

MARINE CARPUAT

Department of Computer Science
Brendan Iribe Center
University of Maryland
8125 Paint Branch Drive
College Park, MD 20742, USA

<http://www.cs.umd.edu/marine>
marine@cs.umd.edu
+1 202-378-4910

I Employment

- 01/2015 – Present **Assistant Professor, University of Maryland**
Department of Computer Science & Institute for Advanced Computer Studies (UMIACS)
Affiliate appointments in the Applied Mathematics & Statistics, and Scientific Computation
graduate program, and in the College of Information Studies
- 05/2013 – 12/2014 **Research Officer, National Research Council Canada**
Multilingual Text Processing Group
- 11/2010 – 05/2013 **Research Associate, National Research Council Canada**
Multilingual Text Processing Group
- 09/2008 – 10/2010 **Postdoctoral Researcher, Columbia University**
Center for Computational Learning Systems

II Educational Background

- Ph.D. Computer Science and Engineering, Hong Kong University of Science and Technology, 2008.
- M.Phil. Electrical and Electronic Engineering, Hong Kong University of Science and Technology, 2002.
- Diplôme d'ingénieur, Ecole Supérieure d'Electricité (Supélec), 2002.

III Awards

- NSF CAREER Award 2018
- Industry research awards: Amazon 2016 and 2018, Google 2016
- Best Paper Awards: *-SEM 2017, TALN 2010
- Computer Science Department Teaching Award 2016
- Best Reviewer Award EMNLP-CoNLL 2012

IV Publications

IV.A Book Chapters

1. Dekai Wu, Pascale Fung, **Marine Carpuat**, Chi-kiu Lo, Yongsheng Yang, and Zhaojun Wu, “Lexical Semantics for Statistical Machine Translation”, in *Handbook of Natural Language Processing and Machine Translation, DARPA Global Autonomous Language Exploitation*, edited by Joseph Olive, Caitlin Christianson and John McCary, pp 236-252, 2011

IV.B Journal Articles

1. Aleš Tamchyna, Fabienne Braune, Alexander Fraser, **Marine Carpuat**, Hal Daumé III and Chris Quirk. “Integrating a Discriminative Classifier into Phrase-based and Hierarchical Decoding”. *Prague Bulletin of Mathematical Linguistics*, 101(April), pp 29-42, 2014

2. Ann Irvine, John Morgan, **Marine Carpuat**, Hal Daumé III and Dragos Munteanu. “Measuring Machine Translation Errors in New Domains”. *Transactions of the Association for Computational Linguistics (TACL)*, 1(October):429–440, 2013
3. **Marine Carpuat**, Yuval Marton and Nizar Habash. “Improving Arabic-to-English Statistical Machine Translation by Reordering Post-verbal Subjects for Alignment”. *Machine Translation, Special Issue on Machine Translation for Arabic*, 26(1-2), pp 105-120, 2012
4. **Marine Carpuat**, Pascale Fung and Grace Ngai, “Aligning Word Senses using Parallel Corpora”, *ACM Transactions on Asian Language Processing*, 5(2), pp 89-120, 2006

IV.C Refereed Conference Papers

1. Sweta Agrawal and **Marine Carpuat**. “Controlling Text Complexity in Neural Machine Translation”. *Empirical Methods in Natural Language Processing (EMNLP)*. 2019.
2. Yogarshi Vyas and **Marine Carpuat**. “Weakly Supervised Cross-lingual Semantic Relation Classification via Knowledge Distillation”. *Empirical Methods in Natural Language Processing (EMNLP)*. 2019.
3. Marianna J. Martindale, **Marine Carpuat**, Kevin Duh and Paul McNamee. “Identifying Fluently Inadequate Output in Neural and Statistical Machine Translation”. *Machine Translation Summit*. 2019.
4. Eleftheria Briakou and **Marine Carpuat**. “The University of Maryland’s Kazakh–English Neural Machine Translation System at WMT19”. *Conference on Machine Translation (WMT)*. 2019.
5. Xuan Zhang, Pamela Shapiro, Gaurav Kumar, Paul McNamee, **Marine Carpuat** and Kevin Duh. “Curriculum Learning for Domain Adaptation in Neural Machine Translation”. *Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*. 2019.
6. Weijia Xu, Xing Niu and **Marine Carpuat**. “Differentiable Sampling with Flexible Reference Word Order for Neural Machine Translation”. *Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL) - Short Papers*. 2019.
7. Xing Niu, Weijia Xu and **Marine Carpuat**. “Bi-Directional Differentiable Input Reconstruction for Low-Resource Neural Machine Translation”. *Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL) - Short Papers*. 2019.
8. Ugur Koc, Shiyi Wei, Jeffrey Foster, **Marine Carpuat** and Adam Porter. “An Empirical Assessment of Machine Learning Approaches for Triaging Java Static Analysis Reports”. *IEEE International Conference on Software Testing, Verification and Validation (ICST)*. 2019.
9. Weijia Xu and **Marine Carpuat**. “The University of Marylands Chinese-English Neural Machine Translation Systems at WMT18”. *Conference on Machine Translation (WMT)*. 2018.
10. Xing Niu, Sudha Rao and **Marine Carpuat**. “Multi-Task Neural Models for Translating Between Styles Within and Across Languages”. *International Conference on Computational Linguistics (COLING)*. 2018.
11. Shyam Upadhyay, Yogarshi Vyas, **Marine Carpuat**, and Dan Roth. “Robust Cross-lingual Hypernymy Detection using Dependency Context”. *Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*. 2018.
12. Yogarshi Vyas, Xing Niu and **Marine Carpuat**. “Identifying Semantic Divergences in Parallel Text without Annotations”. *Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*. 2018.
13. Marianna J. Martindale and **Marine Carpuat**. “Fluency Over Adequacy: A Pilot Study in Measuring User Trust in Imperfect MT”. *Association for Machine Translation in the Americas*. 2018.
14. Xing Niu and **Marine Carpuat**. “A Study of Style in Machine Translation: Controlling the Formality of Machine Translation Output”. *Empirical Methods in Natural Language Processing - Short Papers*. 2017.

15. Yogarshi Vyas and **Marine Carpuat**. “Detecting Lexical Entailment in Context”. Joint Conference on Lexical and Computational Semantics (\star SEM). 2017.
16. Ahmed Elgohary and **Marine Carpuat**. “Learning Easy to Compose Word Representations via Bilingual Supervision”. *Annual Conference of the Association for Computational Linguistics (ACL) - Short Papers*. 2016.
17. Yogarshi Vyas and **Marine Carpuat**. “Sparse Bilingual Word Representations for Cross-Lingual Lexical Entailment”. *Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*. 2016.
18. Allyson Ettinger, Philip Resnik and **Marine Carpuat**. “Retrofitting Sense-Specific Word Vectors Using Parallel Text”. *Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL) - Short Papers*. 2016.
19. Junyi Jessy Li, **Marine Carpuat**, Ani Nenkova. “Cross-lingual Discourse Relation Analysis: A corpus study and a semi-supervised classification system”. *International Conference on Computational Linguistics (COLING)*. 2014
20. Junyi Jessy Li, **Marine Carpuat**, Ani Nenkova. “Assessing the Discourse Factors that Influence the Quality of Machine Translation”. *53rd Annual Conference of the Association for Computational Linguistics (ACL)*. 2014
21. **Marine Carpuat**, Hal Daumé III, Katie Henry, Ann Irvine, Jagadeesh Jagarlamudi and Rachel Rudinger. “SenseSpotting: Never let your parallel data tie you to an old domain”. *52nd Annual Conference of the Association for Computational Linguistics (ACL)*. 2013
22. Atsushi Fujita and **Marine Carpuat**. “FUN-NRC: Paraphrase-augmented Phrase-based SMT Systems for NTCIR-10 PatentMT”. *10th NTCIR Conference*. 2013
23. Cyril Goutte, **Marine Carpuat**, and George Foster. “The Impact of Sentence Alignment Errors on Phrase-Based Machine Translation Performance”. *Conference of the Association for Machine Translation in the Americas (AMTA)*. 2012
24. **Marine Carpuat**, Cyril Goutte and Pierre Isabelle. “Filtering and Routing Multilingual Documents for Translation”. *IEEE Symposium on Computational Intelligence for Security and Defence Applications*. 2012
25. **Marine Carpuat**, Yuval Marton and Nizar Habash. “Improving Arabic-to-English Statistical Machine Translation by Reordering Post-verbal Subjects for Alignment”. *48th Annual Conference of the Association for Computational Linguistics (ACL)*. 2010
26. **Marine Carpuat**, Yuval Marton and Nizar Habash. “Reordering Matrix Post-verbal Subjects for Arabic-to-English SMT”. *17ème Conférence sur le Traitement des Langues Naturelles (TALN)*. 2010.
27. **Marine Carpuat** and Mona Diab. “Task-based Evaluation of Multiword Expressions: a Pilot Study in Statistical Machine Translation”. *11th Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT)*. 2010
28. **Marine Carpuat** and Dekai Wu. “Evaluation of Context-Dependent Phrasal Translation Lexicons for Statistical Machine Translation”. *6th Language Resources and Evaluation Conference (LREC)*. 2008
29. **Marine Carpuat** and Dekai Wu. “Context-Dependent Phrasal Translation Lexicons for Statistical Machine Translation”. *Machine Translation Summit XI*. 2007
30. **Marine Carpuat** and Dekai Wu. “How Phrase Sense Disambiguation outperforms Word Sense Disambiguation for Statistical Machine Translation”. *11th Conference on Theoretical and Methodological Issues in Machine Translation (TMI)*. 2007

31. **Marine Carpuat** and Dekai Wu. “Improving Statistical Machine Translation using Word Sense Disambiguation”. *2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL)*. 2007.
32. **Marine Carpuat** and Dekai Wu. “Evaluating the Word Sense Disambiguation Performance of Statistical Machine Translation”. *Second International Joint Conference on Natural Language Processing (IJCNLP)*. 2005
33. **Marine Carpuat** and Dekai Wu. “Word Sense Disambiguation vs. Statistical Machine Translation”. *43rd Annual Meeting of the Association for Computational Linguistics (ACL)*. 2005.
34. Weifeng Su, **Marine Carpuat**, and Dekai Wu. “Semi-Supervised Training of a Kernel PCA Model for Word Sense Disambiguation”. *20th International Conference on Computational Linguistics (COLING)*. 2004
35. Dekai Wu, Grace Ngai, and **Marine Carpuat**. “Why Nitpicking Works: Evidence for Occam’s Razor in Error Correctors”. *20th International Conference on Computational Linguistics (COLING)*. 2004
36. Dekai Wu, Weifeng Su, and **Marine Carpuat**. “A Kernel PCA Method for Superior Word Sense Disambiguation”. *42nd Annual Meeting of the Association for Computational Linguistics (ACL)*. 2004
37. Dekai Wu, Grace Ngai, and **Marine Carpuat**. “Raising the Bar: Stacked Conservative Error Correction Beyond Boosting”. *Fourth International Conference on Language Resources and Evaluation (LREC)*. 2004
38. Lufeng Zhai, Pascale Fung, Richard Schwartz, **Marine Carpuat** and Dekai Wu. “Using N-best Lists for Named Entity Recognition from Chinese Speech”. *Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT)*. 2004
39. Dekai Wu, Grace Ngai, and **Marine Carpuat**. “N-fold Templated Piped Correction”. *First International Joint Conference on Natural Language Processing (IJCNLP)*. 2004
40. Dekai Wu, Grace Ngai, and **Marine Carpuat**. “A Stacked, Voted, Stacked Model for Named Entity Recognition”. *Computational Natural Language Learning (CoNLL)*. 2003
41. Dekai Wu, Grace Ngai, **Marine Carpuat**, Jeppe Larsen, and Yongsheng Yang. “Boosting for Named Entity Recognition”. *Computational Natural Language Learning (CoNLL)*. 2002
42. Grace Ngai, **Marine Carpuat** and Pascale Fung. “Identifying Concepts Across Languages: A First Step towards a Corpus-based Approach to Automatic Ontology Alignment”. *19th International Conference on Computational Linguistics (COLING)*. 2002
43. **Marine Carpuat**, Grace Ngai, Pascale Fung, and Kenneth Church. “Creating a Bilingual Ontology: a Corpus-based Approach for Aligning WordNet and HowNet”. *1st Global WordNet conference*. 2002

IV.D Refereed Workshop Papers

1. Douglas Oard, Petra Galuscakova, Kathleen McKeown, **Marine Carpuat**, Mohamed Elbadrashiny, Ramy Eskander, Kenneth Heafield, Efsun Kayi, Chris Kedzie, Smaranda Muresan, Suraj Nair, Xing Niu, Dragomir Radev, Anton Ragni, Han-Chin Shing, Yan Virin, Weijia Xu, Rui Zhang, Elena Zotkina, Joseph Barrow and Mark Gales. “Surprise Languages: Rapid-Response Cross-Language IR”. *9th International Workshop on Evaluating Information Access (EVIA)*. 2019.
2. Xing Niu, Michael Denkowski and **Marine Carpuat**. “Bi-Directional Neural Machine Translation with Synthetic Parallel Data”. *ACL Workshop on Neural Machine Translation*. 2018.
3. Alexander Zhang and **Marine Carpuat**. “UMD at SemEval-2018 Task 10: Can Word Embeddings Capture Discriminative Attributes?”. *12th International Workshop on Semantic Evaluation (SemEval)*. 2018.
4. Xing Niu and **Marine Carpuat**. “Discovering Stylistic Variations in Distributional Vector Space Models via Lexical Paraphrases”. *Workshop on Stylistic Variation at EMNLP*. 2017.

5. **Marine Carpuat**, Yogarshi Vyas and Xing Niu. “Detecting Cross-Lingual Semantic Divergence for Neural Machine Translation”. *ACL Workshop on Neural Machine Translation*. 2017.
6. Nathan Schneider, Dirk Hovy, Anders Johannsen and **Marine Carpuat**. “SemEval-2016 Task 10: Detecting Minimal Semantic Units and their Meanings (DiMSUM)”. *10th International Workshop on Semantic Evaluation (SemEval)*. 2016.
7. Xing Niu and **Marine Carpuat**. “The UMD Machine Translation Systems at IWSLT 2016: English-to-French Translation of Speech Transcripts”. *International Workshop on Spoken Language Translation*. 2016.
8. Amittai Axelrod, Yogarshi Vyas, Marianna Martindale, and **Marine Carpuat**. “*n*-gram Language Difference Models for Data Selection”. *International Workshop on Spoken Language Translation*. 2015.
9. Amittai Axelrod, Ahmed Elgohary, Marianna Martindale, Khanh Nguyen, Xing Niu, Yogarshi Vyas, and **Marine Carpuat**. “The UMD Machine Translation Systems at IWSLT 2015”. *International Workshop on Spoken Language Translation*. 2015.
10. **Marine Carpuat**. “Connotation in Translation”. *6th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*. 2015.
11. Cyril Goutte, Michel Simard and **Marine Carpuat**. “CNRC-TMT: Second Language Writing Assistant System Description”. *International Workshop on Semantic Evaluation (SEMEVAL)*. 2014.
12. Cyril Goutte, Serge Léger and **Marine Carpuat**. “The NRC System for Discriminating Similar Languages”. *Workshop on Applying NLP Tools to Similar Languages, Varieties and Dialects at COLING*. 2014.
13. **Marine Carpuat**, Cyril Goutte and George Foster. “Linear Mixture Models for Robust Machine Translation”. *Workshop on Statistical Machine Translation*. 2014.
14. **Marine Carpuat**. “NRC: A Machine Translation Approach to Cross-Lingual Word Sense Disambiguation (SemEval 2013 Task 10)”. *7th International Workshop on Semantic Evaluation at NAACL-HLT*. 2013
15. **Marine Carpuat**. “A Semantic Evaluation of Machine Translation Lexical Choice”. *7th Workshop on Syntax, Semantics and Structure in Statistical Translation at NAACL-HLT*. 2013.
16. Cyril Goutte, Serge Léger and **Marine Carpuat**. “Feature Space Selection and Combination for Native Language Identification”. *8th Workshop on Innovative Use of NLP for Building Educational Applications at NAACL-HLT*. 2013.
17. **Marine Carpuat** and Michel Simard. “The Trouble with SMT Consistency”. *NAACL-HLT Workshop on Statistical Machine Translation*. 2012
18. **Marine Carpuat**. “One Translation per Discourse”. *NAACL-HLT Workshop on Semantic Evaluations: Recent Achievements and Future Directions*. 2009
19. **Marine Carpuat**. “Toward using Morphology in French-English Phrase-Based SMT”. *EACL Workshop on Statistical Machine Translation*. 2009
20. Yihai Shen, Chi-kiu Lo, **Marine Carpuat**, and Dekai Wu. “HKUST Statistical Machine Translation Experiments for IWSLT 2007”. *Fourth International Workshop on Spoken Language Translation*. 2007
21. Xiaofeng Yu, **Marine Carpuat**, and Dekai Wu. “Boosting for Chinese Named Entity Recognition”. 5th SIGHAN Workshop on Chinese Language Processing. 2006
22. Dekai Wu, **Marine Carpuat**, and Yihai Shen. “Inversion Transduction Grammar Coverage of Arabic-English Word Alignment for Tree-Structured Statistical Machine Translation”. *IEEE/ACL 2006 Workshop on Spoken Language Technology*. 2006
23. **Marine Carpuat**, Yihai Shen, Xiaofeng Yu, and Dekai Wu. “Toward Integrating Word Sense and Entity Disambiguation into Statistical Machine Translation”. *Third International Workshop on Spoken Language Translation*. 2006

24. **Marine Carpuat**, Weifeng Su, and Dekai Wu. “Augmenting Ensemble Classification for Word Sense Disambiguation with a Kernel PCA Model”. *Third International Workshop on the Evaluation of Systems for the Semantic Analysis of Text (Senseval-3)*. 2004
25. Grace Ngai, Dekai Wu, **Marine Carpuat**, Chi-Shing Wang, and Chi-Yung Wang. “Semantic Role Labeling with Boosting, SVMs, Maximum Entropy, SNOW, and Decision Lists”. *Third International Workshop on the Evaluation of Systems for the Semantic Analysis of Text (Senseval-3)*. 2004
26. Richard Wicentowski, Grace Ngai, Dekai Wu, **Marine Carpuat**, Emily Thomforde, and Adrian Packel. “Joining forces to resolve lexical ambiguity: East meets West in Barcelona”. *Third International Workshop on the Evaluation of Systems for the Semantic Analysis of Text (Senseval-3)*. 2004

V Funding

- Maryland Catalyst Fund New Directions Award. “Smart Machine Translation With Social Sensitivity: Facilitating Workplace Inclusion Through Socio-technical Solutions”, co-PI with Ge Gao. \$37,909. 2019–2020.
- National Science Foundation. CAREER: Semantic Divergences Across the Language Barrier. \$549,930. 2018–2022.
- Brain and Behavior Initiative Seed Grant. “Cognitive and Neural Precursors to Semantic Word Learning and Math Development”, co-PI with Richard William Prather, Donald Joseph Bolger and Joseph Dien. \$53,892. 2018–2019.
- Amazon Web Services Machine Learning Award. \$50,000. 2018.
- IARPA. System for Cross-Language Information Processing, Translation, and Summarization (SCRIPTS). Co-Investigator of UMD subcontract from Columbia University with PI Doug Oard. \$3,540,186. 2017–2021.
- National Science Foundation. Student Travel Support for 2017 Workshop for Women and Underrepresented Minorities in NLP. \$20,000. 2017–2018.
- National Science Foundation. ACL 2017 Student Research Workshop. \$15,000. 2017-2018.
- Amazon Academic Research Award. \$57,899. 2016.
- Google Faculty Research Award. \$50,000. 2016.

VI Teaching, Mentoring and Advising

VI.A Courses Taught

- CMSC470: Introduction to Natural Language Processing. Department of Computer Science. Fall 2018 (16 students), Fall 2019 (36 students).
- CMSC422: Introduction to Machine Learning. Department of Computer Science. Spring 2016 (75 students), Spring 2017 (80 students), Spring 2018 (87 students).
- CMSC723: Computational Linguistics I. Department of Computer Science. Fall 2015 (31 students), Fall 2016 (45 students), Fall 2017 (35 students).
- CMSC828I: Advanced Topics in Information Processing (Multilingual Natural Language Processing). Spring 2015.

VI.B Teaching Innovations

- Introduced new undergraduate natural language processing course (CMSC470) in Fall 2018.

VI.C Advising

VI.C.1 Undergraduate

- Alexander Zhang (Computer Science). 2017-2018.
 - Alexander Zhang and Marine Carpuat. “UMD at SemEval-2018 Task 10: Can Word Embeddings Capture Discriminative Attributes?”. 12th International Workshop on Semantic Evaluation (SemEval). 2018.
- Kaman Phamdo (Computer Science & Mathematics). Fall 2017. Independent Study.
- Barbara Vasilchenko (Computer Science & Linguistics). Summer 2016. Research Assistant.
- Phong Dinh (Computer Science). Fall 2016-Spring 2017.

- 3 papers by UMD undergraduate students accepted at the 2017 Mid-Atlantic Student Colloquium on Speech, Language and Learning
 - Semantic Evaluation of Cross-lingual Word Representations Within and Across Languages, by Alexander Zhang, Yogarshi Vyas and Marine Carpuat.
 - A Case Study on Low Resource Neural Machine Translation: Translating TED talks from English to Vietnamese, by Phong Dinh and Marine Carpuat.
 - Unsupervised Induction of Word Connotation from Large Corpora, by Barbara Vasilchenko and Marine Carpuat

VI.C.2 Doctoral

PhD Advisor (Committee Chair) for

- Xing Niu (PhD, Computer Science): defended on 7/19/2019
- Yogarshi Vyas (PhD, Computer Science): defended on 9//2019
- Marianna Martindale (PhD, iSchool): passed integrative paper review on 10/26/2017
- Aquia Richburg (PhD, AMSC)
- Weijia Xu (PhD, Computer Science)

Independent Study Advisor

- Eleftheria Briakou (Spring 2019, Computer Science)
- Sweta Agrawal (Fall 2018 and Spring 2019, Computer Science)
- Stephanie Zhou (Fall 2018 and Spring 2019, Computer Science)
- Carolin Arnold (Spring-Fall 2018, Computer Science)
- Ahmed Elgohary Ghoneim (Spring-Fall 2016, Computer Science)

Other Thesis Committees

- Ann Currey. PhD, Informatics. University of Edinburgh. 2019.
- Weiwei Yang. PhD in Computer Science. University of Maryland. 2019.
- Sudha Rao. PhD in Computer Science. University of Maryland. 2018.
- Jinfeng Rao. PhD in Computer Science. University of Maryland. 2018.
- Hua He. PhD in Computer Science. University of Maryland. 2018.
- Mohit Iyyer. PhD in Computer Science. University of Maryland. 2017.
- Jessie Li. PhD in Computer Science. University of Pennsylvania. 2017.
- Benjamin Marie. PhD in Computer Science. Université Paris-Sud. 2016.
- Snigdha Chatuverdi. PhD in Computer Science. University of Maryland. 2016.
- Ann Clifton. PhD in Computer Science. Simon Fraser University. 2015.

PhD Proposal Committees

- Nidhi Shah. PhD in Computer Science. University of Maryland. 2019.
- Shi Feng. PhD in Computer Science. University of Maryland. 2019.
- Pallabi Ghosh. PhD in Computer Science. University of Maryland. 2019.
- Ugur Koc. PhD in Computer Science. University of Maryland. 2018.
- Shuoyang Ding. PhD in Computer Science. Johns Hopkins University. 2018.
- Jinfeng Rao. PhD in Computer Science. University of Maryland. 2017.
- Jessie Li. PhD in Computer Science. University of Pennsylvania. 2016.

VI.D Other Teaching Activities

- Local organizer of summer school and hackathon “Machine Translation Marathon in the Americas”. 2019.
- Lectures on “Introduction to Machine Translation” at The Cornell, Maryland, Max Planck Pre-doctoral Research School. 2019.

- Invited talk at UMD Department of Computer Science Diversity and Inclusion Speaker Series. 2019.
- Invited talk at Artificial Intelligence, Culture and Storytelling Symposium at Morgan State University. 2018.
- Gemstone thesis committee member. 2018.
- Faculty mentor at the CRA-W Graduate Cohort Workshop. 2017.
- Faculty advisor for the Student Research Workshop (SRW) at ACL 2017.
- Lecture on “Semantics & Neural Machine Translation”, Machine Translation Marathon in the Americas (Dayton, Ohio, USA), May 2017.
- Lecture on “Semantics and Machine Translation”, Machine Translation Marathon in the Americas, University of Notre Dame, June 2016.
- Outreach talks to the UMD Language Science Community (Winter Storm, Jan 2015, Language Science Day, Sep 2015, PULSAR Undergraduate Seminar in Language Science Dec 2016).

VII Service

VII.A Editorial Boards and Reviewing Activities

VII.A.1 Journal Editorial Board Member

- *Machine Translation Journal* (2019–Present)
- *Computational Linguistics* (2016–2017)

VII.A.2 Journal Reviewer

- *Transactions of the Association for Computational Linguistics* (Standing reviewer team): 2014–2018
- *ACM Transactions on Asian Language and Information Processing*: 2013, 2010, 2008
- *ACM Transactions on Speech and Language Processing*: 2014, 2013
- *Computational Linguistics*: 2018, 2016, 2015, 2012, 2010, 2009
- *Language Resources and Evaluation*: 2013
- *Machine Translation*: 2014, 2011, 2010, 2009
- *Journal of Natural Language Engineering*: 2011
- *Traitement du Langage Naturel*: 2010, 2009
- *Journal of Artificial Intelligence Research*: 2008
- *Learning Machine Translation*, edited volume in the “NIPS Workshop Series”: 2008
- *Computational Intelligence*: 2008
- *Information Research*: 2008
- *Linguistic Issues in Language Technology*: 2008

VII.A.3 Reviewing Activities for Conferences

Senior Program Committee Roles

- Area chair for Machine Translation and Multilinguality at the 2019 Conference for Empirical Methods in Natural Language Processing (EMNLP 2019)
- Area chair for Machine Translation at the 2019 Meeting of the Association for Computational Linguistics (ACL 2019)
- Tutorial co-chair for Empirical Methods in Natural Language Processing (EMNLP 2019)
- Steering Committee for the International Workshop on Semantic Evaluations (2019, 2018)
- Area chair for the 2018 Conference on Computational Natural Language Learning (CoNLL 2018)
- Area chair for Machine Translation at the 2018 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL 2018)
- Area chair for Machine Translation and Multilinguality at the 2017 Conference for Empirical Methods in Natural Language Processing (EMNLP 2017)
- Co-chair of the International Workshop on Semantic Evaluations (2016, 2017)
- Area chair for Machine Translation and Multilinguality at the 2015 Conference for Empirical Methods in Natural Language Processing (EMNLP 2015)

- Co-chair of the Workshop on Discourse in Machine Translation (2015)
- Area chair for Machine Translation at the 2013 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL 2013)
- Co-chair of the Workshop on Syntax, Semantics and Structure in Statistical Translation: SSST-5 at ACL 2011, SSST-6 at ACL 2012, SSST-7 at NAACL 2013, SSST-8 at EMNLP 2014, SSST-9 at NAACL 2015
- Co-chair of the Student Research Workshop at the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics (COLING/ACL 2006 Student Research Workshop)

Conference/Workshop Reviewer Roles

- Association for Computational Linguistics (ACL): 2018, 2017, 2015, 2014, 2011, 2010
- North American Chapter of the Association for Computational Linguistics (NAACL): 2019, 2015, 2012, 2011, 2010, 2009
- European Chapter of the Association for Computational Linguistics (EACL): 2014, 2012, 2009
- Empirical Methods in Natural Language Processing (EMNLP): 2016, 2014, 2013, 2012 (Best Reviewer Award), 2011, 2009, 2006, 2004
- Joint Conference on Lexical and Computational Semantics (*SEM): 2015, 2013
- Machine Translation Summit: 2015, 2013
- International Joint Conference on Natural Language Processing (IJCNLP): 2013, 2011
- Conference of the European Association for Machine Translation (EAMT): 2012
- Conference of the Association for Machine Translation in the Americas (AMTA): 2018, 2014, 2012
- Conference on Computational Linguistics (Coling): 2018, 2014, 2012, 2010, 2008
- Workshop on Discourse in Machine Translation: 2019
- Human in the Loop: Workshop on Humans and Computer-assisted Translation: 2014
- Workshop/Conference on Statistical Machine Translation (WMT): 2017, 2015, 2014, 2013, 2012, 2011, 2010
- Workshop on Ethics in Natural Language Processing: 2017
- Workshop on Semantic Evaluation (SemEval): 2018, 2014, 2013
- Workshop on Hybrid Approaches to Translation: 2013
- Workshop on Syntax and Structure in Statistical Machine Translation (SSST): 2010, 2007
- Student Research Workshops: NAACL-HLT (2009), ACL (2008, 2005, 2003), AMTA (2008)

VII.A.4 Reviewing Activities for Agencies and Foundations

- National Science Foundation panel and ad-hoc reviews: 2019, 2018, 2017
- Netherlands Organisation for Scientific Research: ad-hoc review 2018
- Swiss National Science Foundation: ad-hoc review 2018
- Selection Panel for Jelinek Memorial Workshop on Speech and Language Technology: 2016, 2017
- European Commission Panel: 2013

VII.B Committees, Professional & Campus Service

VII.B.1 Campus Service – Department

- Teaching Evaluation Committee (2017–2018, 2018–2019)
- Teaching Awards Committee (2017–2018, 2018–2019)
- Chair Search Committee (2017)
- Faculty Search Committee (2016)
- Lecturer Hiring Committee (2015–2016)
- Faculty advisor for the Association for Women in Computing (2015–present)

VII.B.2 Campus Service – Other

- Organizer of Computational Linguistics & Information Processing Colloquium (2017–2019)
- UMIACS APT Committee (2018–2019)
- UMIACS Steering Committee (2017–2018)

VII.B.3 Other Non-University Committees, Memberships, Panels, etc.

- Founding Chair for the Broad Interest Group for Equity in ACL (EquiCL) 2019
- Board member of the ACL Special Interest Group on Lexical Semantics (SIGLEX) 2016-2019
- Advisory board member for the first workshop for Women in Natural Language Processing (WiNLP 2017)
- Ph.D. Thesis Award Panel for ATALA (Association pour le Traitement Automatique des Langues) 2011

VIII Talks

VIII.A Keynotes

1. *Semantic, Style and Other Data Divergences in Neural Machine Translation*, Conference on Machine Translation (Florence, Italy). August 2019.
2. *Semantic and Stylistic Divergences in Machine Translation*, Machine Translation Summit XVI (Nagoya, Japan), September 2017.
3. *Domain and Other Data Divergences in Machine Translation*, Workshop on Semitic Machine Translation (Austin, TX, USA), November 2016.

VIII.B Other Invited Talks

1. *Lost in translation: neural models for low-resource machine translation & detecting semantic divergence across languages*, University of Pennsylvania, CLunch talk (Philadelphia, PA, USA), May 2019.
2. *Lost in translation: neural models for low-resource machine translation & detecting semantic divergence across languages*, Graduate Center, City University of New York, NLP Colloquium (New York, NY, USA), April 2019.
3. *Lost in translation: neural models for low-resource machine translation & detecting semantic divergence across languages*, Columbia University NLP Colloquium (New York, NY, USA), March 2019.
4. *Semantic and Stylistic Divergences in Machine Translation*, New York University “Text as Data” Colloquium Series (New York, NY, USA), October 2018.
5. *Semantic and Stylistic Divergences in Machine Translation*, Carnegie Mellon University Language Technologies Institute (Pittsburgh, PA, USA), September 2018.
6. *The Promise of Natural Language Processing*, AI, Culture and Storytelling Workshop. Morgan State University (Baltimore, MD, USA), April 2018.
7. *Semantic Divergences across Languages*, Natural Language Understanding Workshop at Google Research (New York, NY, USA), July 2017.
8. *Detecting Meaning Divergences in Bilingual Text for Machine Translation*, Amazon Machine Translation Workshop (Seattle, WA, USA), May 2017.
9. *Toward Natural Language Inference Across the Language Barrier*, Microsoft Research (New York, NY, USA), December 2016.
10. *Toward Natural Language Inference Across the Language Barrier*, Linguistics Department, Georgetown University (Washington DC, USA), November 2016.
11. *Domain and Discourse Effects on Machine Translation Lexical Choice*, NII Shonan Meeting on Synchronous Machine Translation. (Shonan, Japan), August 2015
12. *Global and Local Models for Translating across Domains*, Center of Excellence in Human Language Technologies, Johns Hopkins University (Baltimore, MD, USA), June 2015
13. *Second Languages as Noisy Supervision for Natural Language Processing*, Language Science Center, University of Maryland (College Park, MD, USA), January 2015
14. *Machine Translation for/of Code-Switched Data*, panelist at Workshop on code-switching (Doha, Qatar), October 2014

15. *Second Languages as Meaning Annotation*, Department of Computer Science, University of Maryland (College Park, MD, USA), September 2014
16. *Domain Adaptation in Machine Translation*, Keynote talk at the Ninth Machine Translation Marathon (Trento, Italy), September 2014
17. *Second Languages as Noisy Supervision for Natural Language Processing*, Natural Language Processing Seminar, University of Sheffield (Sheffield, UK), September 2014
18. *Second Languages as Noisy Supervision for Natural Language Processing*, Yahoo! Labs (New York, NY, USA), July 2014
19. *Context Models for Meaning Preserving Machine Translation*, Institute for Language, Cognition and Computation, University of Edinburgh (UK), April 2014
20. *Context Models for Meaning Preserving Machine Translation*, Department of Linguistics and Center for Translation Studies, University of Illinois at Urbana-Champaign (Urbana, IL, USA), February 2014
21. *Making Sense in Translation: Addressing Lexical Choice Errors When Translating Across Domains*, Department of Computer Science, Hong Kong University of Science & Technology (Hong Kong SAR), October 2013
22. *Making Sense in Translation: Addressing Lexical Choice Errors When Translating Across Domains*, Department of Computer Science, University of Ottawa, September 2013
23. *Discriminative Translation Lexical Choice as Sense Disambiguation*, AT&T Labs (Florham Park, NJ, USA), February 2013
24. *Discriminative Translation Lexical Choice as Sense Disambiguation*, IBM Thomas J. Watson Research Center (Yorktown Heights, NY, USA), February 2013
25. *Domain Adaptation in Machine Translation*, Keynote talk at the tenth Biennial Conference of the Association for Machine Translation in the Americas (San Diego, CA, USA), October 2012
26. *Phrase Sense Disambiguation for Context-rich Lexical Choice in Statistical Machine Translation*, Linguistics Department, University of Stuttgart (Stuttgart, Germany), April 2012
27. *Toward Learning Corpora Structure for Statistical Machine Translation*, Department of Computer Science, Hong Kong University of Science & Technology (Hong Kong SAR), October 2011
28. *Why context-dependent lexical choice in SMT is nothing more than word sense disambiguation*, panel on semantics in statistical translation at SSST-5 (Portland, OR, USA), July 2011
29. *Toward Semantic Statistical Machine Translation*, Institute for Logic, Language and Computation, University of Amsterdam (Amsterdam, Netherlands), July 2010
30. *Lexical Choice in Statistical Machine Translation*, Columbia University (New York, NY, USA), May 2010
31. *Désambiguïsation lexicale pour une approche sémantique de la traduction automatique statistique*, LIMSI-CNRS (Orsay, France), December 2009
32. *Phrase Sense Disambiguation: Word Sense Disambiguation for Semantic Statistical Machine Translation*, AT&T Labs (Florham Park, NJ, USA), November 2009