

CURRICULUM VITAE

Sanghoun Song (Ph.D. in Linguistics)

Associate Dean for Planning Affairs

Director, Center for Digital Humanities

Vice-Director, Institute for Digital HUSS

sanghoun@korea.ac.kr

(+82) 2-3290-2177

<http://corpus.mireene.com>

ACADEMIC POSITIONS:

- | | |
|---------------------|---|
| Mar. 2019- | Associate Professor (2021~),
Assistant Professor (2019~2021)
Dept. of Linguistics, Korea University |
| Aug. 2015-Feb. 2019 | Assistant Professor
Dept. of English Language & Literature,
Incheon National University |
| Aug. 2014-Aug. 2015 | Research Fellow
Computational Linguistics Lab,
Division of Linguistics and Multilingual Studies,
Nanyang Technological University |
| Mar. 2014-Aug. 2014 | Research Professor
Research Institute of Language & Information, Korea University |

RESEARCH INTERESTS:

HPSG, computational syntax/semantics, deep learning, medical/legal NLP
experimental syntax, corpus-based methods for linguistic studies

TEACHING INTERESTS:

grammatical theory, computational/corpus linguistic,
programming for linguists & linguistics for programmers

EDUCATION:

- | | |
|-----------|---|
| 2009-2013 | Ph.D.
Dept. of Linguistics, Univ. of Washington
Dissertation: <i>A Grammar Library for Information Structure</i>
Committee: Prof. Emily M. Bender (chair), Prof. Toshiyuki Ogihara,
Prof. Fei Xia, Prof. Luke Zettlemoyer (GSR) |
| 2008-2009 | Ph.D. course
Dept. of Linguistics, Korea Univ. (Supervisor: Prof. Jae-Woong Choe) |
| 2006-2008 | M.A. Dept. of Linguistics, Korea Univ.
Thesis: <i>A Constraint-based Analysis of Passive Constructions in Korean</i>
Committee: Prof. Jae-Woong Choe (chair),
Prof. Jong-Bok Kim, Prof. Beom-mo Kang |
| 2007 | Visiting Research Student.
Dept. of Chinese, Translation, and Linguistics, City University of Hong Kong.
Advisor: Prof. Haihua PAN & Prof. Alex Chengyu Fang |
| 1996-2002 | B.A. Dept. of Linguistics, Korea Univ. |

PATENT:

- 2021 (application) Artificial intelligence system and method for Korean medical natural language processing based on deep learning (10-2021-0179385)
Hyung Joon Joo (PI), **Sanghoun Song (30%)**, Yoojoong Kim, Kim Jong-Ho, & Hakje Yoo
- 2023 (transferred) Artificial intelligence system and method for medical specialty recommendation using Korean medical natural language processing artificial intelligence engine based on deep learning (P2023-08-086)
Hyung Joon Joo (PI), **Sanghoun Song (20%)**, Kim Jong-Ho, & Eunbin Cho

BOOKS:

- 2017 Song, Sanghoun. *Modeling Information Structure in a Cross-linguistic Perspective*. Berlin: Language Science Press.
Open Access: <http://langsci-press.org/catalog/book/111>

JOURNAL ARTICLES:

- 2024 Kim, Jungsoo and **Sanghoun Song**. An in-depth corpus-based investigation of unbound reflexives. *English Studies*. (A&HCI, In Press)
- 2024 Cho, Eunbi, Subin Lee, and **Sanghoun Song**. Analysis of Gender Bias in Korean Language Models Utilizing the Winogender Dataset. *Korean Journal of Linguistics* 49(1): 173-205.
- 2024 Noh, Kangsan and **Sanghoun Song**. Inter-sentential Processing of Language Models: The Case of Felicity Conditions. *Studies in Linguistics* 70: 195-224.
- 2023 Yun, Hongoak, Eunkyung Yi, and **Sanghoun Song**. Exploring AI-Generated English Relative Clauses in Comparison to Human Production. *Journal of Cognitive Science* 24(4): 465-496. (SCOPUS)
- 2023 Lee, Jieun and **Sanghoun Song**. How do Korean speakers receive internally headed relative clauses?. *Linguistic Research* 40(3): 463-508. (SCOPUS)
- 2023 **Song, Sanghoun**. A Year's Change in Educational Environment Triggered by ChatGPT - Retrospect and Prospect. *Journal of AI Humanities* 15: 158-188.
- 2023 Oh, Eunjin, Sunwoo Park, Hyesun Cho, and **Sanghoun Song**. ChatGPT takes on a phonology exam: Analyzing its characteristics, errors, and reasoning ability. *Language and Information* 27(2): 73-115.
- 2023 Noh, Kangsan and **Sanghoun Song**. A Diachronic Collostructional Analysis of Phrasal Verbs. *English Language and Linguistics* 29(4): 73-93.
- 2023 Noh, Kangsan and **Sanghoun Song**. Cwuta as the Marker for the Benefactive Construction - A Constructionist Approach -. *Language Facts and Perspectives* 60: 85-105.
- 2023 Lee, Gyu-min and **Sanghoun Song**. Can Korean Language Models Detect Social Registers in Utterances? *Korean Journal of Linguistics* 48(3): 585-605.
- 2023 Lee, Jieun, Say Young Kim, and **Sanghoun Song**. What Restricts or Boosts the Use of Internally-Headed Relative Clauses in Korean. *The Linguistic Association of Korea Journal* 31(3): 113-137.
- 2023 Hong, Seung-hye, Unsub Shin, Gyu-min Lee, Dongjin Hwang, and **Sanghoun Song**. Reasoning Ability of Deep Learning Models on Tautological Expressions in Korean. *Language and Information Society* 49: 119-149.

- 2023 Lee, Jieun, Dongjin Hwang, and **Sanghoun Song**. A case study on the clause linking ability of AI language models: Utilizing the Turbo and Davinci models. *The Journal of Linguistics Science* 105: 27-61.
- 2023 Lee, Jieun, Dongjin Hwang, and **Sanghoun Song**. A case study on the clause linking ability of AI language models: Utilizing the Turbo and Davinci models. *The Journal of Linguistics Science* 105: 27-61.
- 2023 Noh, Kang San, and **Sanghoun Song**. Korean EFL Learners' Underuse of the English Perfect: A Collostructional Analysis of Written Learner Corpora. *Korean Journal of English Language and Linguistics* 23: 342-357. (SCOPUS)
- 2023 **Song, Sanghoun**. How AI Language Models Judge Common Senence. *Journal of Korean Linguistics*. 105: 375-400. (written in Korean)
- 2023 Shin, Unsub, Eunkyung Yi & **Sanghoun Song**. Investigating a neural language model's replicability of psycholinguistic experiments: a case study of NPI licensing. *Frontiers in Psychology*. 14: 937656. (SSCI)
- 2023 Kim, Yoojoong, Jong-Ho Kim, Young-Min Kim, **Sanghoun Song**, & Hyung Joon Joo. Predicting medical specialty from text based on a domain-specific pre-trained BERT. *International Journal of Medical Informatics*. 170: 104956. (SCIE)
- 2022 **Song, Sanghoun**. Using the Deep Learning Models and Surprisal for Language Analysis. *Journal of AI Humanities*. 12: 9-39. (written in Korean)
- 2022 Hwang, Dongjin, Unsub Shin, Jieun Lee, & **Sanghoun Song**. The Ability of an English Language Model to Understand of World Knowledge with Backshift. *Language and Information* 26(2): 153-169.
- 2022 Lee, Jieun, Gyu-min Lee, & **Sanghoun Song**. Does the Korean Deep Learning Model Identify the Gap in the Relative Clauses?. *DONAM OHMUNHAK* 42: 159-195.
- 2022 Yi, Eunkyung, Hyowon Cho, and **Sanghoun Song**. An Experimental Investigation of Discourse Expectations in Neural Language Models. *Korean Journal of English Language and Linguistics* 22: 1101-1115. (SCOPUS)
- 2022 Shin, Unsub, Arum Kang, and **Sanghoun Song**. A Data-based Analysis of Entailment-canceling Operators in Korean. *Studies in Linguistics* 65, 113-148.
- 2022 Lee, Gyu-min and **Sanghoun Song**. The Korean Coronavirus Corpus: A Large-Scale Analysis Using Computational Skills. *The Sociolinguistic Journal of Korea* 30(3): 213-243.
- 2022 Shin, Unsub and **Sanghoun Song**. BERT Learns More than Word Frequency Information: A Case Study of Do-Be Constructions. *Korean Journal of Linguistics* 47(3): 467-489.
- 2022 Shin, Unsub and **Sanghoun Song**. Commonsense Reasoning of Neural Language Models and Trends in Evaluation Benchmarks. *The Journal of Humanities and Social Sciences*. 23(3): 133-166. (written in Korean)
- 2022 Kim, Yoojoong, Jong-Ho Kim, Jeong Moon Lee, Moon Joung Jang, Yun Jin Yum, Seongtae Kim, Unsub Shin, Young-Min Kim, Hyung Joon Joo, & **Sanghoun Song**. A pre-trained BERT for Korean medical natural language processing. *Scientific Reports*. 12:13847. (Nature Portfolio / SCIE)
- 2022 Lee, Junghyun, HyunA Park, and **Sanghoun Song**. Computational analysis of the language usage related to COVID-19 in Chousun-Ilbo and Hangyeore. *Language Facts and Perspectives* 56, 293-330. (written in Korean)
- 2022 Park, Hayul, HyunA Park, and **Sanghoun Song**. Countering the hatred: The counter-speech dataset in Korean for evaluating hate speech detection models.

- Discourse and Cognition* 29(2), 1-23. (written in Korean)
- 2022 Hwang, Dongjin and **Sanghoun Song**. What Can BERT Do for English Linguistics?: A Case Study of Examining *It*-cleft Constructions. *English Language and Linguistics* 28(1): 135-165.
- 2022 **Song, Sanghoun**, Kang San Noh, Kwonsik Park, Un-sub Shin, & Dongjin Hwang. Adversarial example-based evaluation of how language models understand Korean case alternation. *The Linguistic Association of Korea Journal* 30(1): 45-72. (written in Korean)
- 2022 **Song, Sanghoun**. Evaluation of the Deep Learning-based Language Models and Linguistics. *Language and Information Society* 45: 169-191. (written in Korean)
- 2022 **Song, Sanghoun**. Within/Between or Latin Square: Experimental Designs of Acceptability Judgment Testing. *Journal of Language Sciences* 29(1): 339-360.
- 2022 Kim, Kyeong-min and **Sanghoun Song**. Distributional Asymmetry in Subject and Object Internally headed Relative clauses in Korean. *Korean Linguistics* 94: 57-86.
- 2022 Kang, Daeun and **Sanghoun Song**. Comparative Analysis of the Corpus, Experiment, Deep Learning Methods: Focusing on the ‘-tolok’ Control Construction. *Studies in Linguistics* 62: 133-161. (written in Korean)
- 2021 Cheong, Yunam, Guehyun Wang, & **Sanghoun Song**. Detecting the Lexical Variation between South and North Koreans Using the Deep Learning Techniques. *Korean Semantics*. 74: 113-139. (written in Korean)
- 2021 Lee, Hyungjong, Gyu-min Lee, Seok-Hoon You, & **Sanghoun Song**. A deep learning-based analysis of awareness changes in North Korea - Focusing on Rodong Sinmun Big Datas from 2016 to 2019. *North Korean Studies*. 17(2): 89-121. (written in Korean)
- 2021 Park, Kwonsik, Seongtae Kim, & **Sanghoun Song**. Verification of Korean Pre-trained Models' Feasibility of Syntactic Research Using Pairwise Sentences. *Language and Information*. 25(3): 1-21. (written in Korean)
- 2021 Noh, Kang San and **Sanghoun Song**. How Korean EFL Learners Use the Progressives: A Collostructional Analysis. *English Language and Linguistics* 27(2): 109-136.
- 2021 **Song, Sanghoun**, Gyu-Min Lee and Kyeong-min Kim. Probing the Unbound Reflexives in English via the Deep Learning-based Language Model. *Journal of Language Sciences*. 28(3): 51-78. (written in Korean)
- 2021 **Song, Sanghoun**. Different Types of Internally Headed Relative Clauses in Korean: A Corpus-based Analysis. *Studies in Linguistics*. 60: 89-114.
- 2021 Yum, Yunjin, Jeong Moon Lee, Moon Joung Jang, Yoojoong Kim, Jong-Ho Kim, Seongtae Kim, Unsub Shin, **Sanghoun Song***, and Hyung Joon Joo*. A Word Pair Dataset for Semantic Similarity and Relatedness in Korean Medical Vocabulary: Reference Development and Validation. *JMIR Medical Informatics* 9(6): e29667. (SCIE)
- 2021 Park, Kwonsik and **Sanghoun Song**. A deep learning-based understanding of nativelikeness: A linguistic perspective. *Korean Journal of English Language and Linguistics*. 21: 487-509. (SCOPUS)
- 2021 Park, Kwonsik, Myung-Kwan Park, and **Sanghoun Song**. Deep learning can contrast the minimal pairs of syntactic data. *Linguistic Research*. 38(2): 395-424. (SCOPUS)
- 2021 Kim, Seongtae and **Sanghoun Song**. Embedding Calculus with Nonword Properties Improves Word Sense Disambiguation. *Korean Journal of Linguistics*.

- 46(2): 259-292.
- 2021 Kang, Arum and **Sanghoun Song**. Semantic properties of irrealis mood in question: non-veridicality and non-factivity. *Language and Linguistics*. 92: 79-106. (written in Korean)
- 2021 **Song, Sanghoun** & Eunjeong Oh. A multimethod approach to Korean speakers' acquisition of unaccusativity in English. *Journal of Cognitive Science*. 22(1): 135-186. (SCOPUS)
- 2021 Shin, Unsub, Myung-Kwan Park, and **Sanghoun Song**. Is c-command machinelearnable? *The Linguistic Association of Korea Journal*. 29(1): 183-204.
- 2021 Wang, Guehyun and **Sanghoun Song**. Evaluation of Neural Machine Translation Trained by Augmented Data Using English Resource Grammar. *Language and Information Society*. 42: 179-200. (written in Korean)
- 2021 Lee, Gyu-min, Seongtae Kim, Hyunsoo Kim, Kwonsik Park, Un-sub Shin, Guehyun Wang, Myung-Kwan Park, & **Sanghoun Song**. DeepKLM — A Computational Language Model-based Library for Syntactic Experiments — *Language Facts and Perspectives*. 52: 265-306. (written in Korean)
- 2020 **Song, Sanghoun**, HeeCheong Chon, & Soo Bok Lee. Analyzing the Disfluency of Reading Tasks of Persons Who Stutter Based on Deep Learning and Word Embedding. *Communication Sciences & Disorders*. 25(3): 721-737. (SCOPUS)
- 2020 Chung, Woojin, Hyowon Cho, & **Sanghoun Song**. Dictionary-based MWE Level List and its Application. *Korean Journal of Linguistics*. 45(3): 671-701.
- 2020 Cheong, Yunam, Guehyun Wang, & **Sanghoun Song**. A Comparative Study of Lexical Meanings in Chosun-ilbo and Rodong-shinmun. *Korean Semantics*. 69: 253-281. (written in Korean)
- 2020 Cheong, Yunam, Guehyun Wang, & **Sanghoun Song**. A Deep Learning-based Analysis of Ideological Words in Rodong Sinmun. *Korean Linguistics*. 88: 213-245. (written in Korean)
- 2020 Park, Kwonsik, Seok-Hoon You, & **Sanghoun Song**. Not Yet as Native as Native Speakers: Comparing Deep Learning Predictions and Human Judgments. *English Language and Linguistics*. 26(1): 199-228.
- 2020 Lee, Juwon & **Sanghoun Song**. Revisiting the persuade-constructions in Korean with empirical evidence. *Linguistic Research*. 37(1): 29-70. (SCOPUS)
- 2020 **Song, Sanghoun**. Several Misunderstandings about Machine Translation. *TransLinguaTech*. 1:5-23. (written in Korean)
- 2019 Park, Kwonsik, Seok-Hoon You, & **Sanghoun Song**. Using the Deep Learning Techniques for Understanding the nativelikeness of Korean EFL Learners. *Language Facts and Perspectives*. 48: 195-227. (written in Korean)
- 2019 **Song, Sanghoun**, Jae-Woong Choe, & Eunjeong Oh. An empirical study of honorific mismatches in Korean. *Language Sciences*. 75: 47-71. (A&HCI)
- 2018 Lee, Juwon, Eunjeong Oh, & **Sanghoun Song**. A Clausal Analysis of Resultative Constructions with X-key: Evidence from Corpus and Experiments. *Language & Information Society*. 35: 207-244.
- 2018 Lee, Juwon, Eunjeong Oh, & **Sanghoun Song**. An Experimental Study of Resultative Constructions with X-key. *Studies in Linguistics*. 49: 253-278.
- 2018 Lauzon, Kory, Eunjeong Oh, & **Sanghoun Song**. How Korean EFL Learners Comprehend Negative Frequency Adverbs. *Language Facts and Perspectives*. 43: 119-1.
- 2018 **Song, Sanghoun**. Phraseological Analysis of Learner Corpus Based on Language

- Model. *Language and Information*. 22(1): 123-152.
- 2017 **Song, Sanghoun** & Eunjeong Oh. What do you mean by contrast in syntax?. *Linguistic Research*. 34(3): 387-426.
- 2017 **Song, Sanghoun**, Sang-Geun Lee, Jae-Woong Choe, & Eunjeong Oh. Measuring the Degree of Convergence between Informal and Formal Acceptability Judgments for Korean. *Korean Journal of Linguistics*. 42(4): 761-804.
- 2017 **Song, Sanghoun** & Eunjeong Oh. In Defense of Comparisons between Formal and Informal Acceptability Judgments. *Studies in Generative Grammar*. 27(4): 893-902.
- 2017 **Song, Sanghoun**, Duk-Ho Jung & Eunjeong Oh. Acceptability Judgments of Double Case-Marking Constructions in Korean. *Language and Information*. 21(2): 79-113.
- 2017 **Song, Sanghoun**. A Corpus Study of Unbound Reflexive Pronouns in English. *Korean Journal of English Language and Linguistics*. 17(2): 275-305.
- 2017 Lee, Sang-Geun & **Sanghoun Song**. An Experimental Study on Korean EFL Students' Acceptability Judgment on Korean Sentences. *Language Information*. 24: 29-50. (written in Korean)
- 2017 Choe, Jae-Woong, Eunjeong Oh, & **Sanghoun Song**. Do Background Variables of Subjects Affect Acceptability Judgements? *Language and Linguistics*. 74: 181-206. (written in Korean)
- 2016 Lauzon, Kory & **Sanghoun Song**. Korean EFL Learners Can Hardly Understand *hardly*: A Corpus-based Study. *Korean Journal of English Language and Linguistics*. 16(4): 745-771.
- 2016 **Song, Sanghoun** & Eunjeong Oh. 5 or 7: Is the Choice So Important in Acceptability Judgment Testing? *Korean Journal of Linguistics*. 41(3): 449-480.
- 2016 Oh, Eunjeong & **Sanghoun Song**. An Experimental Study of Yeo (2006, 2011): -*Key/-Tolok* Alternation of Resultative Constructions in Korean. *Language Information*. 23: 25-54.
- 2016 **Song, Sanghoun**. A Multilingual Grammar Model of Honorification: Using the HPSG and MRS Formalism. *Language and Information*. 20(1): 25-49.
- 2015 **Song, Sanghoun**, Sang-Geun Lee, Jae-Woong Choe, and Eunjeong Oh. Different Ways to Test Acceptability Judgments. *Studies in Generative Grammar*. 25(2): 447-471.
- 2015 **Song, Sanghoun**. Automatic Acquisition of "noun+verb" Idiomatic Compounds in Korean. *Linguistic Research*. 32(1): 253-280.
- 2015 **Song, Sanghoun**. A Parallel Corpus-based Study of the Past Tense Morpheme in Korean. *Language Information*. 20: 75-104.
- 2014 **Song, Sanghoun**. Information Structure of Relative Clauses in English: a Flexible and Computationally Tractable Model. *Language and Information*. 18(2):1-29.
- 2014 **Song, Sanghoun** & Ji-Young Song. A Data Compilation of Multiple Case-marking Constructions: using the *Sejong* Spoken Corpus. *Language Information*. 19:57-90. (written in Korean)
- 2014 **Song, Sanghoun**. Subject-drop vs. Topic-drop: Evidence from a Multilingual Text. *Language and Linguistics*. 64:149-182.
- 2014 **Song, Sanghoun**, Jae-Woong Choe, & Eunjeong Oh. FAQ: Do Non-linguists Share the Same Intuition as Linguists?. *Language Research*. 50(2):357-386.
- 2014 **Song, Sanghoun** & Jae-Woong Choe. Selectional Preferences of Korean Verbal Items. *Linguistic Research*. 31(2):249-273.

- 2014 **Song, Sanghoun**. On Subjunctives in Korean: Exploiting a Bilingual Corpus. *Language and Information*. 18(1):1-32.
- 2012 **Song, Sanghoun** & Jae-Woong Choe. Automatic Construction of Verbal Type Hierarchy for Korean. *Language and Linguistics*. 54:201-238. (written in Korean)
- 2011 Kim, Jong-Bok, Jaehyung Yang, **Sanghoun Song**, and Francis Bond. Deep Processing of Korean and the Development of the Korean Resource Grammar. *Linguistic Research* (Center for the Study of Language, Kyung Hee Univ.). 28(3):635-672.
- 2010 **Song, Sanghoun** & Jae-Woong Choe. The Strength of the Relationship between Semantic Similarity and the Subcategorization Frames of the English Verbs: a Stochastic Test based on the ICE-GB and WordNet. *Language and Information*. 14(1):113-144 (written in Korean)
- 2010 Kim, Jong-Bok, Jaehyung Yang, & **Sanghoun Song**. Parsing Korean Comparative Constructions in a Typed-Feature Structure Grammar. *Language and Information*. 14(1):1-24
- 2010 **Song, Sanghoun**. Pragmatic Usage of Wh-elements in Korean. *Language and Information*. 11:91-113. (written in Korean)
- 2009 **Song, Sanghoun**. NARA ver. 2.0: Search System for the Sejong Korean-Japanese Bilingual Corpus. *The Journal of the Japanese Language Association of Korea*. 26:133-146. (written in Korean and Japanese)
- 2009 **Song, Sanghoun**. HPSG-based Generation for Korean Sentences and STYLE Features. *Linguistic Research* (Center for the Study of Language, Kyung Hee Univ.). 26(3):67-90. (written in Korean)
- 2009 **Song, Sanghoun** & Francis Bond. Checking the Feasibility of the Sejong Bilingual Corpus for Statistical Machine Translation. *Language and Linguistics*. 46:53-84.
- 2008 Choe, Jae-Woong, **Sanghoun Song**, & Jieun Jeon. Probabilistic Context-Free Grammar Rules based on Sejong Korean Treebank. *Language and Information*. 9:87-139. (written in Korean)
- 2008 **Song, Sanghoun**, Jieun Jeon, & Jae-Woong Choe. Semantic Restriction for the 'Adjective + Prepositional Phrase' Construction in English: a Statistical Test Using ICE-GB and WordNet. *Language and Linguistics*. 41:75-103. (written in Korean)

PEER-REVIEWED CONFERENCE PROCEEDINGS:

- 2021 Lee, Gyu-Min & **Sanghoun Song**. *HPSG/MRS-Based Sentence Generation with Transformer*. The 35th Pacific Asia Conference on Language, Information and Computation. Shanghai International Studies University (online).
- 2021 Kang, Arum & **Sanghoun Song**. *A study on subjunctive mood in Korean: Using corpus and experimental linguistic data*. The 23rd Seoul International Conference on Generative Grammar. Seoul (online).
- 2021 Lee, Gyu-Min & **Sanghoun Song**. *HPSG/MRS-Based Sentence Generation with Transformer*. The 28th International Conference on Head-Driven Phrase Structure Grammar. Goethe-Universität, Frankfurt am Main (online).
- 2020 Park, Kwonsik & **Sanghoun Song**. *Plausibility and Well-formedness Acceptability Test on Deep Neural Nativeness Classification*. The 34th Pacific Asia Conference on Language, Information and Computation (PACLIC 34). VNU University of Science (online).
- 2019 Lee, Juwon & **Sanghoun Song**. *The persuade-construction in Korean controls nothing*.

- The 33rd Pacific Asia Conference on Language, Information and Computation (PACLIC 33). Hakodate, Japan.
- 2015 Moeljadi, David, Francis Bond, and **Sanghoun Song**. *Building an HPSG-based Indonesian Resource Grammar (INDRA)*. Grammar Engineering Across Frameworks 2015 (in conjunction with ACL 2015).
- 2015 Fan, Zhenzhen, **Sanghoun Song**, and Francis Bond. *An HPSG-based Meta-Grammar for Chinese: ZHONG [|]*. Grammar Engineering Across Frameworks 2015 (in conjunction with ACL 2015).
- 2015 **Song, Sanghoun**. *Representing Honorifics via Individual Constraints*. Grammar Engineering Across Frameworks 2015 (in conjunction with ACL 2015).
- 2015 **Song, Sanghoun**, Bo Chen, Joanna Ut Seong Sio and Francis Bond. *An HPSG-based Analysis of Resultative Compounds in Chinese*. The 22nd International Conference on Head-Driven Phrase Structure Grammar.
- 2015 Wang, Wenjie, **Sanghoun Song**, and Francis Bond. *A Constraint-based Analysis of A-NOT-A Questions in Mandarin Chinese*. The 22nd International Conference on Head-Driven Phrase Structure Grammar.
- 2015 Sio, Joanna Ut Seong and **Sanghoun Song**. *Divergence of Expressing Definiteness between Mandarin and Cantonese*. The 22nd International Conference on Head-Driven Phrase Structure Grammar.
- 2015 Fan, Zhenzhen, **Sanghoun Song**, and Francis Bond. *Building Zhong [|], a Chinese HPSG Meta-Grammar*. The 22nd International Conference on Head-Driven Phrase Structure Grammar.
- 2012 **Song, Sanghoun** & Jae-Woong Choe. *Calculating Selectional Preferences of Transitive Verbs in Korean*. The 26th Pacific Asia Conference on Language, Information and Computation (PACLIC 26). Bali, Indonesia.
- 2012 **Song, Sanghoun** & Emily M. Bender. *Individual Constraints for Information Structure*. The 19th International Conference on Head-Driven Phrase Structure Grammar (HPSG 2012). Daejeon, Korea.
- 2011 **Song, Sanghoun** & Emily M. Bender. *Using Information Structure to Improve Transfer-based MT*. Workshop on Information Structure and Formal Grammar at the 18th International Conference on Head-Driven Phrase Structure Grammar (HPSG 2011). University of Washington. Seattle, USA.
- 2010 **Song, Sanghoun**, Jong-Bok Kim, Francis Bond, & Jaehyung Yang. *Development of the Korean Resource Grammar: Towards Grammar Customization*. The 8th Workshop on Asian Language Resources (in conjunction with COLING2010). Beijing, China.
- 2010 Kim, Jong-Bok, Jaehyung Yang, & **Song, Sanghoun**. *Korean Comparative Constructions: A Constraint-Based Approach and Computational Implementation*. The 17th International Conference on Head-Driven Phrase Structure Grammar (HPSG 2010). Paris Diderot, France.
- 2009 **Song, Sanghoun** & Francis Bond. *Online Search Interface for the Sejong Korean-Japanese Bilingual Corpus and Auto-interpolation of Phrase Alignment*. The 3rd Linguistic Annotation Workshop (LAW III) – ACL-IJCNLP. Suntec, Singapore.
- 2008 **Song, Sanghoun** & Jae-Woong Choe. *The Relationship between Semantic Similarity and Subcategorization Frames in English: A Stochastic Test Using ICE-GB and WordNet*. The 22nd Pacific Asia Conference on Language, Information and Computation (PACLIC 22). De La Salle University-Manila. Cebu City, Philippines.
- 2008 **Song, Sanghoun** & Jae-Woong Choe. *Automatic Construction of Korean Verbal Type Hierarchy using Treebank*. The 15th International Conference on Head-Driven

- Phrase Structure Grammar (HPSG 2008). ATR. Keihanna, Japan.
- 2008 **Song, Sanghoun** & Jieun Jeon. *The Xavier Module – Information Processing of Treebanks*. International Conference of Cognitive Science (ICCS 2008). Yonsei University. Seoul, Korea.
- 2007 **Song, Sanghoun**. *Constraints and Type for Hierarchies Korean Serial Verb Constructions*. The 21st Pacific Asia Conference on Language, Information and Computation (PACLIC 21). Seoul National University. Seoul, Korea.
- 2007 **Song, Sanghoun** & Jae-Woong Choe. *Type Hierarchies for Passive Forms in Korean*. The 14th International Conference on Head-Driven Phrase Structure Grammar (HPSG07). Stanford University. Stanford, USA.

GRANTS:

- 2023 Mid-Career Researcher Program
(The National Research Foundation of Korea)
'A Linguistic Re-evaluation of the Korean Language Models'
KRW 40,000,000 (2 years)
PI: Sanghoun Song
- 2022 'Korean ROCStories Dataset' (National Institute of Korean Language)
KRW 91,400,400 (7 months)
PI: Sanghoun Song
- 2022 Research Grant on Interdisciplinary Studies (Korea University)
'Development and Application of the Medical AI in Korean'
KRW 5,000,000 (1 year)
PI: Sanghoun Song
- 2021 'Korean Commitmentbank' (National Institute of Korean Language)
KRW 121,800,000 (7 months)
PI: Sanghoun Song
- 2021 Young Researcher Program (The National Research Foundation of Korea)
'Computational Linguistic Analysis and Implementation toward the
Coronavirus Crisis'
KRW 40,000,000 (2 years)
PI: Sanghoun Song
- 2021 Research Grant (Korea University)
'A Data Analysis of Relative Clauses'
KRW 5,000,000 (1 year)
PI: Sanghoun Song
- 2021 Research Grant on Interdisciplinary Studies (Korea University)
'An Interdisciplinary Study for Analyzing the North Korean Data'
KRW 25,000,000 (1 year)
PI: Seok-Hoon You
- 2020 Research Grant on Interdisciplinary Studies (Korea University)
'A Deep Learning-based Analysis of Speech-Pathologic Data'
KRW 2,500,000 (1 year)
PI: Sanghoun Song
- 2020~
(7 years) The 4th Brain Korea 21 (The National Research Foundation of Korea)
'Flagship Deep Learning Program for Language Scientists'
KRW 60,740,000 (6 months)
PI: Seok-Hoon You

- 2020 Young Researcher Program (The National Research Foundation of Korea)
‘An Empirical Base for the Study of Internally Headed Relative Clauses in Korean’
KRW 20,000,000 (1 year)
PI: Sanghoun Song
- 2020 KU Future Research Grant (Korea University)
‘Comparing the South and North Korean Languages with Deep Learning’
KRW 7,000,000 (1 year)
PI: Sanghoun Song
- 2019 Research Grant on Interdisciplinary Studies (Korea University)
‘An Analysis of *Rodong Paper* Using the NLP Techniques’
KRW 2,500,000 (1 year)
PI: Sanghoun Song
- 2019 Young Researcher Program (The National Research Foundation of Korea)
‘An Empirical Approach to Unbound Reflexives in English’
KRW 20,000,000 (1 year)
PI: Sanghoun Song
- 2019 Start-up Grant (Korea University)
‘A Corpus-based Study of Possessive Relative Clauses in English’
KRW 10,000,000 (1 year)
PI: Sanghoun Song
- 2017 Research Grant (Incheon National University)
‘A Corpus Study of Possessive Relative Clauses in English’
KRW 5,000,000 (1 year)
PI: Sanghoun Song
- 2016 Young Researcher Program (The National Research Foundation of Korea)
‘A Grammar Engineering-based Analysis of Korean Learner Corpora of English’
KRW 20,000,000 (2 years)
PI: Sanghoun Song
- 2016 Research Grant (Incheon National University)
‘A Corpus Study of Unbound Reflexives in English’
KRW 5,000,000 (1 year)
PI: Sanghoun Song
- 2015 Basic Research Lab Program (The National Research Foundation of Korea)
‘Modeling an Optimal Research Program for Experimental Syntax in Korean’
KRW 278,600,000 (3 years)
PI: Eunjeong Oh
- 2015 Start-up Grant (Incheon National University)
‘Methodology of Testing Acceptability Judgments’
KRW 12,000,000 (7 months)
PI: Sanghoun Song

PROJECTS:

- Jun. 2018-Dec. 2018 Building the Korean National Corpus
A Research of How to Use the Korean National Corpus
[Korean National Academy]

May. 2016-Apr. 2018	A Grammar Engineering-based Analysis of Korean Learner Corpora of English [National Research Foundation of Korea]
Nov. 2015-Oct. 2018	Modeling an Optimal Research Program for Experimental Syntax in Korean [National Research Foundation of Korea]
Aug. 2014-Aug. 2015	<i>That's what you meant.</i> A Rich Representation for Representing Meaning across Languages [Ministry of Education, Singapore Government]
Jan. 2010-Feb. 2014	The LinGO Grammar Matrix [National Science Foundation]
May. 2012-Apr. 2013	Selectional Preferences of Transitive Verbs in Korean [National Research Foundation of Korea]
May. 2011-Sep. 2011	Development of Korean MT Models with Deep Processing [National Research Foundation of Korea]
May. 2010-Apr. 2011	Automatic Construction of Korean Verbal Type Hierarchy using Treebank [National Research Foundation of Korea]
Nov. 2008-Oct. 2009	The Relationship between Semantic Similarity and Subcategorization Frames in English: A Stochastic Test Using ICE-GB and WordNet [Korean Research Foundation]
Mar. 2006-Feb. 2009	The 2 nd Brain Korea 21 Project – KLing Forum Foundational Studies for Applications of Linguistic Theories [Korean Research Foundation]
Mar. 2001-Aug. 2002	The Project for the Globalization of Korean, The Portal Web-Site Building [The Ministry of Culture and Tourism, Rep. of Korea]

AWARDS, HONORS, AND FELLOWSHIPS:

2023	The Granite Research Award
2022	The Granite Research Award The Granite Lecture Award (Spring)
2021	The Granite Research Award Lecture Award (Spring, Autumn)
2020	Lecture Award (Spring, Autumn)
2019	The Granite Lecture Award (Autumn)
2013	Graduate School Dissertation Fellowship (from the Graduate School, Univ. of Washington)
2009-2013	Top Scholar Award (from the Univ. of Washington)
2009	The Global Internship Fellow (from the Korean Research Foundation)
2008-2009	The Doctoral Scholarship (from the Seoul Metropolitan Government)
2007-2009	The 2nd Brain Korea 21 Project Scholarship

TEACHING EXPERIENCE:

Lecturer,

Dept. of Linguistics, Korea Univ.

Linguistics of Big Data (graduate)

SPR 2024

Linguistics of Artificial Intelligence (graduate)

AUT 2023

Theory of Grammar (undergraduate)

SPR 2023

Quantitative Foundation for Linguistics (graduate)

AUT 2022

Language and Artificial Intelligence (undergraduate)

AUT 2022

Seminar in Natural Language Processing (graduate)

SPR 2022

Ethics in Artificial Intelligence Society (undergraduate)

SPR 2022

Language and Statistics (undergraduate)

SPR 2022

Syntactic Theory (graduate)

AUT 2021

Language and Computer II (undergraduate)

SPR 2021

Computational Syntax (graduate)

SPR 2021

Language and Computer I (undergraduate)

AUT 2020, SPR 2024

Medical NLP (graduate)

AUT 2020

Syntax (undergraduate)

SPR 2020

Natural Language Processing (graduate)

SPR 2020, SPR 2023

Generative Grammar (undergraduate)

AUT 2019, AUT 2021, AUT 2023

Experimental Syntax (graduate)

AUT 2019

Deep Learning for Linguistic Studies (under/graduate)

SPR 2019

Grammar Engineering (graduate school)

SPR 2018

Linguistics 211: **Using Computer as a Tool**

SPR 2014 (teaching programming skills for linguists)

Linguistics 823: **Special Topics in**

Linguistic Information Processing

SPR 2014 (graduate, co-teaching with Prof. Jae-Woong Choe)

Dept. of English Language and Literature, Incheon National Univ.

Advanced English Grammar

SPR 2018
Grammar Engineering
SPR 2018
Computational Linguistics
SPR 2016, SPR 2017
Corpus Linguistics
SPR 2016
Programming Skills for Linguists
AUT 2015, AUT 2016, AUT 2017
Statistics for Linguists
AUT 2016, AUT 2017
Comparative Linguistics
SPR 2017
Practical English
AUT 2015, SPR/AUT 2016, SPR 2017, AUT 2017

School of English, Kyung Hee Univ.

English Syntax (teaching in English)
SPR 2014 (HPSG, co-teaching with Prof. Jong-Bok Kim)

Comparative Linguistics between English and Korean
(teaching in English)
SPR 2014 (co-teaching with Prof. Jong-Bok Kim)

Division of English Language & Literature, Dongguk Univ.

Applied Linguistics (graduate school)
SPR 2017, AUT 2017

Teaching Assistant,

Dept. of Linguistics, Univ. of Washington

Linguistics 472: **Introduction to Computational Linguistics**
SPR 2013

Linguistics 570: **Shallow Processing Techniques for NLP**
AUT 2012, AUT 2011, AUT 2010

Linguistics 571: **Deep Processing Techniques for NLP**
WIN 2013, WIN 2012, WIN 2011

STUDENTS:

MA (completed)

Kang San Noh (Feb. 2023): Korean EFL Learners' Underuse of the English Verb-Particle Constructions: A Collostructional Analysis

Junghyun Lee (Feb. 2023): Exploring the language use of newspapers on Covid-19: Focusing on four major newspapers

Hayul Park (Feb. 2023)

Gyu-min Lee (Feb. 2022): HPSG/MRS-based natural language generation using transformer

Un-sub Shin (Feb. 2022): Investigating neural network's sensitivity to negative polarity items

Seongtae Kim (Feb. 2022)

Guehyun Wang (Aug. 2021): Paraphrasing training data for neural machine translation

Kwonsik Park (Feb. 2020): A comparison of nativeness judgments by the deep learning system and humans on Korean EFL learner sentences

PhD (on-going) Kwonsik Park, Kang San Noh

MA (on-going) Yebin Lee, Subin Lee

Alumni Un-sub Shin (UCSD)

SERVICE:

Standing Committee, The International Conference on HPSG, 2017-2023

Board Member, The Linguistic Society of Korea
Korea Association for Corpus Linguistics (Vice)
The English Linguistics Society of Korea
The Korean Association for the Study of English Language & Linguistics
The Korean Generative Grammar Circle
The Discourse and Cognitive Linguistics Society of Korea

Conference Committee, ICKL 2015-2017, 2017-2019
PACLIC 30 (Organization Committee)
PACLIC 34 (Program co-chair)

Editor, UWWPL (University of Washington Working Papers in Linguistics)
[2009-2014]

Program Committee, HPSG 2015, 2017, 2018
GEAF 2015 (Grammar Engineering Across Frameworks)

Local Committee, HPSG 2015

Secretary, The 2nd Brain Korean 21 Project – KLing Forum
[2006-2009]
The Discourse and Cognitive Linguistics Society of Korea
[2001-2002]

WORK EXPERIENCE:

Internship, Swype, Inc. (now, NUANCE), Seattle
[Aug. 2010-Sep. 2010]

NICT, Japan
(National Institute of Information and Communications Technology)
Knowledge Creating Communication Research Center
Language Infrastructure Group
Supervisor: Prof. Francis Bond
[Feb. 2009-Aug. 2009]

Researcher, Eoneo Inc., *Language Technology Total Solution*
Natural Language Processing Team
[2002-2005]

Research Assistant, LinGO Grammar Matrix
[Mar. 2012-Jun. 2012, Mar. 2011-Jun. 2011]

Dept. of Linguistics, Univ. of Washington
[Dec. 2009-Jun. 2010]

DEVELOPMENT EXPERIENCE:

Computational Grammars:

- Resource Grammars: Korean, Chinese, Japanese, English, Indonesian
- Experimental Grammars: Vietnamese, Russian, Spanish

Language Database: Korean, Korean-Japanese, Korean-English

Stemmer: English (C++, Java, Python)

POS Taggers: Korean (C++), English (Python)

Syntactic Parsers: Korean (Java), English (Python, C++)

Machine Translators: English / Korean / Japanese / Chinese

NLP Engines: English Grapheme To Phoneme (G2P) Transducer, *_isA_* Recognizer, Korean Expression Searching Engine, English Spell Checker, Sentence Splitter etc.

Interfaces: NARA (<http://corpus.mireene.com/nara.php>), English Electronic Dictionary

PROGRAMMING SKILLS:

Languages: Linux Shell Script, C/C++/C#, Python, Perl, Java

Web Programming: XML/XSLT, PHP, Javascript, CSS

DBMS: MySQL, sqlLite, MS-SQL, mdb

etc.: LaTeX, R, Tensorflow, PyTorch

LANGUAGES:

Korean: Native

English: Fluent

Mandarin Chinese, Japanese: Intermediate

CONTACT:

Department of Linguistics
Korea University
[02841] 145 Anam-ro, Seongbuk-gu, Seoul, Korea
Tel: +82-2-3290-2177 (office)
E-mail: sanghoun@korea.ac.kr