

# Shuoyang Ding

RESEARCH ASSISTANT · PHD CANDIDATE IN NATURAL LANGUAGE PROCESSING

Hackerman 225A, Johns Hopkins University, 3300 N. Charles St., Baltimore, MD 21210, United States

✉ dings@jhu.edu | 🏠 sding.org | 📧 shuoyangd | 📷 shuoyangd

## Education

---

### Johns Hopkins University

Baltimore, United States

PH.D. IN COMPUTER SCIENCE

Sep. 2014 - PRESENT

- Advisor: Philipp Koehn

### Beijing University of Posts & Telecommunications

Beijing, China

B.E. IN COMPUTER SCIENCE

Sep. 2010 - Jun. 2014

- GPA: 89.41/100 (rank 1<sup>st</sup> among 298)
- Thesis: *Prototype-driven, treebank-oriented grammar engineering*

## Experience

---

### Center for Speech and Language Processing, Johns Hopkins University

Baltimore, United States

RESEARCH ASSISTANT

Aug. 2014 - PRESENT

- Advised by Philipp Koehn and Kevin Duh
- Participated in various evaluation campaigns (WMT17', WMT16', OpenMT15')
- Currently working on interpretability and confidence measure for neural machine translation

### Microsoft

Redmond, United States

RESEARCH INTERN/CONTRACTOR, MICROSOFT TRANSLATOR

Mar. 2020 - May 2020, Sep 2020 - Present

- Mentored by Marcin Junczys-Dowmunt
- Studied data augmentation methods for building unified MT evaluation and quality estimation models

### Salesforce Research

Palo Alto, United States

INTERN, DEEP LEARNING RESEARCH

May. 2018 - Aug. 2018

- Mentored by James Bradbury
- Studied models that jointly learn to translate and to induce universal grammars for the source language

### Amazon

Pittsburgh, United States

APPLIED SCIENTIST INTERN

May. 2016 - Aug. 2016

- Mentored by Alon Lavie, Michael Denkowski and Ann Clifton
- Designed and implemented an automatic hyper-parameter optimization paradigm for machine translation systems

### Institute for Language, Cognition and Computation, University of Edinburgh

Edinburgh, Scotland

VISITING RESEARCHER

Jun. 2015 - Aug. 2015

- Worked with Philip Williams, Rico Sennrich and Matthias Huck
- Incorporated operation sequence model into the latest Moses syntax-based decoder as a feature function

### Language Computing and Web Mining Group, Peking University

Beijing, China

RESEARCH ASSISTANT

Jul. 2013 - Jun. 2014

- Advised by Weiwei Sun
- Undergraduate research on Chinese semantic parsing and word segmentation

## Service

---

- 2020 **Program Committee / Reviewer**, AACL, ACL, IWSLT, COLING, CoNLL, EMNLP
- 2019 **Program Committee / Reviewer**, NAACL, ACL, CoNLL, EMNLP, DeepLo
- 2018 **Program Committee / Reviewer**, ACL, CoNLL, EMNLP
- 2017 **Teaching Assistant**, CS 468 (Machine Translation), Johns Hopkins University
- 2017 **PhD Recruitment Committee**, Department of Computer Science, Johns Hopkins University
- 2015 **Program Committee**, Mid-Atlantic Student Colloquium on Speech, Language and Learning

## Publications

---

- Espresso: A Fast End-to-end Neural Speech Recognition Toolkit** *ASRU 2019*  
Yiming Wang, Tongfei Chen, Hainan Xu, [Shuoyang Ding](#), Hang Lv, Yiwen Shao, Nanyun Peng, Lei Xie, Shinji Watanabe, Sanjeev Khudanpur
- A Call for Prudent Choice of Subword Merge Operations in Neural Machine Translation** *MT Summit 2019*  
[Shuoyang Ding](#), Adithya Renduchintala, Kevin Duh
- An Exploration of Masking for Neural Machine Translation** *MT Summit 2019*  
Matt Post, [Shuoyang Ding](#), Marianna Martindale and Winston Wu
- Saliency-driven Word Alignment Interpretation for Neural Machine Translation** *WMT 2019*  
[Shuoyang Ding](#), Hainan Xu, Philipp Koehn
- Parallelizable Stack Long Short-Term Memory** *SPNLP 2019*  
[Shuoyang Ding](#), Philipp Koehn
- Improving End-to-end Speech Recognition with Pronunciation-assisted Sub-word Modeling** *ICASSP 2019*  
Hainan Xu, [Shuoyang Ding](#), Shinji Watanabe
- Multi-Modal Data Augmentation for End-to-end ASR (Best Student Paper Award)** *Interspeech 2018*  
Adithya Renduchintala, [Shuoyang Ding](#), Matthew Wiesner, Shinji Watanabe
- The JHU Machine Translation Systems for WMT 2017** *WMT 2017*  
[Shuoyang Ding](#), Huda Khayrallah, Philipp Koehn, Matt Post, Gaurav Kumar and Kevin Duh
- The JHU Machine Translation Systems for WMT 2016** *WMT 2016*  
[Shuoyang Ding](#), Kevin Duh, Huda Khayrallah, Philipp Koehn, and Matt Post
- Grammatical Relations in Chinese:GB-Ground Extraction and Data-Driven Parsing** *ACL 2014*  
Weiwei Sun, Yantao Du, Xin Kou, [Shuoyang Ding](#), Xiaojun Wan
- How Do Source-side Monolingual Word Embeddings Impact Neural Machine Translation?** *Preprint*  
[Shuoyang Ding](#), Kevin Duh
- Backstitch: Counteracting Finite-sample Bias via Negative Steps** *Preprint*  
Yiming Wang, Hossein Hadian, [Shuoyang Ding](#), Ke Li, Hainan Xu, Xiaohui Zhang, Daniel Povey, Sanjeev Khudanpur

## Honors & Awards

---

- 2018 **Best Student Paper Award**, International Speech Communication Association (3 selected among 700+ papers)
- 2013 **Finalist**, Interdisciplinary Contest in Modeling (7 selected among 957 participants)
- 2013 **First Class Scholarship**, Beijing University of Posts & Telecommunications
- 2012 **First Class Scholarship**, Beijing University of Posts & Telecommunications
- 2011 **National Scholarship**, Ministry of Education, China (top 1%-ranked students)

## Skills

---

- Programming Languages** Python, C++, Bash, Java, Haskell
- Tookits and Libraries** PyTorch, Theano, Fairseq, OpenNMT, Nematus, Mosesdecoder
- Natural Languages** Mandarin (Native), English (Full Professional), German (Elementary)