Chuyuan Li

Department of Computer Science, The University of British Columbia

web: chuyuanli.github.io
email: chuyuan.li@ubc.ca

Education

Ph.D. in Computer Science 2023
Lorraine Research Laboratory in Computer Science and its Applications & Université de Lorraine *France*

- **Dissertation:** Facing Data Scarcity in Dialogues for Discourse Structure Discovery and Prediction.
- **Advisors:** Maxime Amblard (Professor, LORIA), Chloé Braud (Rechercher, CNRS).
- **Research Internship:** University of British Columbia (Nov 2021-May 2022), advisor: Giuseppe Carenini.

M.Sc. in Computational Linguistics 2019
Université Paris Cité (formerly Paris Diderot). Graduated with distinction (mention très bien). *France*

B.A. in French 2015
Wuhan University of Technology. Graduated summa cum laude. *China*

Professional Experience

Post-doctoral Researcher Nov 2023-Present
Computer Science Department, University of British Columbia *Canada*

- **Research projects:**
  - Large Language Models: In-context learning (ACL 2025); reinforcement learning (EMNLP 2025); test-time scaling (NewSumm workshop 2025); multi-agent systems (collaboration with Salesforce AI Research).
  - Discourse Analysis: LLM-based discourse parsing (SIGDIAL 2024, **Best Paper Award**); relation prediction (CODI workshop 2024-25); benchmarking (EACL 2026), and interpretability.
  - Machine Learning for Health: Dementia detection with multimodal data (PMLR 2025); principled zero-shot (CL4Health workshop 2025) and in-context learning approaches (ACL 2025).
- **Advisor:** Giuseppe Carenini, carenini@cs.ubc.ca

Temporary Teaching and Research Associate (ATER) Sep 2022-Sep 2023
Institut des sciences du Digital, Management & Cognition, Université de Lorraine *France*

- 100 hours of teaching, including plenary lectures and tutorials to L1-M2 students.
- Subjects: introduction to NLP, formal logic and reasoning, discourse and dialogue, introduction to Grid'5000 HPC, programming languages (Python, C).

Teaching Assistant during Ph.D. Sep 2019-Sep 2022
LORIA, Inria, Université de Lorraine *France*

Industrial Experiences

***Synomia**, Paris, France* May-Sep 2019

- Computational linguist: Developed an automatic language processing chain for Chinese for automatic extraction of entities and sentiments. The processing chain includes word segmentation, syntactic analysis, and entity extraction.

***SemantiWeb France**, Paris, France* Jul-Aug 2018

- Computational linguist: Machine learning for sentiment detection in SNCF-related tweets; linguistic resource development (segmentation, dictionaries, named entities).

Lionbridge**, remote, in collaboration with **Amazon Alexa May-Jun 2018

- Computational linguist: Created and implemented Backus–Naur Form (BNF) grammars to improve Chinese speech recognition for Alexa.

Publications and Communication

REFERRED JOURNAL ARTICLE

[1] Jonathan Ginzburg, Zulipiye Yusupujiang, **Chuyuan Li**, Kexin Ren, Aleksandra Kucharska and Pawel Lupkowski. Characterizing the Response Space of Questions: data and theory. In *Journal Dialogue & Discourse (D&D)* 13.2 (2022). https://doi.org/10.5210/dad.2022.203.

REFERRED CONFERENCE ARTICLES

[2] **Chuyuan Li**, and Giuseppe Carenini. BeDiscovER: The Benchmark of Discourse Understanding in the Era of Reasoning Language Models. In *Proceedings of the 19th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2026, to appear)*, Rabat, Morocc, March 2026. Acceptance rate: 20%.

[3] **Chuyuan Li**, Austin Xu, Shafiq Joty, and Giuseppe Carenini. Topic-guided reinforcement learning with LLMs for enhancing multi-document summarization. In *Findings of the Association for Computational Linguistics: EMNLP 2025*, Suzhou, China, November 2025. https://aclanthology.org/2025.findings-emnlp.662/. Acceptance rate: 17%.

[4] Amirhossein Abaskohi, Raymond Li, **Chuyuan Li**, Shafiq Joty, and Giuseppe Carenini. CEMTM: Contextual Embedding-based Multimodal Topic Modeling. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Suzhou, China. 2025. https://aclanthology.org/2025.emnlp-main.590/. Acceptance rate: 22%.

[5] **Chuyuan Li**, Raymond Li, Thalia S. Field, and Giuseppe Carenini. Delta-KNN: Improving Demonstration Selection in In-Context Learning for Alzheimer's Disease Detection. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (ACL)*, Vienna, Austria. 2025. https://aclanthology.org/2025.acl-long.1253/. Acceptance rate: 20%.

[6] Raymond Li, **Chuyuan Li**, Gabriel Murray, and Giuseppe Carenini. Explicit Bayesian Inference to Uncover the Latent Themes of Large Language Models. In *Findings of the Association for Computational Linguistics (ACL): ACL 2025*, Vienna, Austria. 2025. https://aclanthology.org/2025.findings-acl.1123/. Acceptance rate: 16.7%.

[7] Shih-Han Chou, Miini Teng, Harshinee Sriram, **Chuyuan Li**, Giuseppe Carenini, Cristina Conati, Thalia S. Field, Hyeju Jang, Gabriel Murray. Multimodal Classification of Alzheimer's Disease by Combining Facial and Eye-tracking Data. In *Proceedings of the 4th Machine Learning for Health Symposium (ML4Health)*, Vancouver, Canada. PMLR. 2024. https://proceedings.mlr.press/v259/chou25a.html. Acceptance rate: 40%.

[8] **Chuyuan Li**, Yuwei Yin, and Giuseppe Carenini. Dialogue Discourse Parsing as Generation: a Sequence-to-Sequence LLM-based Approach. In *Proceedings of the 25th Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGdial)*, Kyoto, Japan. 2024. https://aclanthology.org/2024.sigdial-1.1/. Acceptance rate: 41%. **Best Paper Award**.

[9] Gaetano Cimino, **Chuyuan Li**, Giuseppe Carenini, and Vincenzo Deufemia. Coherence-based Dialogue Discourse Structure Extraction using Open-Source Large Language Models. In *Proceedings of the 25th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, Kyoto, Japan. 2024. https://aclanthology.org/2024.sigdial-1.26/. Acceptance rate: 41%

[10] **Chuyuan Li**, Patrick Huber, Wen Xiao, Maxime Amblard, Chloé Braud and Giuseppe Carenini. Discourse Structure Extraction from Pre-Trained and Fine-Tuned Language Models in Dialogues. In *Findings of the Association for Computational Linguistics: EACL 2023*, Dubrovnik, Croatia. 2023. https://aclanthology.org/2023.findings-eacl.194/. Acceptance rate: 17%

[11] **Chuyuan Li**, Chloé Braud and Maxime Amblard. Multi-Task Learning for Depression Detection in Dialogs. In *Proceedings of the 23rd Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGdial)*, Edinburgh, UK. 2022. https://aclanthology.org/2022.sigdial-1.7/. Acceptance rate: 45%.

[12] Maxime Amblard, Chloé Braud, **Chuyuan Li**, Michel Musiol, Caroline Demily et Nicolas Franck. Investigation par Méthodes d'Apprentissage des Spécificités Langagières propres aux Personnes avec Schizophrénie (Investigating Learning Methods Applied to Language Specificity of Persons with Schizophrenia). In *Actes de la 6e conférence conjointe JEP (33e édition), TALN (27e édition), RÉCITAL (22e édition).* 2020. https://www.aclweb.org/anthology/2020.jeptalnrecital-taln.2/. Acceptance rate: 63%.

[13] Jonathan Ginzburg, Zulipiye Yusupujiang, **Chuyuan Li**, Kexin Ren and Paweł Łupkowski. Characterizing the Response Space of Questions: a Corpus Study for English and Polish. In *Proceedings of the 20th Annual SIGdial Meeting on Discourse and Dialogue (SIGdial)*, Stockholm, Sweden. 2019. https://www.aclweb.org/anthology/W19-5937/. Acceptance rate: 35%.

ARTICLES IN PEER-REVIEWED WORKSHOP PROCEEDINGS

[14] Shuhaib Mehri, **Chuyuan Li**, Giuseppe Carenini. Discourse Relation Recognition with Language Models Under Different Data Availability. In *Proceedings of the 6th Workshop on Computational Approaches to Discourse, Context and*

Document-Level Inferences (CODI 2025), EMNLP, Suzhou, China. 2025. https://aclanthology.org/2025.codi-1.13/.

[15] Chloé Braud, Amir Zeldes, **Chuyuan Li**, Yang Janet Liu, Philippe Muller. The DISRPT 2025 Shared Task on Elementary Discourse Unit Segmentation, Connective Detection, and Relation Classification. In *Proceedings of the 4th Shared Task on Discourse Relation Parsing and Treebanking (DISRPT 2025)*, EMNLP, Suzhou, China. 2025. https://aclanthology.org/2025.disrpt-1.1/.

[16] Raymond Li, **Chuyuan Li**, Gabriel Murray, Giuseppe Carenini. Hierarchical Attention Adapter for Abstractive Dialogue Summarization. In *Proceedings of The 5th New Frontiers in Summarization Workshop*, Hybrid. 2025. https://aclanthology.org/2025.newsum-main.2/.

[17] Juntai Cao, Xiang Zhang, Raymond Li, Jiaqi Wei, **Chuyuan Li**, Shafiq Joty, Giuseppe Carenini. Multi2: Multi-Agent Test-Time Scalable Framework for Multi-Document Processing. In *Proceedings of The 5th New Frontiers in Summarization Workshop*, Hybrid. 2025. https://aclanthology.org/2025.newsum-main.10/.

[18] **Chuyuan Li**, Giuseppe Carenini, and Thalia Field. On Large Foundation Models and Alzheimer's Disease Detection. In *Proceedings of the Second Workshop on Patient-Oriented Language Processing (CL4Health)*, NAACL, Albuquerque, New Mexico. 2025. https://aclanthology.org/2025.cl4health-1.13/.

[19] **Chuyuan Li**, Chloé Braud, Maxime Amblard, and Giuseppe Carenini. Discourse Relation Prediction and Discourse Parsing in Dialogues with Minimal Supervision. In *Proceedings of the 5th Workshop on Computational Approaches to Discourse (CODI)*, St. Julians, Malta. 2024. https://aclanthology.org/2024.codi-1.15.

[20] **Chuyuan Li**, Maxime Amblard and Chloé Braud. A Semi-supervised Dialogue Discourse Parsing Pipeline. In *Actes des 5èmes journées du Groupement de Recherche CNRS "Linguistique Informatique, Formelle et de Terrain"*. (2023): 85. https://inria.hal.science/hal-04356416/document.

[21] **Chuyuan Li**, Maxime Amblard, Chloé Braud, Caroline Demily, Nicolas Franck and Michel Musiol. Investigating non lexical markers of the language of schizophrenia in spontaneous conversations. In *Proceedings of the 2nd Workshop on Computational Approaches to Discourse (CODI)*, Punta Cana, Dominican Republic and Online. 2021. https://aclanthology.org/2021.codi-main.3/.

DOCTORAL DISSERTATION

[22] **Chuyuan Li**. Facing Data Scarcity in Dialogues for Discourse Structure Discovery and Prediction. *Université de Lorraine*, 2023. https://theses.hal.science/tel-04218515/file/DDOC_T_2023_0107_LI.pdf.

INVITED TALKS

- Guest lecture in COGS 401, the senior Cognitive Systems seminar. Philosophy and Cognitive Systems departments at UBC, Vancouver, Canada. February 2026 (to appear)
- "Beyond Benchmarks: Enhancing LLMs for Discourse Analysis and Healthcare Applications", invited talk at School of Computing Science, Simon Fraser University, Vancouver, Canada. March 2025
- "LLM-based Dialogue Discourse Parsing", seminar at Melodi, IRIT, Toulouse, France. September 2024
- "Discourse Structure Extraction in Dialogues", seminar Café TAL, ATILF, Nancy, France. September 2022
- "Discourse Structure Extraction in Dialogues", seminar at Melodi, IRIT, Toulouse, France. June 2022

Awards & Honors

2024 **Best Paper Award**, 25th Annual Meeting of the Special Interest Group on Discourse and Dialogue
2024 **Postdoctoral Fellows Travel Award ($1,500)**, UBC Faculty of Science
2022 **DrEAM International Mobility Grant (€9,000)**, Université de Lorraine & Mitacs
2019 **Summa Cum Laude**, M. Sc., Université Paris Cité, France
2017 **Summa Cum Laude**, Master, ESSEC Business School, France
2015 **Summa Cum Laude**, Bachelor, 1st in class, Wuhan University of Technology, China
2011-2014 **National Scholarship** (3 years), Ministry of Education, China

Service

Co-Organizer	CODI-CRAC'26, "2nd Joint Workshop on CODI (7th) and CRAC (9th)", to be held at ACL 2026
	DISRPT'25, "Shared Task on Discourse Segmentation, Connective and Relation Identification across Formalisms", held at CODI-CRAC'25
	CODI-CRAC'25, "1st Joint Workshop on CODI (6th) and CRAC (8th), EMNLP 2025
	CODI'24, "5th Workshop on Computational Approaches to Discourse, Context and Document-Level Inferences", EACL 2024
	IWCS 2023, "15th International Conference on Computational Semantics"
	JEP-TALN-RECITAL 2020, "National Joint Conference on Speech Studies (33rd), Natural Language Processing (27th), and Student Researchers in Computer Science for Natural Language Processing (22nd)", France
Area Chair	Conference of the Nations of the Americas Chapter of the ACL (NAACL), 2025
	International Conference on Computational Linguistics (COLING), 2025
	Conference on Empirical Methods in Natural Language Processing (EMNLP), 2024
	Conference of the European Chapter of the ACL (EACL), 2024
Program Committee	International Conference on Language Resources and Evaluation (LREC), 2026
	Annual AAAI Conference on Artificial Intelligence (AAAI), 2025
	Conference on Computational Semantics (IWCS), 2025
	Annual Meeting of Special Interest Group on Discourse and Dialogue (SIGdial), 2024-25
	Workshop on Linguistic Analysis for Health (HeaLing), 2026
	Workshop on Patient-Oriented Language Processing (CL4Health), 2025-26
Reviewer	ACL Rolling Review (standing reviewer since 2023)
	Transactions of the Association for Computational Linguistics (TACL), temporal 2024

Supervision

Ph.D. student	Raymond Li, UBC, research mentorship, since 2025
	Gaetano Cimino, University of Salerno, internship, Sep 2023-Jul 2024 ⇒ Now postdoc at UBC
Master student	Amirhossein Abaskohi, UBC, research mentorship, since 2025 ⇒ Now Ph.D. student at UBC
	Albert Millert, University of Lorraine, directed study, Sep 2020-Sep 2021
	Anar Yeginbergenova, University of Lorraine, directed study, Sep 2020-Sep 2021
Undergraduate	Armin Talaie, UBC, honors thesis, Oct 2024-Jul 2025
	Ronald Liu, UBC, USRA, May-Sep 2024
	Ahmad Shuhaib Mehri, UBC, COGS 402, Sep 2023-Jun 2024 ⇒ Now Ph.D. student at UIUC
	Ryan Lazenby, UBC, COGS 402, Jan-Jun 2024

Languages & Skills

Languages: Chinese (native), English (fluent), French (fluent)
Programming Languages: Python, Bash
Machine Learning: PyTorch, TensorFlow, HuggingFace, DeepSpeed, Weights & Biases, Scikit-Learn
High-Performance Computing: Grid'5000 (OAR), Compute Canada (Slurm)
Operating Systems: macOS, Linux