

Lifeng Jin

✉ jin.544@osu.edu

🌐 <https://lifengjin.github.io>



Education

- 2014 – Now **Ph.D. candidate, The Ohio State University, USA** in Computational Linguistics and Computational Psycholinguistics.
Areas of research: unsupervised learning of syntax, parsing with cognitive constraints, sentence classification in the medical domain.
Interests: Bayesian models. deep generative models for language. NLP for low-resource languages.
- 2006 – 2008 **M.A. University of Sheffield, UK** in Intercultural Communication.
Main focus: the effect of negative transfer from first language to second language in second language learners.
- 2002 – 2006 **B.A. Beijing Language and Cultural University, China.** major in Teaching Chinese as a Second Language, minor in Computer Science.

Refereed Research Publications

- 1 **Jin, L., Doshi-Velez, F., Miller, T., Schwartz, L., & Schuler, W. (2019).** Unsupervised Learning of PCFGs with Normalizing Flow. In *ACL*.
- 2 **Jin, L. & Schuler, W. (2019).** Variance of average surprisal: a better predictor for quality of grammar from unsupervised PCFG induction. In *ACL*.
- 3 Maicher, K., Zimmerman, L., Wilcox, B., Liston, B., Cronau, H., Macerollo, A., ... Danforth, D. (2019). Using Virtual Standardized Patients to Accurately Assess Information Gathering Skills in Medical Students. *Medical Teacher*.
- 4 **Jin, L., Doshi-Velez, F., Miller, T. A., Schuler, W., & Schwartz, L. (2018).** Depth-bounding is effective: Improvements and evaluation of unsupervised PCFG induction. In *EMNLP*. [🔗 https://github.com/lifengjin/dimi%7B%5C_%7Demnlp18](https://github.com/lifengjin/dimi%7B%5C_%7Demnlp18)
- 5 **Jin, L., King, D., Hussein, A., White, M., & Danforth, D. (2018).** Using Paraphrasing and Memory-Augmented Models to Combat Data Sparsity in Question Interpretation with a Virtual Patient Dialogue System. In *Proceedings of the 13th workshop on innovative use of nlp for building educational applications* (pp. 13–23). [🔗 http://www.aclweb.org/anthology/W18-0502](http://www.aclweb.org/anthology/W18-0502)
- 6 **Jin, L., Schuler, W., Doshi-Velez, F., Miller, T. A., & Schwartz, L. (2018).** Unsupervised Grammar Induction with Depth-bounded PCFG. *Transactions of the Association for Computational Linguistics TACL*. [🔗 https://github.com/lifengjin/db-pcfg](https://github.com/lifengjin/db-pcfg)
- 7 **Jin, L., White, M., Jaffe, E., Zimmerman, L., & Danforth, D. (2017).** Combining CNNs and Pattern Matching for Question Interpretation in a Virtual Patient Dialogue System. In *Proceedings of the 12th workshop on innovative use of nlp for building educational applications* (pp. 11–21). [🔗 https://aclanthology.coli.uni-saarland.de/papers/W17-5002/w17-5002%20http://aclweb.org/anthology/W17-5002](https://aclanthology.coli.uni-saarland.de/papers/W17-5002/w17-5002%20http://aclweb.org/anthology/W17-5002)
- 8 **Shain, C., Bryce, W., Jin, L., Krakovna, V., Doshi-Velez, F., Miller, T., ... Schwartz, L. (2017).** Modeling syntax acquisition via cognitively-constrained unsupervised grammar induction. In *CUNY 2017*. Cambridge.

- 9 Duan, M., **Jin, L.**, & Schuler, W. (2016). OSU_CHGCG at SemEval-2016 Task 9: Chinese Semantic Dependency Parsing with Generalized Categorical Grammar. In *Proceedings of the 10th international workshop on semantic evaluation (SemEval-2016)* (pp. 1218–1224). <http://www.aclweb.org/anthology/S16-1189>
- 10 **Jin, L.**, Duan, M., & Schuler, W. (2016). OCLSP at SemEval-2016 Task 9: Multilayered LSTM as a Neural Semantic Dependency Parser. In *Proceedings of the 10th international workshop on semantic evaluation (SemEval-2016)* (pp. 1212–1217). <http://www.aclweb.org/anthology/S16-1188>
- 11 Shain, C., Bryce, W., **Jin, L.**, Krakovna, V., Doshi-Velez, F., Miller, T., ... Schwartz, L. (2016). Memory-Bounded Left-Corner Unsupervised Grammar Induction on Child-Directed Input. In *Proceedings of COLING 2016, the 26th international conference on computational linguistics: technical papers* (pp. 964–975). <https://github.com/tmills/uhhmm>
- 12 Jaffe, E., **Jin, L.**, King, D., & van Schijndel, M. (2015). AZMAT: Sentence Similarity Using Associative Matrices. In *Proceedings of the 9th international workshop on semantic evaluation (semeval-2015)*. <http://www.aclweb.org/anthology/S15-2029>
- 13 **Jin, L.** & de Marneffe, M.-C. (2015). The Overall Markedness of Discourse Relations. In *EMNLP*. <http://aclweb.org/anthology/D15-1132>
- 14 **Jin, L.** & Schuler, W. (2015). A Comparison of Word Similarity Performance Using Explanatory and Non-explanatory Texts. In *NAACL-HLT*. <http://www.aclweb.org/anthology/N15-1101>
- 15 Ye, M., Tang, Z., Xu, J., & **Jin, L.** (2015). Recommender System for E-Learning Based on Semantic Relatedness of Concepts. *Information*, (6), 443–453. <https://www.semanticscholar.org/paper/Recommender-System-for-E-Learning-Based-on-Semanti-Ye-Tang/600c9f9249d1ce5df8cd0b96b7ea8dcf5d18387e>
- 16 Ye, M., **Jin, L.**, Tang, Z., & Xu, J. (2014a). A Semantic Recommender System for Learning Based on Encyclopedia of Digital Publication. In *Human-computer interaction international conference 2014*. https://link.springer.com/content/pdf/10.1007/978-3-319-07854-0%7B%5C_%7D34.pdf
- 17 Ye, M., **Jin, L.**, Tang, Z., & Xu, J. (2014b). Sentences Extraction from Digital Publication for Domain-Specific Knowledge Service. In *Human-computer interaction international conference 2014*. https://link.springer.com/chapter/10.1007/978-3-319-07857-1%7B%5C_%7D49
- 18 **Jin, L.**, Ye, M., & Fu, Y. (2013). Number in Chinese: A corpus-based computational investigation. In *Chinese lexical semantics workshop*. https://link.springer.com/content/pdf/10.1007/978-3-642-45185-0%7B%5C_%7D67.pdf
- 19 Ye, M., **Jin, L.**, Li, Y., Tang, Z., & Xu, J. (2013). Computing semantic relatedness for domain entities from encyclopaedias of digital publishing resource. In *2013 international conference on information engineering*.
- 20 Zhou, J., **Jin, L.**, & Han, S. (2009). Unified hierarchical iterate model of human conceptualization and cognition. In *Proceedings of the 2009 8th ieee international conference on cognitive informatics, icci 2009* (pp. 44–51). <https://scholars.opb.msu.edu/en/publications/unified-hierarchical-iterate-model-of-human-conceptualization-and>

Research Projects

- 2016 – Now **Unsupervised Grammar Induction (DARPA LORELEI)**
 Unsupervised grammar induction for low-resource languages with human memory-like constraints for efficient inference with Bayesian models.

Research Projects (continued)

- 2015 **Virtual Patient Project (HRSA, NBME and NSF)**
Question classification for interactive dialogue system for training student doctors to interact with patients with CNN-based memory networks.
- 2015 **Working memory for Sentence processing with MEG**
Experiments with magnetoencephalography for probing human memory effects while processing complex syntactic structures.

Awards and Financial Support

- 2019-2020 **Presidential fellowship**, Graduate School, The Ohio State University. Awarded for one year of full-time dissertation research.
- 2018 **First place: Flash talks**, The Ohio Supercomputer Center. Awarded for being the best flash talk in the biannual Statewide Users Group Conference of the Ohio Supercomputer Center.
- 2017 **Summer visiting student**, SLATE lab, OSU. For collaboration on the Virtual Patient project.
- CCBS GRA Fellowship**, Center for Cognitive and Brain Sciences, OSU. Awarded for the Virtual Patient project.
- 2015 **Summer visiting student**, University of Trento, Italy. Awarded for the sentence processing with the MEG project.
- Targeted Investment in Excellence grant**, Department of Linguistics, OSU. Awarded for the sentence processing with the MEG project.
- 2014 – Now **GRA and GTA**, Department of Linguistics, OSU. GTA awarded for teaching *LING 3801: Code making and code breaking*. GRA awarded for the research projects above.

Conference Program Committee

EMNLP 2019, NAACL 2019, ACL 2019, BEA 2019, COLING 2018, BEA 2018, CMCL 2017