) l'existence théorique d'une structure arborescente optimale, 2) les propriétés du langage 'retenues' dans la tâche artificielle cible, et 3) la quantité de données disponible pour l'apprentissage, et conséquemment de tirer de meilleures conclusions sur les propriétés de ces méthodes.

2.6. Didactique du lexique

Des séquences d'intervention pour soutenir l'apprentissage de mots ciblés en classe de français auprès d'élèves de 4e secondaire éprouvant des difficultés langagières (chercheur principal : Dominic Anctil, cochercheuse : Rachel Berthiaume,

collaboratrices : Vanessa Dandavino et Isabelle Ducharme) : Ce projet *CRSH – Développement Savoir* (2019-2021) s’inscrit dans le cadre d’un partenariat entre la Faculté d’éducation de l’Université de Montréal et l’École Vanguard, qui a pour mission d’agir auprès d’élèves manifestant des difficultés graves d’apprentissage. Les élèves éprouvant des troubles du langage peinent à comprendre et retenir les nouveaux mots auxquels ils sont exposés et n’arrivent pas toujours à utiliser un vocabulaire précis, tant à l’oral qu’à l’écrit, à cause d’une connaissance peu profonde des mots et de difficultés d’accès lexical. Le projet vise à tirer profit des travaux sur l’enseignement direct du vocabulaire pour vérifier en quoi ils peuvent contribuer à l’accroissement du vocabulaire d’élèves avec difficultés langagières en fin de parcours secondaire. Dans une approche collaborative, trois enseignantes de 4^e secondaire de Vanguard seront accompagnées par l’équipe de recherche dans l’élaboration et la mise à l’essai de séquences d’interventions sémantiques et phonologiques visant l’apprentissage d’un vocabulaire ciblé dans des textes lus en classe. Les apprentissages lexicaux des élèves seront mesurés pour vérifier l’efficacité des séquences proposées.

Développer l’habileté à inférer, le vocabulaire et la sensibilité lexicale pour soutenir l’apprentissage de la lecture : une recherche collaborative avec des enseignantes de maternelle et de 1^{re} année du primaire intervenant en milieu défavorisé (chercheur principal : Dominic Anctil, cochercheuses : Ophélie Tremblay et Andréanne Gagné — UQAM, collaboratrices : Dominique Chouinard et Lude Pierre) : Ce projet CRSH — Développement de partenariat (2020-2023) est mené en collaboration avec le programme Une école montréalaise pour tous du ministère de l’Éducation. Il vise à répondre aux besoins des élèves issus de milieux défavorisés qui, dès le début de leur parcours scolaire, sont plus à risque de connaître des difficultés dans leur apprentissage de la lecture si l’école n’intervient pas pour développer leur vocabulaire et leurs habiletés de base en compréhension. Dans une approche collaborative, deux cohortes (maternelle et 1^{re} année) de 14 enseignantes et de 4 conseillères pédagogiques recevront une formation visant l’appropriation de nouvelles pratiques par un accompagnement dans a) la planification et la mise en œuvre d’interventions pour travailler la compréhension et le vocabulaire à partir de la lecture interactive d’œuvres jeunesse et b) la conception d’activités de consolidation de vocabulaire ciblé et de développement de la sensibilité lexicale. Les apprentissages lexicaux chez les élèves seront aussi mesurés.

Identification, description et enseignement du vocabulaire transdisciplinaire en usage au primaire (responsable : Ophélie Tremblay; Patrick Drouin et Fabienne Venant : cochercheurs) Ce projet vise à identifier et décrire les mots de ce vocabulaire transdisciplinaire, à travers l’analyse d’un corpus de manuels scolaires en usage au Québec, de la 1^{ère} à la 6^e année. La description des mots retenus, sous la forme de réseaux lexicaux, et la constitution de corpus d’exemples illustrant les emplois des différents sens des mots ciblés soutiendront l’élaboration d’une démarche innovante d’enseignement du vocabulaire transdisciplinaire. La validation de cette démarche par des didacticiens des disciplines (français, mathématiques, sciences, histoire, géographie et éducation à la citoyenneté, éthique et culture religieuse) et des enseignants (un pour

chacun des 6 niveaux du primaire) permettra d'évaluer, de façon exploratoire, le potentiel des réseaux lexicaux et de l'utilisation de corpus d'exemples pour l'enseignement et l'apprentissage du vocabulaire (en français et dans les autres matières).

Pratiques d'utilisation des dictionnaires papier et électronique dans les classes du primaire et du secondaire en français langue première au Québec (responsable : Ophélie Tremblay). Ce projet, subventionné par le FRQSC, Établissement de nouveaux professeurs-chercheurs (2014-2017), vise à dresser un portrait des pratiques déclarées d'enseignement de l'utilisation du dictionnaire (papier et électronique) en classe, au primaire et au secondaire en français langue d'enseignement

2.7. Représentation et visualisation de données

Données du patrimoine culturel québécois : représentation et visualisation (responsables : Lyne Da Sylva et Thomas Hurtut, de l'École Polytechnique de Montréal). Ce projet mandaté et subventionné par le MCCQ s'est déroulé en trois phases : (i) des données descriptives d'œuvres de différents organismes (essentiellement des musées québécois et de Bibliothèque et Archives nationales du Québec) ont été colligées à partir des organismes individuels ; (ii) les données ont été encodées en données ouvertes et liées (formalisme RDF), sur la base d'une modélisation à l'aide de l'ontologie CIDOC-CRM (CIDOC Conceptual Reference Model) ; (iii) enfin, un outil d'exploration et de visualisation de ces données a été développé. Initialement prévu pour se terminer en mars 2021, le projet s'est prolongé jusqu'au 31 mai 2021.

3. Encadrement d'étudiants aux études supérieures

Desislava Aleksandrova (2e cycle, linguistique). *La détection automatique d'énoncés biaisés* (François Lareau, terminé).

Elham Aryapour (3e cycle, didactique). *L'anxiété langagière d'apprenants adultes des cours de francisation au Québec en situation de communication orale : ses manifestations en classe et hors classe* (Dominic Anctil, en cours).

Silvia Calvi (3e cycle, linguistique). *Caractérisation et extraction des unités polylexicales dans le domaine du droit du commerce international*, Université de Vérone (Patrick Drouin codirecteur, en cours).

Ménaïc Champoux-Chouinard (3e cycle, didactique). *Développement d'un test adaptatif informatisé visant à faire le portrait des connaissances grammaticales explicites des élèves de secondaire* (Pascale Lefrançois, en cours).

Bruno Charrette (2e cycle, didactique). *Les notions de « prédicat sémantique » et de « régime verbal » pour soutenir l'analyse syntaxique au 3e cycle du primaire* (Dominic Anctil, en cours).

Myriam Circé (3e cycle, didactique). *Conception et mise à l'essai d'une méthode itérative biphase pour le développement de l'expertise en écriture des cégépiens* (Pascale Lefrançois, en cours).

Ludivine Claus (2e cycle, traduction). *Contribution à l'élaboration de l'Encyclopédie humanitaire : mesures des apports et limites de l'approche en corpus de la terminologie textuelle* (titre provisoire) (Aurélie Picton, en cours).

Hubert Corriveau (2e cycle, linguistique). *De FrameNet à la Théorie Sens-Texte : Conversion et correspondance* (François Lareau, en cours).

Jessica Cyr (2e cycle, traductologie). *Les technologies langagières dans les cours de traduction : Un bilan canadien* (Elizabeth Marshman, en cours).

Michaëlle Dubé (2e cycle, linguistique). *Le traitement des locutions en génération automatique de texte multilingue* (François Lareau, en cours).

Catherine Fleurent (3e cycle, didactique). *Étude des constructions syntaxiques problématiques chez les étudiants postsecondaires et mise en place d'un module d'enseignement de la révision* (Pascale Lefrançois, en cours).

Grigorios, Giannakis (2e cycle, traduction). *Création de ressources terminologiques gouvernementales. Comparaison Suisse-Québec dans le domaine du changement climatique* (titre provisoire) (Aurélie Picton, en cours).

Nicolás Gonzales Granado (2e cycle, traduction). *Contribution à l'Encyclopédie humanitaire : analyse en corpus comparables de concepts entre organisations* (Aurélie Picton, terminé)

Elizabeth Haikings (2e cycle, linguistique) *L'infinitif en espagnol et en français à travers Cien años de soledad / Cent ans de solitude* (Jasmina Milićević, en cours).

Linna He (2e cycle, linguistique). *Un dictionnaire de régimes verbaux en mandarin* (François Lareau, terminé).

Julie Humbert-Droz (3 cycle, traitement informatique multilingue). *Circulation des termes entre langues de spécialité et langue générale : proposition d'un cadre théorique et méthodologique d'analyse du phénomène de la déterminologisation* (co-direction : Aurélie Picton et Anne Condamines, Université Toulouse Jean Jaurès, en cours).

Gabrielle Joli-Coeur (2e cycle, didactique). *L'apprentissage du vocabulaire par la littérature jeunesse chez les élèves allophones du primaire : l'impact de la culture première dans les œuvres lues* (Dominic Anctil, en cours).

Zhiwei Han (3e cycle, traduction). *Description du lexique spécialisé chinois pour locuteurs non sinophones : apport de la Lexicologie explicative et combinatoire pour l'acquisition de compétences lexicales* (Marie-Claude L'Homme, en cours).

Sylvie Marcotte (3e cycle, didactique). *Des stratégies pédagogiques utilisées en classe de français pour développer la compétence scripturale des élèves* (Pascale Lefrançois, terminé).

Montalvan Ayala, Luz de Maria (2e cycle, traduction). *Analysis of terminological definitions in fisheries sustainability* (co-direction : Marie-Claude L'Homme et Antonio San Martín, Université du Québec à Trois-Rivières, en cours).

Cheryl McBride (3e cycle traductologie). *User experience in translation technologies* (Elizabeth Marshman, en cours).

Fiona Meister (2^e cycle, linguistique). *Cohérence discursive et implicatures conversationnelles : analyses empiriques et théoriques vers un modèle pragmatique bayésien à l'échelle de la conversation*. (Antoine Venant, en cours).

Julia Njuguna (2e cycle, traductologie). *Terminology work in Kenya* (Elizabeth Marshman, en cours).

Charlotte Portenseigne (2e cycle, linguistique). *Le traitement des relatives dans un réalisateur profond* (François Lareau, en cours).

Caroline Proulx (3e cycle, didactique). *Élaboration et mise à l'essai d'un répertoire de dispositifs et de tâches intégrés à un réseau littéraire pour soutenir les enseignants dans l'enseignement lexical* (co-direction : Dominic Anctil et Isabelle Montesinos-Gelet, en cours).

Shahrazad Rahme (3e cycle, sciences de l'information). *Indexation automatique et repérage des images numériques* (titre provisoire) (Lyne Da Sylva, en cours).

Judith Rémillard (3e cycle traduction). *Intégration de la traduction automatique dans l'environnement de traductologie* (Elizabeth Marshman, en cours).

Marielle St-Germain (3e cycle, sciences de l'information). *Étude de l'impact des ontologies sur la visualisation de données patrimoniales dans un contexte de recherche exploratoire* (titre provisoire) (Lyne Da Sylva, en cours).

Claudine Sauvageau (passage accéléré au 3e cycle en septembre 2018, didactique). *L'oral réflexif entre pairs pour favoriser la consolidation du vocabulaire rencontré en lecture chez des élèves à risque du 1er cycle du primaire* (Dominic Anctil, en cours).

Allison Stentaford (2e cycle, traductologie). *The DiCoEnviro, Specialized Terminology and Renewable Energies: Assessing Student Translators' Comprehension and Usage of Specialized Terms* (Elizabeth Marshman, en cours).

Kizito Tekwa (3e cycle, traductologie). *Real-Time Machine-Translated Instant Messaging: A New Technology-based Approach Towards Improving Second Language Willingness to Communicate* (co-direction : Elizabeth Marshman et Marie-Josée Hamel, en cours).

Emma Thibert (3e cycle, traitement informatique multilingue). *Circulation et appropriation terminologique chez les patients experts : une analyse des liens entre expertise et terminologie chez les patients diabétiques* (co-direction : Aurélie Picton et Valérie Delavigne, en cours)

Yeng Zing (2e cycle, traduction). *Frame Semantics for the field of climate change: discovering frames based on Chinese and English terms* (co-direction : Marie-Claude L'Homme et Jian Yun Nie, Département d'informatique et de recherche opérationnelle, en cours).

4. Personnel

4.1. Étudiants

Clémence Bideaux (maîtrise en lettres, UQTR). Contribution au projet *Développement d'une méthodologie d'élaboration de définitions terminologiques : analyse de corpus et variation contextuelle*.

Sarah Choukeir (baccalauréat en traduction). Relecture et correction de documents numérisés pour le projet *Identification, description et enseignement du vocabulaire transdisciplinaire en usage au primaire*.

Florence Bois-Villeneuve (baccalauréat en traduction). Relecture et correction de documents numérisés dans le projet d'élaboration d'un lexique de l'ébénisterie.

Hubert Corriveau (maîtrise en linguistique). Génération automatique de texte ; annotations à la FrameNet.

Jessica Cyr (maîtrise en traductologie). Construction du corpus de l'environnement.

Michaëlle Dubé (maîtrise en linguistique). Génération automatique de texte.

Zhiwei Han (3e cycle, traduction). Encodage de liens lexicaux dans le DiCoEnviro chinois.

Linna He (maîtrise en linguistique). Génération automatique de texte.

Natasha Herger (DESS en traduction, UQTR). Contribution au projet *Développement d'une méthodologie d'élaboration de définitions terminologiques : analyse de corpus et variation contextuelle*.

Julie Humbert-Droz (doctorat, Université de Genève). Assistanat en terminologie, création de TP, co-organisation de séminaires.

Svetlana Komarova (maîtrise en traduction). Aide au travail éditorial pour un manuscrit portant sur les approches théoriques à la terminologie.

Maryam Mohammadi (doctorat en traductologie). Analyse de données sur l'apport des contextes annotées.

Sophie Lavoie (baccalauréat en traduction). Enrichissement des fiches de LexiTrans.

Luz de Maria Montalvan Ayala (2e cycle, traduction). Enrichissement de la version espagnole du DiCoEnviro.

Jean-Benoît Morel (baccalauréat en traduction). Enrichissement des fiches de LexiTrans.

Caroline Naud (baccalauréat en traduction). Relecture et correction de documents numérisés dans le projet *Identification, description et enseignement du vocabulaire transdisciplinaire en usage au primaire*.

Charlotte Portenseigne (maîtrise en linguistique). Génération automatique de texte.

Elizabeth Saint (doctorat en traductologie). Collecte de données sur l'apport des contextes annotées.

Jean-Baptiste Shamba (doctorat en traductologie). Collecte de données sur l'apport des contextes annotées.

Allison Stentaford (maîtrise en traductologie). Rédaction de fiches dans le DiCoEnviro.

Salomé Tapin (maîtrise en traduction). Élaboration d'un corpus anglais portant sur l'agriculture durable, enrichissement du DiCoEnviro.

Mylène Tardif (maîtrise en didactique). Revue de littérature et analyse de matériel didactique afin de consigner des propositions d'activités lexicales pour le primaire; entrevues avec des enseignants du primaire sur leurs pratiques d'enseignement du lexique.

Emma Thibert (doctorat, Université de Genève). Assistanat en terminologie, création de TP, construction de corpus.

4.2. Stagiaires

Des projets de stage ont été reportés en raison de la situation sanitaire.

4.3. Professionnels de recherche

Catherine Trekker-Séguin (auxiliaire de recherche). Contribution au projet *Développement d'une méthodologie d'élaboration de définitions terminologiques : analyse de corpus et variation contextuelle*.

5. Autres activités

5.1 Séminaires et colloques

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

[<http://rali.iro.umontreal.ca/rali/?q=fr/node/1222/list/2018>]

5.2 Professeurs, chercheurs invités et séjours de recherche

Les projets de visites de chercheurs invités ont été reportés en raison de la situation sanitaire.

5.3 Collaborations

Maria Francesca Bonnadona (Département des langues et littératures étrangères, Université de Vérone) : Version italienne du DiCoEnviro.

Anne Condamines (CLLE-ERSS, CNRS et Université Jean-Jaurès) : Co-direction de la thèse de Julie Humbert-Droz; analyse diachronique du processus de déterminologisation.

Valérie Delavigne (CLESTHIA, Université Sorbonne-nouvelle) : Montage de projet sur les patients-experts et la circulation des termes (projet de thèse d'Emma Thibert).

Isabel Durán-Muñoz (Université de Córdoba) : DiCoAdventure.

- Pamela Faber (Groupe de recherche LexiCon, Université de Grenade) : Extraction de relations conceptuelles à l'aide de « Word sketches » basés sur des patrons de connaissances.
- Paolo Frassi (Département des langues et littératures étrangères, Université de Vérone) : Co-supervision de la thèse de Silvia Calvi.
- Marie-Josée Hamel (Université d'Ottawa) : Projet de dictionnaire électronique de reformulations pour apprenants du français langue seconde.
- Amélie Josselin-Leray, Amaël Arguel, Franck Amadiou et Aline Chevalier (CETIM, CLLE) : Élaboration d'un projet de recherche (FNS-Sinergia) sur le poste de travail des sous-titres-euses.
- Amélie Josselin-Leray (Elizabeth Marshman et Aurélie Picton) : Réflexion sur les suites du colloque ACFAS 2018 > projet de valorisation de l'enseignement en terminologie (en cours).
- Sylvain Kahane (Modyco, Université Paris Ouest) : Élaboration du formalisme des grammaires d'unification polarisées.
- Flávia Lamberti (Département de traduction, Université de Brasilia) : Version portugaise du DiCoEnviro.
- Pilar León Araúz (Groupe de recherche LexiCon, Université de Grenade) : Extraction de relations conceptuelles à l'aide de « Word sketches » basés sur des patrons de connaissances.
- Jian Yun Nie (Département d'informatique et de recherche opérationnelle, Université de Montréal) : Encadrement de Zheng Ying.
- Teresa Ortego (Universidad de Valladolid) : Aproximación a las estrategias de traducción EN-BR-ES del léxico de las redes sociales en un corpus comparable periodístico.
- Alain Polguère (Université de Lorraine) : Projet financé (CRSH—Développement Savoir) pour la génération automatique de définitions lexicographiques.
- Carlos Subirats Rüggeberg (Université autonome de Barcelone) : Découverte de cadres sémantiques dans le DiCoEnviro.
- Tiago Torrent (Université Juiz de Fora, Brésil) et Collin Baker (Berkeley University) : Global FrameNet.

6 Publications et communications

6.1 Livres et chapitres de livres

- Alipour, M. & M.C. L'Homme. 2021. "Une typologie revisitée de la synonymie en terminologie". Agot Canós, R. & D. ar Rouz (dirs.) *Traductologie, terminologie et traduction*, Paris : L'Harmattan. pp. 157-173.
- Drouin, P. 2021. "Repérage outillé de la néologie: Apports de la linguistique de corpus et du traitement automatique de la langue". Díaz, M.B.V., J. C. D. Hoyos, P. Dury, J. Makri-Morel & V. Renner (dirs.) *La néologie des langues romanes : Nouvelles approches, dynamiques et enjeux*. Berne : Peter Lang. 18 pages.
- Gagné, A. et Tremblay, O. 2020. "Développement d'outils de mesure du vocabulaire au préscolaire : enjeux théoriques et méthodologiques d'une approche centrée sur l'enfant". Côté, I., K. Lavoie, R.-P. Trottier-Cyr (dirs.). *La méthodologie de recherche centrée sur l'enfant : recueillir et restituer la parole des enfants et des adolescent-e-s*. Presse de l'Université Laval, 361-384.
- L'Homme, M. 2020. *La terminologie : principes et techniques*, 2e édition, Montréal : Presses de l'Université de Montréal.
- Lefrançois, P. & S. Marcotte. Sous presse. "La correction de textes variés". Montésinos-Gelet, I., M. Dupin de Saint-André & O. Tremblay (dirs.) *Lecture et écriture*. Montréal : Chenelière.
- Marengo, S. 2020. "La théorie Sens-Texte". In Dostie, G. (dir.) *Combats pour la linguistique au Québec (1960-2000). Courants, théories, domaines*. Bruxelles : Peter Lang. pp. 81-115.
- Mel'čuk, I. 2020. "Stat'ja kak stat'ja: pisat' tak pisat'! O sintaksičeskix idiomax. — Stat'ja kak stat'ja: pisat' tak pisat'! Idiomes syntaxiques". Keidan, A. (dir.) *Druz'ja, slova, tablicy. Sbornik k 75letiju A. K. Polivanovoj*. Ekaterinburg: Kabinetnyj učěnyj. pp. 259-278.
- Mel'čuk, I. 2021. *Ten Studies in Dependency Syntax*. Berlin: De Gruyter Mouton, 2021. 443 pages.
- Mel'čuk, I. 2021. "Représentation sémantique". Marengo, S. (dir.) *La théorie Sens-Texte et ses applications. Lexicologie, lexicographie, terminologie, didactique des langues*. Paris : L'Harmattan. pp. 21-47.

- Mel'čuk, I. & L. Iordanskaja. 2020. "Mestoimenija KTO i KOTORYJ v opredelitel'nom pridatočnom". — Les pronoms KTO 'qui' et KOTORYJ ≈ 'lequel' dans une subordonnée modificatrice. In Leonid Iomdin & Igor' Boguslavskij (dirs.) *Ot semantičeskix kvarkov do vselennoj v alfavitnom porjadke. K 90-letiju akademika Jurija Derenikoviča Apresjana [Trudy Instituta ruskogo jazyka im. V. V. Vinogradova, № 2]*. Moskva: IRJaz RAN, pp. 107-129.
- Mel'čuk, I. & J. Milićević. 2020. *An Advanced Introduction to Semantics. A Meaning-Text Approach*. Cambridge: Cambridge University Press. 421 pages.
- Mel'čuk, I. & A. Polguère. 2021. "Les fonctions lexicales dernier cri". Marengo, S. (dir.) *La théorie Sens-Texte et ses applications. Lexicologie, lexicographie, terminologie, didactique des langues*. Paris : L'Harmattan, pp. 73-153.
- Mel'čuk, I. & T. Reuther. 2020. "Explanatory Combinatorial Dictionary". Greenberg, M.L. (dir.) *Encyclopedia of Slavic Languages and Linguistics Online*. First published online.
- Picton A., P. Drouin & J. Humbert-Droz. 2021. "Exploration et visualisation de la variation terminologique en corpus spécialisés complexes : réflexions et propositions méthodologiques". Mangeot, M. & A. Tutin (dirs.) *Lexique(s) et genre(s) textuel(s) : approches sur corpus*. Paris : Éditions des archives contemporaines. pp. 99-116.

6.2 Articles dans des revues avec comité de lecture

- Joan Casademont, A. 2020. "Los verbos fraseológicos especializados en el ámbito de los museos". *Revista de Lenguas Modernas*, 33. pp. 85-109.
[<https://r-libre.telug.ca/2129/>]
- Joan Casademont, A. 2020. "Estudio etiológico de los errores del alumnado francófono de catalán nivel B1". *Revista Fuentes*, 22(1). pp. 65-74.
[<https://r-libre.telug.ca/2029/>]
- Joan Casademont, A. 2020. "Types d'erreurs chez des francophones qui apprennent le catalan dans un examen de niveau B1 (Seuil)". *Revista Íkala*, 25(2). pp. 417-433.
[<https://r-libre.telug.ca/2029/>]
- Joan Casademont, A. 2020. "Analysis of compositions by B1 Level (Threshold) Francophone learners of Catalan: Typology of errors and correspondences". *Journal of Catalan Studies*, 1(22). pp. 33-58.
- Durán-Muñoz, I. & M.C. L'Homme. 2020. "Diving into English motion verbs from a lexico-semantic approach: a corpus-based analysis on adventure tourism". *Terminology*, 26(1). pp. 34-60.

- L'Homme, M.C. 2020. "Being a privileged witness of twenty years of research in terminology: An editorial statement". *Terminology*, 26(1), pp. 1-6.
- Marcotte, S. & P. Lefrançois. Soumis. "Quelles stratégies pédagogiques participent au développement de la compétence scripturale ? Analyse secondaire d'une enquête à grande échelle". *Revue des sciences de l'éducation*.
- Mel'čuk, I. 2020. "The Russian Idiom 'КТО-КТО, А...'". *Russian Linguistics*. DOI 10.1007/s11185-020-09224-z.
- Mel'čuk, I. 2020. "Clichés and Pragmatemes". *Neophilologica*, 32. pp. 9-20. [<http://doi.org/10.31261/NEO.2020.32.01>]
- Mel'čuk, I. & L. Iordanskaja. 2021. "Minimodel' semantičeskogo sinteza russkix fraz". — Minimodèle de synthèse sémantique de phrases russes. *Russkij jazyk v naučnom osveščenii*.
- Mel'čuk, I. & L. Iordanskaja. 2021. "Minimodel of semantic synthesis of Russian sentences". *Lingvisticae Investigationes*.
- Mel'čuk, I., L. Iordanskaja, S. Krylosova & P. Mikhel'. 2020. "Kategorija napravlenosti u russkix parnyx glagolov peremeščenija (na primere glagolov LETET'/LETAT')". — The category of directionality in Russian paired verbs of motion (a case study of LETET'/LETAT'). *Voprosy jazykoznanija*, 1. pp. 27-64.
- Mel'čuk, I. & L. Iordanskaja. 2021. "Neobyčnye russkie parnye podčinitel'nye sojuzy (tipa NE USPEL..., KAK... ili EDVA..., KAK...)". — Quelques conjonctions subordinatives binaires russes inusuelles]. *Izvestija Rossijskoj akademii nauk. Serija literatury i jazyka*, 80(4). pp. 5-20.
- Mel'čuk, I. & L. Iordanskaja. 2021. "OBA – unikal'noe slovo russkogo jazyka". — OBA 'tous les deux', mot unique du russe. *Voprosy jazykoznanija*, 1. 57-69.
- Picton, A., A. Condamines & J. Humbert-Droz. 2021. "Analyse diachronique du processus de déterminologisation. Une réflexion en diachronie courte en physique des particules". Jacquet-Pfau, C. & M.T. Zanola (dirs.) *Terminologie diachronique : méthodologies et études de cas*. Paris : Classiques Garnier, *Cahiers de lexicologie*, 118 (1). pp. 193-225.
- San Martín, A. Sous presse. "A Flexible Approach to Terminological Definitions: Representing Thematic Variation". *International Journal of Lexicography*. [<http://doi.org/10.1093/ijl/ecab013>]

San Martín, A., M. Cabezas-García, M. Buendía, B. Sánchez-Cárdenas, P. León-Araúz, A. Reimerink & P. Faber. 2020. “Presente y futuro de la base de conocimiento terminológica EcoLexicon”. *Onomázein*, 49. pp. 174-202.
[<http://doi.org/10.7764/onomazein.49.09>]

Tremblay, O. et Anctil, A. 2020. Recherches actuelles en didactique du lexique : avancées, réflexions, méthodes. *Revue LIDIL*, 52. pp. 1-14.

6.3 Articles dans des actes de congrès

Drouin, P., J.B. Morel & M.C. L’Homme. 2020. “Automatic Term Extraction from Newspaper Articles: Making the Most of Specificity and Common Features”. *6th International Workshop on Computational Terminology (COMPUTERM 2020). Terminology processing within automatic and human applications, Language Resources and Evaluation (LREC 2020)*. Marseille. France.

L’Homme, M.C. 2020. “Revisiting polysemy for terminology.” Gavriilidou, Z., M. Mitsiaki & A. Fliatouras (dirs.) *Lexicography for Inclusion. Euralex 2020 Proceedings*. pp. 415-424. Alexandroupolis, Greece.

L’Homme, M.C., B. Robichaud & C. Subirats. 2020. “Building multilingual specialized resources based on FrameNet: Application to the field of the environment”. *International Global FrameNet Workshop (IFNW 2020), Language Resources and Evaluation (LREC 2020)*, Marseille. France.

Rigouts-Terryn, A., Drouin, P., Hoste, V., & Lefever, E. (2020). “TermEval 2020: Shared Task on Automatic Term Extraction Using the Annotated Corpora for Term Extraction Research (ACTER) Dataset”. *Proceedings of Computerm 2020*. Marseille, France, pp. 85–94.

San Martín, A., C. Trekker & P. León-Araúz. 2020. “Extraction of Hyponymic Relations in French with Knowledge-Pattern-Based Word Sketches”. *12th Language Resources and Evaluation Conference*. Marseille, France. pp. 5953-5961.
[<http://www.lrec-conf.org/proceedings/lrec2020/pdf/2020.lrec-1.729.pdf>]

6.4 Communications non publiées

Calvi, S., P. Drouin & P. Frassi. 2021. “Acquisition automatique des collocations terminologiques : exploration dans le domaine du droit du commerce international”. *Phraséoterm*. Vérone, Italie.

Dow, M. & P. Drouin. 2021. “L’évolution du genre du mot ‘COVID-19’ dans le français de trois continents : Une étude de deux corpus médiatiques”. *Congrès de l’ACL 2021*.

- Han, Z. 2021. “Description du lexique spécialisé chinois pour locuteurs non sinophones”. *XXXIV^{es} Journées de linguistique (JDL 2021)*.
- Joan Casademont, A., C. Bach & È. Viladrich Castellanas. 2020. “Anàlisi de necessitats formatives a partir de corpus: càpsules didàctiques autocorrectives en línia per als estudiants de català com a llengua addicional”. *Baby-TaLC / E-GRAPELES : Langages et identités*. Perpignan, France. Octobre 2020.
- Gagné, N. & A. Joan Casademont. 2020. “Complexité et précision : Développement oral des enfants débutants en apprentissage de l’anglais langue seconde”. In *Baby-TaLC / E-GRAPELES : Langages et identités*. Perpignan, France. Octobre 2020.
- Joan Casademont, A. & N. Gagné. 2020. “Crosslinguistic influences and L3/Ln teaching: A corpus-based analysis of L3/Ln Catalan intermediate learners' errors in writing.” *TaLC2020: 14th Teaching and Language Corpora Conference*. Perpignan, France. Juillet 2020.
- Humbert-Droz, J. & A. Picton. 2020. “Analysis of terminological variation in specialized corpora / Analyse et traitement de la variation terminologique en corpus spécialisé”, *11e Journée de linguistique suisse*. Fribourg, Suisse. Novembre 2020.
- L’Homme, M.C. 2021. “Las palabras olvidadas de la terminología: verbos, adjetivos y unidades predicativas”. Conférence en ligne donnée dans le cadre de journées d’études. Université de Jaen, Espagne.
- Tremblay, O. et Gagné, A. 2020. “Aimer les mots pour mieux les apprendre”. *Séminaires RALI-OLST*. Université de Montréal, 16 décembre 2020.

6.5 Articles dans des revues professionnelles

- Tremblay, O. et Anctil, D. 2020. Compte rendu du symposium « Concepts émergents, pratiques innovantes et réflexions méthodologiques actuelles en didactique du lexique ». *La Lettre de l’AIRDF*, 67, pp. 26-29.
- Tremblay, O. et Gagné, A. (2020). Chronique « Au fil des mots ». Aimer les mots pour mieux les apprendre : la sensibilité lexicale au service de l’enrichissement du vocabulaire. *Vivre le primaire*, 33(2), pp. 16-18.

6.6 Autres

- Dow, M., & P. Drouin. 2021. Evolution of the gender of “COVID-19” in traditional and social media in the French of three continents. 20 pages.
[\[https://ling.auf.net/lingbuzz/005711\]](https://ling.auf.net/lingbuzz/005711)

- Han, Z. 2021, à paraître. Compte-rendu de Tyulenev, Sergey (2014). Translation and Society: An Introduction, London and New York : Routledge, 210 pages. *Sessions sociologiques*, 11(1).
- Picton, A. 2020. “Renforcement de la pédagogie en classe inverse et vidéos interactives”. Témoignages et partages de pairs, Pôle SEA, Université de Genève. [<https://www.unige.ch/enseignement-a-distance/communautes-pratique/temoignages-pairs/>]
- Picton, A. 2020. “Inverser sa classe et structurer le travail à distance pour autonomiser les étudiant-es dans leurs apprentissages”. Atelier partage d’expériences, Pôle SEA, Université de Genève. 19 novembre 2020, [<https://moodle.unige.ch/course/view.php?id=3749>]
- Tremblay, O. 2020. Regards sur un parcours de chercheur en didactique du lexique : Francis Grossmann. Lidil [En ligne], 62. Mis en ligne le 03 novembre 2020. [<http://journals.openedition.org/lidil/8153>]