Jiyoun Kim

jiyoun.kim9218@gmail.com / kite92@skku.edu
25-2, Seonggyungwan-ro, Jongo-gu, Seoul, Republic of Korea

https://kim-ji-youn.github.io

RESEARCH INTEREST

- Computational Linguistics (Semantics, Pragmatics), Machine Translation, Natural Language Processing (NLP, Deep Learning for NLP)

EDUCATION

Mar.2018~ Feb. 2020	**Sungkyunkwan University** *Master of Arts in German Language and Literature*	Seoul, Korea

Thesis: Computational Linguistic Study on Metaphor
      - Focusing on detecting metaphorical expressions in German 'adjective-noun' phrases
Advisor: *Prof. Munpyo Hong*

Apr.2015~ Feb. 2016	**Universität Tübingen** *Department of General and Computational Linguistics* Exchange Student (Apr. 2015 ~ Aug. 2015) Visiting Student (Oct. 2015 ~ Feb. 2016)	Tübingen, Germany

Mar.2013~ Feb. 2018	**Sungkyunkwan University** *Bachelor of Arts in German Language and Literature* *Bachelor of Science in Engineering (with Samsung Convergence Software Course)*	Seoul, Korea

PUBLICATIONS

- **Conferences**

1. Y. E. Koo, **J. Y. Kim**, J. P. Hong, M. P. Hong, S-K. Choi, "Towards a Linguistically Motivated Segmentation for a Simultaneous Interpretation System." *Proceedings of the 34th Pacific Asia Conference on Language, Information and Computation (PACLIC 34)*. Hanoi, Vietnam. 2020. – *oral presentation*

2. Y. E. Koo, **J. Y. Kim**, J. P. Hong, M. P. Hong, S-K. Choi, "Segmentation Methods for Different Speech Rate in Simultaneous Interpretation." *Proceedings of the 32nd Annual Conference on Human and Cognitive Language Technology (HCLT 32)*. Seoul, Korea. 2020. – *oral presentation* [in Korean]

3. **J. Y. Kim**, Y. E. Koo, Y. J. Zhu, "A Study for Categorizing Relations between Headword and Aliases." *Proceedings of the Association for Information Science and Technology (ASIS&T 2019)*. Melbourne, Australia. 2019. – *poster presentation*

4. Y. E. Koo, **J. Y. Kim**, M. P. Hong, "Automatic Speech Act Classification based on the Hierarchical Structure of Speech Act Categories." *Proceedings of the 33rd Pacific Asia Conference on Language, Information and Computation (PACLIC 33)*. Hakodate, Japan. 2019. – *poster presentation*

5. Y. E. Koo, **J. Y. Kim**, M. P. Hong, "A Study on Segmentation Unit for Real-time Simultaneous Interpretation System." *Proceedings of the 31st Annual Conference on Human and Cognitive Language Technology (HCLT 31)*. Daejeon, Korea. 2019. – *oral presentation* [in Korean]

- **Journals**

1. Y. E. Koo, **J. Y. Kim**, M. P. Hong, Y. K. Kim, "A Linguistic Study of Automatic Speech Act Classification for Korean Tutorial Dialog." *Journal of Korean Institute of Information Scientists and Engineers (KIISE).* 45(8). pp. 807-815. 2018. [in Korean]

PATENTS

1. **J. Y. Kim**, S. Y. Moon, Y. T. Kim, J. Y. Seo, J. P. Choi, K. H. Hong, "A System for Generating Catchphrase using Movie Synopsis", Republic of Korea (Apply: 10-2019-0040667)

WORK EXPERIENCE

Feb.2021~ Present	**NC Soft corp.,** Language AI Lab *Intern*	Gyeonggi-do, Korea

- Working as developer intern in Machine Translation Data team

RESEARCH EXPERIENCE

Mar. 2016 ~ Nov. 2020	**Electronics and Telecommunication Research Institute (ETRI)** *Research Assistant (Advisor: Prof. Munpyo Hong)*	Daejeon, Korea

- Applied segmentation strategies for simultaneous interpretation system for English and Korean
- Demonstrated correlation between length of segmentation unit and translation rate depending on language pairs
- Built cognitive-pragmatic model for simultaneous interpretation
- Designed and developed algorithm to implement ideal segmentation strategy to detect optimal segmentation units
- Results of research were applied to *Genie Talk*, the official translation app for the 2018 Pyeongchang Winter Olympic
- Papers related to this research were accepted in *HCLT 31*, *HCLT 32*, and *PACLIC 34*

Mar. 2017 ~ Dec. 2019	**SK Telecom Co., Ltd** *Research Assistant (Advisor: Prof. Munpyo Hong)*	Seoul, Korea

- Participated in as project manager
- Established guidelines for corpus generation and error detection
- Constructed 487,834 sentences-size query corpus for training Korean dialogue systems in K-pop domain
- Detected and categorized error types in morphology analysis of Named Entities (128,133 words in total) and query sentences (16,868 sentences in total) to improve SKT Korean Language Processing engine
- Suggested possible solutions to the errors
- Gave poster presentation on categorizing relations between headwords and aliases at *ASIS&T 2019*

Jul. 2019 ~ Jan. 2020	**Institution of Language and Information, Yonsei University** *Research Assistant*	Seoul, Korea

- Contributed to establish manuals to build co-reference resolution corpus in written Korean
- Built 63,000 tokens-size corpus for co-reference resolution in written Korean
- Scrutinized 270,000 tokens of the corpus (full corpus includes 2,000,000 tokens in total)

ACADEMIC EXPERIENCE / TEACHING

Mar. 2018 ~ Aug. 2018	**Introduction to Computational Linguistics (GER3025)**, Sungkyunkwan University *Teaching Assistant*	Seoul, Korea

- Covered Python programming and other natural language processing tools: TreeTagger and Antconc
- Answered questions in person and online, gave advice on final projects of students

Mar. 2019 ~ Aug. 2019	**Introduction to German Linguistics II (ILI2002)**, Sungkyunkwan University *Teaching Assistant*	Seoul, Korea

- Covered Python programming, Natural Language Toolkit (NLTK) and web-data crawling using HTML and BeutifulSoup
- Answered questions in person and online, gave solutions to install related Python packages and programs, gave comments and advice on final projects of students

Mar. 2019 ~ Aug. 2019	**Machine Translation and Linguistics (ILI2001)**, Sungkyunkwan University *Teaching Assistant*	Seoul, Korea

- Covered Python programming and Natural Language Toolkit (NLTK) for processing English language data

EXTRACURRICULAR ACTIVITIES

- **C-School project**
  - Collaborated with advisor professor (*Prof. Yungyung Cheong*) and students majoring in Computer Science
  - Led project on constructing model for data crawling, data pre-processing and information extraction
  - Made a system using Seq2Seq model to generate catchphrase using movie synopsis
  - Applied for patent based on the project

- **Co-Deep-Learning project**
  - Collaborated with advisor professor (*Prof. Jaekwang Kim*) and students not majoring in Computer Science
  - Organized whole curriculum to learn deep learning and discussed on weekly basis
  - Made own materials based on Lectures of *Neural Networks and Deep Learning* (on Coursera, Andrew Ng), Lectures of *Stanford University CS224n: Natural Language Processing with Deep Learning* (on YouTube) and *Deep Learning from Scratch* (S. Goki, 2017)

- **Peer Leader of Learning Community**
  - Learning community: basic unit of learning, education and college life for first-year students, consists of about 20 students, operated mainly by Sungkyun mentor (Professor), Freshman guide (Senior), and Peer leader
  - Organized self-study group, Sungkyun freshman seminar, and sports events
  - Selected as one of best LC in 2013, received scholarship (Undergraduate College Scholarship, 2013)

- **Sungkyunkwan Orchestra**
  - Second Violin
  - Started as novice and ended with Tchaikovsky, The Nutcracker

- **Yogini** certified as yoga instructor

SCHOLARSHIPS AND AWARDS

2018 ~ 2019	SimSan Scholarship (graduate students, German Language and Literature), Sungkyunkwan University, Korea
2018 ~ 2019	Graduate Merit Scholarship, Sungkyunkwan University, Korea
2018	TA Scholarship, Sungkyunkwan University, Korea
2016 ~ 2018	CORE (COllege of humanities' Research and Education) Scholarship, Sungkyunkwan University, Korea
2014	Global Humanities Program Scholarship, Sungkyunkwan University, Korea
2013	Undergraduate College Scholarship, Sungkyunkwan University, Korea
2013	Academic Excellence Scholarship, Sungkyunkwan University, Korea
2013 ~ 2016	YulGok Scholarship (Outstanding Admission Scholarship, half tuition), Sungkyunkwan University, Korea

SKILLS

- **Programming Skills**
  - Programming Language: Python (Advanced), Java (Advanced)
  - Deep Learning Libraries: PyTorch (Advanced), TensorFlow (Advanced)
  - Others: LISP, PROLOG, HTML, CSS, JavaScript
- **Languages**
  - Korean (Native), English (Fluent), German (Advanced), Chinese (Basic)
    - IBT TOFLE: 100 (Reading: 29, Listening: 26, Writing: 22, Speaking: 23)
    - TestDaf: B2-C1 (Reading: 4 / 5, Listening: 3 / 5, Writing: 4 / 5, Speaking: 3 / 5)
    - HSK: Level 3 (256 / 300)