

Filip Miletic

Room B510 • Maison de la Recherche • Université Toulouse – Jean Jaurès

5, allées Antonio Machado • 31058 Toulouse Cedex 9 • France

filip.miletic@univ-tlse2.fr

EDUCATION

2018–present	PhD in Linguistics Toulouse Jean Jaurès University • CLLE Research Laboratory (UMR CNRS 5263) (France) Topic: <i>Using word embeddings to detect semantic shifts in Quebec English.</i> Supervisors: Anne Przewozny-Desrioux, Ludovic Tanguy. Funding: 3-year doctoral contract.
2016–2018	MA in Modern Languages and Literatures for Cultural Services (English and French) University of Genoa (Italy) Visiting student at the University of Sherbrooke (Quebec, Canada) (August–December 2017). Dissertation: <i>Contact-induced lexical and morphosyntactic phenomena in Quebec English.</i> Supervisors: Cristiano Broccias, Anna Giaufret. Grade: 110/110 <i>cum laude</i> , recommended for publication (It. <i>dignità di stampa</i>).
2014–2018	Certificate in Humanities and Cultural Heritage IANUA School for Advanced Studies • University of Genoa (Italy) Advanced selective program on multidisciplinary approaches to issues in the humanities.
2013–2016	BA in Modern Languages and Cultures (English and French) University of Genoa (Italy) Visiting student at the University of Nantes (France) (January–May 2016). Dissertation: <i>Lexical and morphosyntactic features of Canadian English on Twitter.</i> Supervisor: Cristiano Broccias. Grade: 110/110 <i>cum laude</i> .

TEACHING

2019–present	Department of Linguistics • Toulouse Jean Jaurès University (France)			
	BA	1 st year	48h	Introduction to sociolinguistics
	BA	2 nd year	48h	Computer science for linguistics
	MA	2 nd year	4h	Topics in NLP [seminar]
	MA	2 nd year	12h	Using Twitter to create linguistic corpora [tutorial]
2018	Department of Modern Languages and Cultures • University of Genoa (Italy)			
	BA	1 st year	17h	French morphosyntax and listening comprehension [tutoring]
	BA	3 rd year	17h	French composition [tutoring]

OTHER EXPERIENCE

2017	University Language Center (CLAT) • University of Genoa (Italy) Internship: content editing and translation for EFMR.it (<i>Études françaises mises en réseau</i>).
2016	Parallel Corpora and Bilingual Lexicography: Research and Applications Research project • CLLE, CNRS & University of Toulouse (France) Manual lemmatization and morphosyntactic annotation of a Serbian corpus.

INSTITUTIONAL AND ACADEMIC SERVICE

2020, 2021	Organizing committee member, <i>Le Printemps des Jeunes Chercheurs CLLE</i> , Toulouse (France).
2020	Reviewer, <i>SemEval 2020 Task 1: Unsupervised Lexical Semantic Change Detection</i> .
2019–present	PhD student representative, CLLE Research Laboratory.
2019	Participant, <i>Avez-vous vu l'humain ? Journée découverte des SHS</i> , a general public event marking the 80 th anniversary of the CNRS, Toulouse (France), December 2019.
2019	Organizing committee member, <i>Travel in Quebec: Auto- and Hetero-Representations</i> , Interuniversity Center for Quebec Studies, Genoa (Italy), April 2019.
2018–present	Member, CEPEL (Student Society for the Study of Language), Toulouse Jean Jaurès University.

PUBLICATIONS AND PRESENTATIONS

Miletic, F., Przewozny-Desriaux, A., Tanguy, L. (2021) [accepted]. Modeling fine-grained sociolinguistic variation: The promises and pitfalls of Twitter corpora and neural word embeddings. *Corpus Pitfalls: Dealing with Messy Data (and Other Traps for the Unwary) (Workshop at ICAME42)*, Dortmund.

Miletic, F., Przewozny-Desriaux, A., Tanguy, L. (2021) [accepted]. The status and representation of contact-induced semantic shifts in Quebec English: From Twitter users to sociolinguistic informants. *3rd International Conference on Sociolinguistics: Diversities, New Media and Language Management*, Prague.

Miletic, F., Przewozny-Desriaux, A., Tanguy, L. (2020). Collecting tweets to investigate regional variation in Canadian English. *Proceedings of the 12th Conference on Language Resources and Evaluation (LREC 2020)*, pages 6255–6264.

Miletic, F., Przewozny-Desriaux, A., Tanguy, L. (2020). Methodological issues in using word embeddings in a sociolinguistic perspective: The case of contact-induced semantic variation across Canadian Twitter corpora. *Empirical Studies of Word Sense Divergences across Language Varieties (Workshop at DGfS 2020)*, Hamburg.

Miletic, F. (2019). Contact-induced lexical variation in Quebec English: An accountable description. *RJC2019 – 22èmes Rencontres des jeunes chercheurs en Sciences du Langage*, Paris.

Miletic, F. (2019). Repérage des glissements de sens en anglais québécois à travers l'analyse sémantique distributionnelle : enjeux méthodologiques d'une approche interdisciplinaire. [Detecting semantic shifts in Quebec English using distributional semantic analysis: Methodological issues in an interdisciplinary approach.] *Le Printemps des Jeunes Chercheurs CLLE*, Toulouse.

Miletic, F. (2017). Lexical features of Canadian English on Twitter. *Canada in the Making: 150 Years of Cultural and Linguistic Diversity*, Rende (Italy).

LANGUAGES

native language	Serbian
near-native proficiency	English • French • Italian <i>Cambridge Certificate of Proficiency in English • grade A • score 228/230 • 3 March 2018</i>

IT SKILLS

NLP tools	word2vec • word2vec • DISSECT • BERT • Twython • CMU ARK Twitter NLP Tools
programming	Python • numpy • pandas • scipy • scikit-learn • gensim • matplotlib bash • R • HTML • CSS • PHP
software suites	LaTeX • Microsoft Office • Adobe Creative Suite