

# YANG XU

Curriculum Vitae

Feb 2017

909 Southgate Dr, State College, PA 16801  
Phone: (814) 441-9473  
Email: yang.xu@psu.edu

The Pennsylvania State University  
College of Information Sciences and Technology  
319 IST, University Park, PA 16802

## RESEARCH INTERESTS

---

I am a computational psychologist. I am generally interested in discovering the underlying factors and hidden patterns of human behaviors using computational methods. Most of my current work focuses on a broad range of linguistic aspects of dialogue: linguistic alignment, uniform information density, and power in language. For my Ph.D. thesis proposal, I am working towards an integrated model of dialogue from perspectives of information theory, cognitive science, and social-psychology.

## EDUCATION

---

<i>Ph.D candidate (ABD)</i>	Since Aug 2013
College of Information Sciences and Technology, The Pennsylvania State University, USA	
<i>Master of Science</i>	Sep 2010 - Jul 2013
Department of Psychology, Tsinghua University, China	
<i>Bachelor of Science</i>	Sep 2006 - Jul 2010
Department of Electronic Engineering, Tsinghua University, China	

## PUBLICATIONS

---

### Under review

Yang Xu and David Reitter. Spectral analysis of information density in dialogue predicts collaborative task performance. 2017. (Under review, conference paper)

Yang Xu and David Reitter. Entropy in conversation: Towards an information-theoretic model of dialogue. 2016. (Under review, journal paper)

Yang Xu and David Reitter. Alignment follows the signals of linguistic power, but not social power. 2016. (Under review, journal paper)

### Published

Yang Xu and David Reitter. Entropy converges between dialogue participants: Explanations from an information-theoretic perspective. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 537–546, Berlin, Germany, August 2016. Association for Computational Linguistics

Yang Xu and David Reitter. Convergence of syntactic complexity in conversation. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 443–448, Berlin, Germany, August 2016. Association for Computational Linguistics

Yang Xu and David Reitter. An evaluation and comparison of linguistic alignment measures. In *Proc. Cognitive Modeling and Computational Linguistics (CMCL)*, pages 58–67, Denver, CO, 2015. Association for Computational Linguistics

Alexander G. Ororbia II, Yang Xu, Vito D’Orazio, and David Reitter. Error-correction and aggregation in crowd-sourcing of geopolitical incident information. In N. Agarwal et al., editor, *Social Computing*,

*Behavioral Modeling and Prediction*, volume 9021 of *Lecture Notes in Computer Science*, pages 381–387. Springer, 2015

David Reitter, Yang Xu, Patrick Craven, Anik Sndor, R. Chris Garrett, E. Vince Cross, and Jerry L. Franke. Cognitive models predicting surprise in robot operators. In *Proc. International Conference on Cognitive Modeling*, pages 190–191, Groningen, Netherlands, 2015

Yang Xu and Hong Li. The influence of visibility range and degree of urgency on the efficiency of evacuation: the mediating effect of herding behavior. *Studies of Psychology and Behavior*, 13(3):311–319, 2015. (Published in Chinese)

Hong Li, Yang Xu, Shi Chen, and Anqi Gao. The preference for the sources of information in simulated emergency escape judgment and decision-making. *Studies of Psychology and Behavior*, 10(6):452–458, 2012. (Published in Chinese)

## HONOR AND AWARDS

---

Outstanding Graduates Award	Tsinghua University, 2013
Distinguished Master Thesis Award	Tsinghua University, 2013
Fellowship of Excellence in College Entrance Exam	Changxing, Zhejiang, 2006

## TECHNICAL STRENGTHS

---

### *Programming Languages*

Python and R are my most frequently used languages. I also use Java for certain projects.

### *Natural Language Processing Toolkits*

I use spaCy and NLTK (Python) for most of the natural language processing tasks. For more efficiency-demanding cases, I use Stanford CoreNLP (Java) as well.

### *Data Science Toolkits*

I am familiar with the vectorization programming paradigm of R language, and its state-of-the-art fast aggregating and parallel computing packages (e.g., data.table, foreach).

I have experience using the Java APIs of Weka, which contributes to the SBP paper. I have also used machine learning toolkit scikit-learn (Python) for some course projects.

### *Web Technologies*

I have years of experience developing web-based experimenting systems: Node.js for my master thesis, and Flask for an independent study project in my PhD program.

### *Database Technologies*

I work intensively with MySQL database.