

# Yejin Cho

<http://yejincho.info>  
[ycho@utexas.edu](mailto:ycho@utexas.edu)

## INTERESTS

### Computational Semantics

- Narrative reasoning, event representation, natural language understanding
- Inference over knowledge graphs, representation learning from graphs

## EDUCATION

### University of Texas at Austin

*Ph.D. student*

Sep. 2018 – Present  
Austin, TX

- Ph.D. in Computational Linguistics
- Advisor: *Katrin Erk*
- Total GPA of 3.94 / 4.0

### Korea University

*Masters student*

Sep. 2015 – Feb. 2018  
Seoul, Korea

- M.A. in English Language and Literature
- Advisor: *Hosung Nam*
- Total GPA of 4.5 / 4.5
- Thesis: Functional Awareness of RNNLM-based Word Embeddings. [[pdf](#)]

### Yonsei University

*Undergraduate student*

Mar. 2011 – Aug. 2015  
Seoul, Korea

- B.A. in Korean Language and Literature
- B.A. in English Language and Literature (Double Major)
- Total GPA of 3.87 / 4.3

### University of California, Los Angeles

*Undergraduate Exchange student*

Sep. 2013 – June. 2014  
Los Angeles, CA

- Major in Linguistics
- Total GPA of 3.97 / 4.0 (*Dean's Honors list for full three quarters*)

## PUBLICATIONS

\*: Equal contribution

Yejin Cho\*, Juan Diego Rodriguez\*, Yifan Gao, and Katrin Erk. 2020. **Leveraging WordNet Paths for Neural Hypernym Prediction**. In *Proceedings of the 28th COLING*, pages 3007-3018, Barcelona, Spain (Online). [[pdf](#)]

Heejo You, Hyungwon Yang, Jaekoo Kang, Youngsun Cho, Sunghah Hwang, Yeonjung Hong, Yejin Cho, Seohyun Kim, and Hosung Nam. 2016. **Development of Articulatory Estimation Model Using Deep Neural Network**. *Phonetics and Speech Sciences*. 8:31-38. [[pdf](#)]

## RESEARCH EXPERIENCE

### Enriching Event Representations with Linguistic Knowledge

*Ongoing Project for Qualifying Paper (Advisor: Katrin Erk)*

Fall, 2020 –  
Austin, TX

- Task: Multiple-Choice Narrative Cloze (Granroth-Wilding and Clark, 2015); Given a sequence of events so far, predict the most likely event that will happen next.
- Idea: Exploit fine-grained semantic properties of predicates and arguments (e.g., volition, awareness) annotated in **Semantic Proto-Role** dataset (Reisinger et al., 2015) in learning event representations.
- Developing code for experiments.

**Hypernym Prediction in WordNet**, University of Texas, Austin  
*Research Assistant (Advisor: Katrin Erk)*  
*Collaborated with Juan Diego Rodriguez and Yifan Gao*

Spring, 2019 – Fall, 2019  
Austin, TX

- Task: Given a node in WordNet (e.g., *daisy*), predict its direct hypernym (e.g., *flower*) within the graph.
- Idea: Framed the task into **sequence generation problem** where a model generates from the encoded representation of a given hyponym (i.e., *daisy*) the (reverted) entire taxonomy path to the root (i.e., *flower* → ... → *plant* → *organism* → ... → *object* → *physical entity* → *entity*).
- Evaluated the first node in the generated chain which corresponds to the direct hypernym.
- Ran experiments with a **sequence-to-sequence** encoder-decoder with attention.
- Replicated five benchmark systems in link prediction task on WN18RR dataset by adapting the original source codes to our dataset and task setup:
  - TransE (Bordes et al., NeurIPS 2013)
  - M3GM (Pinter & Eisenstein, EMNLP 2018)
  - Poincaré embeddings (Nickel & Kiela, NeurIPS 2018)
  - CRIM for hypernym discovery (Bernier-Colborne and Barriere, SemEval-2018)
  - text2edges (Prokhorov et al., NAACL 2019)
- Achieved the **new state-of-the-art performance** in hypernym prediction task on WN18RR.
  - 4.11% test accuracy gain over the previous best model (Bernier-Colborne and Barriere, 2018).

**EMCS Laboratory**, Korea University  
*(Education, Mathematics, Computer science and Speech Laboratory)*  
*Graduate Researcher (Advisor: Hosung Nam)*

Fall, 2015 – Fall, 2017  
Seoul, Korea

- Built Korean Large Vocabulary Continuous Speech Recognition (LVCSR) system (800k vocabulary) from text and audio corpora with transcription using Kaldi speech recognition toolkit
- **Subword** (Pseudo-morpheme) **language modeling** for building Korean LVCSR system
- Language modeling experiments using SRILM and RNNLM Toolkit
- Designed and developed Korean text normalization and language preparation package for LM in Kaldi-based ASR system (KoLM) [[code](#)]
- Designed and developed rule-based Korean Grapheme-to-Phone conversion system (KoG2P) as an open python package [[code](#)]

## TEACHING EXPERIENCE

Teaching Assistantship in Computational Linguistics at *UT Austin*

- LIN350 **Computational Semantics** (*Instructor: Katrin Erk*) Spring-2020
- LIN313 **Language and Computers** (*Instructor: Jessy Li*) Fall-2018, Fall-2019

Teaching Assistantship in English Linguistics at *Korea University*

- ENGL238 English Phonetics (*Instructor: Hosung Nam*) Fall-2015, Fall-2016
- ENGL399 Studies on English Linguistics (*Instructor: Hosung Nam*) Spring-2016

Assistant Instructor for MATLAB Programming at *Korea University*

Winter-2015

- Covered basics of MATLAB programming for undergraduate students of humanities majors without previous experience in programming.
- Ran hands-on classes for over 40 students, answered questions, scored and commented on student assignments for 4 weeks.
- Led team project on web crawling and text mining of Korean lyrics.

- HONORS AND AWARDS
- Graduate School Fellowship, University of Texas at Austin, 2018
  - **National Humanities Scholarship**: Graduate Research Scholarship for Humanities and Social Sciences, Korea Student Aid Foundation (KOSAF), 2016-2017
  - Honors Scholarships, Korea University, 2016
  - Teaching Assistant Scholarships, Korea University, 2015-2017

TECHNICAL SKILLS

Languages Python, MATLAB, UNIX shell scripting, R  
Basic readability in Java, Javascript

Toolkits PyTorch, TensorFlow, Kaldi Speech Recognition Toolkit

SELECTED COURSEWORK

*At the University of Texas at Austin (Instructor):*

Topics in Natural Language Processing	(Eunsol Choi)	Fall, 2020
Natural Language Processing	(Greg Durrett)	Fall, 2019
Computational Discourse	(Junyi Jessy Li)	Fall, 2018
Philosophy of Language	(Josh Dever, Ray Buchanan)	Spring, 2020
Semantics I	(David Beaver)	Spring, 2019
Syntax I	(John Beavers)	Fall, 2018

*At Korea University:*

Natural Language Processing Applications	Spring, 2017
Applied Math for Brain and Cognitive Engineering	Spring, 2017
Numerical Linear Algebra	Fall, 2016
Introduction to Neural Networks	Spring, 2016
Introduction to Applied Mathematics (audited)	Fall, 2015

*At the University of California, Los Angeles:*

Phonetic Theory	Spring, 2014
Phonology I	Spring, 2014
Syntax I	Winter, 2014
Introduction to Phonetics	Winter, 2014
Introduction to Linguistics	Fall, 2013

*At Yonsei University:*

Korean Syntax	Spring, 2015
Korean Morphology	Fall, 2014
Korean Semantics	Spring, 2013
Introduction to Korean Linguistics	Fall, 2012
Korean Phonetics	Spring, 2012

LANGUAGE PROFICIENCY

Native in Korean  
Fluent in English (IBT TOEFL 116; Reading: 30, Listening: 29, Speaking: 29, Writing: 28)  
Intermediate in French (DELF B1)  
Basic conversation and readability in Japanese and Modern Standard Arabic (MSA)