

## TITRE COURS : **TA-TAO ET TEXTOMÉTRIE**

**Intervenant** : Maria ZIMINA-POIROT (Université Paris Diderot - USPC)

Maria Zimina. MCF à l'université Paris Diderot-Paris 7. Thèse en sciences du langage « Approches quantitatives de l'extraction de ressources traductionnelles à partir de corpus parallèles » soutenue en 2004 à l'université Sorbonne Nouvelle-Paris 3. Publications sur la textométrie multilingue, traduction assistée par ordinateur (TAO), terminologie, analyse de corpus parallèles et comparables avec prise en compte des annotations multiples. Contribution au développement des logiciels de textométrie multilingue MkAlign (<http://www.tal.univ-paris3.fr/mkAlign>), Le Trameur (<http://www.tal.univ-paris3.fr/trameur>) et le iTrameur (<http://www.tal.univ-paris3.fr/trameur/iTrameur>). Projets de recherche intégrant le traitement automatique des langues, la statistique textuelle, la TA-TAO, l'exploration et la visualisation des alignements textuels (PERTOMed, TextCoop, EVALDA-ARCADE II, FORMMEL, ANR TransRead). Recherches en cours sur l'analyse de la phraséologie prosodique (corpus RHAPSODIE) ; l'anglais médical simplifié (corpus Cochrane). Co-responsable des projets pédagogiques dans le cadre du Master Pro ILTS : TSA (Traduction de Site web vers l'Anglais, démarré en 2014) et ENSUPECO (ENglish SUBtitled of PEdagogical Content, financé dès 2019).

**Cordonnées** : [mzimina@eila.univ-paris-diderot.fr](mailto:mzimina@eila.univ-paris-diderot.fr)

\*\*\*

**Date et horaire** : **jeudi 19 juillet, 9h-9h45**

### DESCRIPTION GÉNÉRALE DU COURS

L'accélération des rythmes de travail et l'accroissement des volumes de textes à traduire rend nécessaire le recyclage optimal des ressources de traductions existantes. Parallèlement, le débat sur l'évaluation de la qualité de la version finale de la traduction valorise une attention soutenue aux détails. Au delà d'expériences particulières et multiples dans l'exploration des corpus de traductions avec des méthodes de la linguistique de corpus, la textométrie utilise des modèles de données qui préservent des liens unissant les textes source et cible à plusieurs niveaux (mots, syntagmes, phrases, cadres discursifs, etc.). Une complémentarité entre l'analyse statistique des unités textuelles dénombrables, issues des traitements informatisés, et l'identification qualitative des équivalences par analyses successives en corpus, permet d'élargir les possibilités d'exploration de textes traduits au-delà de la lecture tabulaire de segments alignés.

### OBJECTIFS A ATTEINDRE

- **cerner les tendances récentes dans la pratique TA-TAO**
- **comprendre les spécificités de la démarche textométrique**
- **observer des liens entre *unités textuelles en corpus et équivalences de traduction***

### BIBLIOGRAPHIE

CRONIN, Michael, Translation in the Digital Age, Milton Park, Abingdon, Oxon, Routledge, 2013.  
FLEURY, Serge, ZIMINA, Maria, "Trameur: A Framework for Annotated Text Corpora Exploration", Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: System Demonstrations, Dublin, Dublin City University, 23-29 August 2014, ed. Tounsi, Lamia, Rak, Rafal, ACL, 2014, p. 57-61.  
SALEM, André, « Introduction à la résonance textuelle », Actes des 7es Journées internationales d'Analyse statistique des Données Textuelles, 10-12 mars 2004, Louvain-la-Neuve, Université catholique de Louvain, éd. Purnelle, Gérard, Fairon, Cédric, Dister, Anne, Louvain-la-Neuve, Presses Universitaires de Louvain, vol. 2, 2004, p. 986-992.  
ZIMINA Maria, FLEURY, Serge, « Approche systémique de la résonance textuelle multilingue », Actes des 12es Journées Internationales d'Analyse Statistiques des Données Textuelles, Paris, INaLCO, 3-6 juin 2014, éd. Née, Émilie, Daube Jean-Michel, Valette, Mathieu, Fleury, Serge, Paris, JADT.org, 2014, p. 717-728.  
ZIMINA, Maria, FLEURY, Serge, « Perspectives de l'architecture Trame/Cadre pour les alignements multilingues ». Nouvelles perspectives en sciences sociales : revue internationale de systémique complexe et d'études relationnelles, vol. 11, no 1, 2015.  
ZIMINA, Maria, « Alignement textométrique des unités lexicales à correspondances multiples dans les corpus parallèles », Actes des 7es Journées internationales d'Analyse statistique des Données Textuelles, 10-12 mars 2004, Louvain-la-Neuve, Université catholique de Louvain, éd. Purnelle, Gérard, Fairon, Cédric, Dister, Anne, Louvain-la-Neuve, Presses Universitaires de Louvain, vol. 2, 2004, p. 1195-1202.

\*\*\*

## **TITRE ATELIER : *OUTILS D'AIDE À LA GESTION DE RESSOURCES TRADUCTIONNELLES INCRÉMENTALES***

**Date et horaire :** jeudi 19 juillet, 14h00-15h30

### **TEXTE(S) ETUDIÉ(S) EN ATELIER**

The Cochrane Collaboration, *Cochrane Database of Systematic Reviews (CDSR)*, 2018.  
<http://www.cochranelibrary.com/cochrane-database-of-systematic-reviews>

### **DESCRIPTION GÉNÉRALE DE L'ATELIER**

Cette session de travail permet d'explorer un corpus parallèle anglais-français composé des titres de résumés d'articles scientifiques en anglais (*Cochrane Systematic Reviews*), leurs réécritures en anglais médical simplifié dans *Plain Language Summaries (PLS)*, et leurs traductions respectives en français (2010-2015). Le concept de *ressource traductionnelle incrémentale* constitue le fil conducteur de cette exploration itérative en corpus menée sur la base de l'action commune des apports quantitatifs et qualitatifs. Cette démonstration vise à montrer comment une Mémoire de Traduction dotée d'un moteur textométrique (*iTrameur* : <http://www.tal.univ-paris3.fr/trameur/iTrameur/>) permet d'induire des correspondances à plusieurs niveaux d'analyse.