Curriculum Vitae

Hakyung Sung

(last revised: October 2024)

Contact Information

Department of Linguistics
1600 Millrace Dr. 201
University of Oregon
Eugene, Oregon 97403-1290

- Learner Corpus Research-Applied Data Science Lab
- hsung@uoregon.edu
- http://hksung.github.io
- https://github.com/hksung

EDUCATIONS

Jun. 2025 (expected)	University of Oregon *Ph.D. Linguistics*
Dec. 2024 (expected)	University of Oregon (Concurrent degree) *M.S., Computer Science*
Feb. 2018	Seoul National University *M.A., English Education*
Feb. 2016	Ewha Womans University *B.A., English Language and Literature (with English Teaching Profession)*

HONORS / AWARDS

- *Professional Development Graduate Scholarship* (2023). University of Oregon
- *Raymund Fellowship* (2021). University of Oregon
- *Promising Scholar Award* (2021). University of Oregon
- *Lockey Graduate Science Award* (2021). University of Oregon
- *Top-off Award in Linguistics Department* (2021). University of Oregon
- *Excellence of Academics* (2017), Seoul National University
- *Scholarship of Graduate Student Instructor* (2017), Seoul National University

EXTERNAL FELLOWSHIPS / GRANTS

- PI (2024.9 - 2025.6). *Evaluating argument structure construction-based indices for construct validity in assessing (English as a) second language grammatical proficiency (Dissertation).* PI: Hakyung Sung. Supervisor: Kristopher Kyle (University of Oregon); Mikyung K. Wolf & Ikkyu Choi (ETS). Harold Gulliksen Psychometric Research Fellowship sponsored by ETS ($33,000)

- Co-researcher (2024.2 - 2025.1). *Towards inclusive NLP research: Developing language models on second language Korean*. PI: Gyu-Ho Shin (University of Illinois, Chicago). Sponsored by the Academy of Korean Studies ($13,200)

PEER-REVIEWED PUBLICATIONS

3: Journal article; †6: Conference proceedings

[9] †**Sung, H**., & Kyle, K. (*in press*). Leveraging pre-trained language models for linguistic analysis: A case of argument structure constructions. In *the Association for Computational Linguistics: EMNLP 2024*.

[8] †**Sung, H**., & Shin, G-H. (2024). Constructing a dependency treebank for second language learners of Korean. In *Proceedings of the International Conference on Language Resources and Evaluation (LREC-COLING 2024)*. https://aclanthology.org/2024.lrec-main.332/

[7] †**Sung, H**., & Kyle, K. (2024). Annotation scheme for English argument structure constructions treebank, In *Proceedings of the 18th Linguistic Annotation Workshop (LAW-XVIII)*. https://aclanthology.org/2024.law-1.2/ [presentation]

[6] **Sung, H**., Cho, S., & Kyle, K. (2024). An empirical evaluation of lexical diversity indices in L2 Korean writing assessment, *Language Assessment Quarterly,* 1-22. https://doi.org/10.1080/15434303.2024.2311728

[5] Kyle, K., **Sung, H.**, Eguchi, M., & Zenker, F. (2024). Evaluating evidence for the reliability and validity of lexical diversity indices in L2 oral task responses, *Studies in Second Language Acquisition*, 1-22. https://doi.org/10.1017/S0272263123000402

[4] †**Sung, H**., & Shin, G-H. (2023). Diversifying language models for lesser-studied languages and language-usage contexts: A case of second language Korean, In *Findings of the Association for Computational Linguistics: EMNLP 2023*. https://aclanthology.org/2023.findings-emnlp.767/

[3] †**Sung, H**., & Shin, G-H. (2023). Towards L2-friendly pipelines for learner corpora: A case of written production by L2-Korean learners, In *Proceedings of the 18th Workshop on Innovative Use of NLP for Building Educational Application*, 72-82, Association for Computational Linguistics. https://aclanthology.org/2023.bea-1.6/ [poster]

[2] †Kyle, K., & **Sung, H.** (2023). An argument structure construction treebank, In *Proceedings of the Workshop on Construction Grammars and NLP (CxGs+NLP, GURT)*, 51–62, Association for Computational Linguistics. https://aclanthology.org/2023.cxgsnlp-1.7/

[1] **Sung, H.** (2019). Korean EFL learners' processing of English caused-motion construction, *English Teaching, 74*(1), 49-73.

WORKSHOPS

- **Sung, H.** & Kyle, K. (Sep. 2024). *The construction of high-quality manually annotated datasets for fine-tuning LLMs for automatic linguistics analyses.* Workshop at 16th American Association for Corpus Linguistics Conference (AACL 2024). Eugene, OR.

PRESENTATIONS

Main presenter

[9] **Sung, H**., & Wolf, M., Suhan, M., & Kyle, K. (Mar. 2025). *Investigating the lexical characteristics of young EFL students' writing across different test tasks.* Paper presenting at AAAL 2025. Denver, CO.

[8] **Sung, H**., & Kyle, K. (Sep. 2024). *Building a robust argument structure construction (ASC) treebank and an automatic ASC tagger.* Paper presenting at AACL 2024. Eugene, OR.

[7] **Sung, H.** (Apr. 2024). *Exploring the relationship between argument structure construction usage and L2 English proficiency.* Paper presenting at LARC 2024. Provo, UT.

[6] **Sung, H**., & Kyle, K. (Mar. 2024). *The relationship between L2 oral proficiency interview scores and characteristics of argument structure construction use.* Paper presenting at AAAL 2024. Houston, TX.

[5] **Sung, H.,** & Shin, G-H. (Sep. 2023). *Comparison of morphological analyzers for L2-Korean written corpora.* Poster presenting at EuroSLA 32, Birmingham, UK.

[4] **Sung, H**., & Kyle, K. (May. 2023*). Automatic identification of motion constructions in L1 and L2 English spoken corpora*. Paper presenting at ICCG12. Prague, Czech Republic.

[3] **Sung, H.,** & Kyle, K. (Mar. 2023). *Exploring evenness lexical diversity indices in L2 written and spoken corpora.* Paper presenting at AAAL 2023. Portland, OR.

[2] **Sung, H.** (Sep. 2022). *On the development of motion constructions of Korean L2 English learners: A cross-sectional spoken corpus-based study.* Paper presenting at AACL 2022. Flagstaff, AZ.

[1] **Sung, H.,** & Cho, S. (Aug. 2022). *Evaluating lexical diversity of Korean as a second language learners' writing using NLP tool.* Paper presenting at EuroSLA 31. Fribourg, Switzerland.

Not main presenter

[9] Kyle, K., **Sung, H**., Biber, D., Reppen, R., & Egbert, J. (Mar. 2025). *The development and evaluation of an open-source lexicogrammatical complexity analysis tool: The Lexicogrammatical Tagger.* Paper presenting at AAAL 2025. Denver, CO.

[8] Diantoro, C., **Sung, H.,** & Baese-Berk, M. (May. 2024). *Differences in prominence-related acoustic measurements in L2 read speech as a function of L1 framing typology.* Poster presenting at ASA. Ottawa, Canada.

[7] Kyle, K., **Sung, H**., & Eguchi, M. (Mar. 2024). *Identifying lexico-grammatically ambiguous language features using large language models.* Invited colloquium at AAAL 2024. Houston, TX.

[6] Kyle, K., & **Sung, H**. (Mar. 2024). *The relationship between L2 holistic writing proficiency scores and characteristics of argument structure construction use.* Paper presenting at AAAL 2024. Houston, TX.

[5] Kyle, K., & **Sung, H**. (May. 2023). *Robust automatic annotation of argument structure constructions*. Paper presenting at ICCG12. Prague, Czech Republic.

[4] Kyle, K., **Sung, H.**, Eguchi, M., & Zenker, F. (Mar. 2023). *The reliability and validity of lexical diversity indices in second language oral productions.* Paper presenting at AAAL 2023. Portland, OR.

[3] Diantoro, C., **Sung, H.**, & Baese-Berk, M. (Mar. 2023). *Comparing stress patterns of verbs and satellites in L1/L2 spoken corpora based on framing typology.* Poster presenting at AAAL 2023. Portland, OR.

[2] Kyle, K., **Sung, H.**, Eguchi, M., & Zenker, F. (Sep. 2022). *Evaluating the reliability and validity of lexical diversity indices in second language oral corpora.* Paper presenting at AACL 2022. Flagstaff, AZ.

[1] Cho, S., & **Sung, H.** (Sep. 2022). *Investigating lexical diversity in L2 Korean writing: Focusing on multiple indices and Korean tokenizers.* Paper presenting at AACL 2022. Flagstaff, AZ.

MASTER'S THESIS

- Sung, H. (2018). *Korean EFL learners' processing of English caused-motion construction*. (Seoul National University). https://s-space.snu.ac.kr/handle/10371/142034

OTHER RESEARCH PRODUCTS

- *English ASC Treebank*: The ASC Treebank is a linguistically annotated dataset of argument structure constructions (ASCs), informed by research in the usage-based constructionist approach to language learning. The first version of the treebank includes 26,437 annotated ASC tokens, while the second version contains 22,069 tokens. Developed in collaboration with Kris Kyle (University of Oregon)
  - ASC treebank v1: https://github.com/LCR-ADS-Lab/ASC-Treebank
  - ASC treebank v2, annotation manual: https://osf.io/v75qu/?view_

- *Korean lexical diversity python package*: The Korean lexical diversity python package extended the previous TAALED (Tool for the Automatic Analysis of Lexical Diversity; Kyle et al., 2021) tool to calculate a wide variety of lexical diversity indices (e.g., MATTR, MTLD, & HD-D) for Korean. As an agglutinative language, Korean requires more fine-grained parsing at the morpheme level (compared to English), so we provide two versions of the morphological parser: (1) baseline Korean morphological parsers and (2) a fine-tuned L2-Korean morphological parser which improved accuracy in Korean lexical diversity analysis. Developed in collaboration with Sooyeon Cho (Zurich University of Applied Sciences): https://github.com/NLPxL2Korean/Korean-Lexical-Diversity

- *UD Korean-KSL dataset*: This UD Korean-KSL dataset is a dependency treebank of L2 Korean, featuring morpheme and Universal Dependency manual annotations. The latest version (Sep. 2024) contains 12,058 annotated sentences from the written data produced by L2-Korean

learners at various proficiency levels, with the dataset continuing to grow. Developed in collaboration with Gyu-Ho Shin (University of Illinois, Chicago): https://github.com/UniversalDependencies/UD_Korean-KSL/tree/dev

PROFESSIONAL/TEACHING EXPERIENCE

Sep. 2022 -   Graduate Employee, University of Oregon   Eugene, OR
†sole instructor/instructor/grader, or teaching assistant in the following courses:

Course	Term	Role
The Structure of English Words (LING 150)	Fall 2024	†Sole instructor (3 hours/week)
Intro to Linguistic Behavior (LING 302)	Spring 2024	Instructor: Leading discussion sessions (2 hours/week)
The Structure of English Words (LING150) *Research Methods in Applied Linguistics* (LING 430/530)	Winter 2024	Grader Assistant: Material refinement
Second Language Acquisition (LING 444/544)	Fall 2023	Instructor: Leading discussion sessions (2 hours/week)
Language & Cognition (LING 416/516) Sociolinguistics (LING 491/591)	Spring 2023	Grader
Sociolinguistics (LING 491/591)	Winter 2023	Grader
The Structure of English Words (LING 150)	Fall 2022	Grader

Jun. 2024 -   Summer Intern, Educational Testing Service (ETS)
Jul. 2024   2024 Ida Lawrence Research Summer Internship Program
- Project: *Linguistic feature extractions and statistical modeling to identify language development patterns in TOEFL Junior writing/speaking test-takers*

Jun. 2022 -   Summer Intern, The American Council on the Teaching of Foreign Languages (ACTFL)
Aug. 2022

Oct. 2021 -   Electronic Resources Student Assistant, University of Oregon Library
Sep. 2022

Mar. 2018 -   English Teacher, *Daewon* & *Gyeonggi* Foreign Language High Schools
Feb. 2021

PROFESSIONAL SERVICES

Upon request   *Ad-hoc* Reviewer (Journal manuscript; †Conference proceeding; ††Conference abstract)
- *Language Learning*

- *Language Assessment Quarterly*
- *Journal of Second Language Writing*
- *International Review of Applied Linguistics in Language Teaching*
- *Bilingualism: Language and Cognition*
- *MwALT Best Student Paper*
- †*Workshop on Innovative Use of NLP for Building Educational Application*
- ††*International American Association for Corpus Linguistics Conference*

Sep. 2024   Conference Committee, the 16th *American Association for Corpus Linguistics Conference*

Jan. 2022 -   Graduate Student Faculty Representative Liaison, Linguistics, University of Oregon
June. 2023

Mar. 2017 -   Graduate Student Instructor, Seoul National University.
Feb. 2018

OTHER SKILLS / CERTIFICATIONS

- Languages: English (Full working proficiency); Korean (Native); Spanish (Reading proficiency)
- Programming Languages: R, Python, SQL
- Documentation/Typesetting: Markdown, LaTeX
- License: Regular Teacher's License in Korean (Teaching English, 2nd degree)