Seongmin Mun

Email: simon.seongmin.mun@gmail.com
Homepage: http://seongminmun.com/

RESEARCH INTERESTS

Main Area Natural Language Processing; Machine learning; Deep learning; Data visualization; Statistics; Korean

Current Focus Word sense disambiguation; Distributional semantic models; Classification/visualization system development

EDUCATION

Université Paris Nanterre

Ph.D. in *Sciences du Langage*, July 2021

Dissertation: *Polysemy resolution with word embedding models and data visualization: the case of adverbial postpositions -ey, -eyse, -(u)lo in Korean* (Chairperson: Dr. Guillaume Desagulier)

Ajou University

M.S. in *Media Contents*, February 2016

Pyeongtaek University

B.S. in *Statistics*, February 2014

SKILLS

Programing Language:

Python, Java, JavaScript, SQL, PHP, HTML/CSS

Web Application:

Bitnami, Flask, Node.js, RShiny, Spring

Natural Language Processing:

nltk, spaCy, KoNLPy, stringr, tm, KoNLP, CoreNLP, KKMA, Eunjeon

Machine (Deep) Learning:

PyTorch, Transformers, TensorFlow, Keras, Gensim, scikit-learn, SciPy, MASS, lmtest

Data Visualization:

D3.js, Google Charts, Chart.js, matplotlib, ggplot2

Statistics Software:

R (RStudio), SAS, SPSS, MATLAB, Stata

PAPERS AND PROCEEDINGS

<u>**Major**</u>

Shin, G-H., & Mun, S. (2021). Bayesian simulation of clause-level constructional knowledge in child language development: Active transitives and suffixal passives in Korean. In D. Dionne and L-A. Vidal Covas (Eds.), *Proceedings of the 45th Annual Boston University Conference on Language Development* (pp. 679–692). Boston, MA: Cascadilla Press.

Ha, H., Han, H., Mun, S., Bae, S., Lee, J., & Lee, K. (2019). An improved study of multilevel semantic network visualization for analyzing sentiment word of movie review data. *Applied Sciences*, *9*, 2419. doi:10.3390/app9122419

Mun, S., Choi, G., Desagulier, G., & Lee, K. (2018). PreechVis: Visual profiling using multiple-word combinations. *Proceedings of the 13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (IVAPP2018)*, 97-107.

Mun, S., Desagulier, G., & Lee, K. (2017). A visual approach for text analysis using multiword topics. *IEEE, EuroVis 2017, 19th EG/VGTC Conference on Visualization*, 57-59.

Mun, S., & Lee, K. (2016). Data analysis by integrating statistics and visualization: visual verification for the prediction model. *Journal of Society of Design Convergence*, 15(6), 195-214.

<u>**Others**</u>

Mun, S. (invited; revised and resubmitted). *Towards a better annotation system for Korean adverbial postpositions: Challenges for development research*.

Jung, B. K., Shin, G-H., & Mun, S. (under review). *What Natural Language Processing techniques tell us about text quality of learner writing: Relationship between similarity scores and proficiency*.

Shin G-H., & Mun, S. (under review). *Korean-speaking children's constructional knowledge about a transitive event: Corpus analysis and Bayesian modelling*.

Ha, H., Mun, S., Choi, G., Hong, E., Lee, S., & Lee, K. (2019). Visual analysis on the political orientation of historical characters in the Joseon Dynasty: Focusing on Seo, Geojeong. *Archives of Design Research*, 32(1), 147-161.

Lee, K., Lee, J., Kim, D., Park, J., Mun, S., Jang, Y., & Park, J. (2017). Controversy visualization- how controversial public discourse in wikipedia articles evolves over time. *Archives of Design Research*, 30(4), 57-69.

Mun, S., Choi, G., Lee, S., & Lee, K. (2017). Visual analysis for voting relationships in Joseon Dynasty, Korea. *Proceedings of the 10th International Conference on ICT, Society and Human Beings*, 111-118.

Park, S., Jo, E., Mun, S., & Lee, K. (2016). Analysis of corporate twitter usage types and Influence visualization on key players. *Journal of Society of Design Convergence*, 15(6), 129-144.

Mun, S., Kim, G., Choi, G., & Lee, K. (2016). Visual trend analysis method for ontology-based opinion mining on movie review. *Proceedings of the 13th International Conference on Web Based Communities and Social Media 2016*, 287-290.

Ha, H., Han, H., Mun, S., Bae, S., Lee, J., & Lee, K. (2016). Visualization of movie recommendation system using the sentimental vocabulary distribution map. *The Korea Society of Computer and Information*, 21(5), 19-29.

Mun, S., Kim, G., Choi, G., & Lee, K. (2016). Movie Recommended system based on analysis for the user review utilizing ontology visualization. *Journal of Society of Design Convergence*, 15(2), 347-368.

Park, S., Jo, E., Lee, H., Mun, S., & Lee, K. (2015). Analysis of corporate twitter usage types and influence visualization on key players. *Proceedings of the IADIS International Conference Web Based Communities*, 122-131.

Ha, H., Mun, S., Kwon, O., & Lee, K. (2015). Proximity based circular visualization for similarity analysis of UNGA voting patterns. *Proceedings of the IEEE International Symposium on Big Data Visual Analytics (BDVA 2015)*, 127-128.

Choi, H. Ha, H., Mun, S., & Lee, K. (2015). Proximity based circular visualization for similarity analysis of voting patterns between nations in UN General Assembly. *Journal of Society of Design Convergence*, 14(4), 133-150.

Mun, S., Ha, H., & Lee, K. (2015). Relationship analysis between the box office performance and sentimental words in movie review. *Journal of Society of Design Convergence*, 14(4), 1-16.

Mun, S., Han, H., Ha, H., & Lee, K. (2014). A visual analysis on factors affecting repurchase intention in social commerce. *Journal of Society of Design Convergence*, 13(6), 139-152.

PRESENTATIONS

Major

Shin, G-H., & Mun, S. (2021, October). Korean-speaking children's constructional knowledge about transitive events: Corpus analysis and Bayesian modelling. Paper presenting at the 30th Association for Korean Studies in Europe, La Rochelle University, France.

Mun, S., & Shin, G-H. (2021, August). *How does context window size address polysemy of adverbial postposition -(u)lo in Korean?* Paper presenting at the 11th International Conference on Construction Grammar (ICCG11).

Mun, S., & Shin, G-H. (2020, June). *Context window and polysemy interpretation: A case of Korean adverbial postposition -(u)lo*. Poster presented at the IMPRS Conference 2020: Interdisciplinary Approaches to the Language Sciences, Max Planck Institute for Psycholinguistics.

Guo, W., Kim, S., Mun, S., & Lee, K. (2018, October). *BubbleUp: Toward better analysis for temporal event data*. Poster presented at the IEEE VIS 2018, Scientific Visualization, Information Visualization and Visual Analytics.

Mun, S., Wang, I., Desagulier, G., Choi, G., & Lee, K. (2018, July). *Time flies like an arrow and fruit flies like a banana; parsing multiword constructions with DepVis*. Paper presented at the 10th International Conference on Construction Grammar (ICCG10), Paris Sorbonne University.

Others

Jung, B. K., Shin, G-H., & Mun, S. (2021, October). Automatic text quality analysis of learner writing: Parallel comparisons of NLP techniques for text similarity measurement. Paper presenting at the 30th Association for Korean Studies in Europe, La Rochelle University, France.

Jung, B. K., Shin, G-H., & Mun, S. (2021, August). NLP-based measurement of text quality for learner writing: Relationship between text similarity and proficiency. Paper presenting at the 22nd Biennial Meeting of the International Circle of Korean Linguistics (ICKL22).

Shin, G-H., & Mun, S. (2021, August). *Bayesian simulation of Korean-speaking children's constructional knowledge about transitive events*. Paper presenting at the 11th International Conference on Construction Grammar (ICCG11).

Shin, G-H., & Mun, S. (2020, November). *Bayesian simulation of clause-level constructional knowledge in child language development: Active transitives and suffixal passives in Korean*. Poster presented at the 45th Annual Boston University Conference on Language Development (BUCLD).

Shin, G-H., & Mun, S. (2020, June). *Bayesian modelling of clause-level constructional knowledge for Korean-speaking preschool children*. Poster presented at the IMPRS Conference 2020: Interdisciplinary Approaches to the Language Sciences, Max Planck Institute for Psycholinguistics.

Mun, S. (2019, November). *A computational approach to resolve the polysemy of postpositions in Korean*. Poster presented at the COLDOC 2019, Colloque de doctorants et de jeunes chercheurs en Sciences du langage.

Mun, S., Wang, I., Desagulier, G., Choi, G., & Lee, K. (2018, December). *Fruit flies like a banana: Parsing multiword constructions with DepVis*. Paper presented at the Digits, internal reflection, and the study of linguistics., Hankuk University of Foreign Studies.

Choi, G., Yoon, T., Park, C., Lee, S., Choi, H., Park, C., Kang, J., Mun, S., & Lee, K. (2018, December). *A visual analysis of powerrRelations among people on time-series data using network visualization: Focusing on 19th national assembly data in Korea*. Paper presented at the 43rd Annual Meeting of the Social Science History Association.

Desagulier, G., Lacheret-Dujour, A., Isel, F., & Mun, S. (2018, May). *Characterizing discourse genres with prosodic features in a reference treebank of spoken French*. Paper presented at the Aflicojet2018, Corpora and Representativeness.

Mun, S., Desagulier, G., & Lee, K. (2017, October). *How can we capture multiword expressions?* Poster presented at the 5th International Conference on Statistical Language and Speech Processing (SLSP2017).

Mun, S., Desagulier, G., Lacheret-Dujour, A., & Lee, K. (2017, June). *How can we capture multiword expressions?* Paper presented at the day of Ph.D. students in MODYCO.

Mun, S., Han, H., Choi, G., Yoon, T., Kim, J., & Lee, K. (2016, November). *Hierarchical network visualization approach to power mechanism in medieval Joseon Korea*. Paper presented at the 41st Annual Meeting of the Social Science History Association.

Bae, S., Mun, S., Choi, G., Lim, S., Bang, S., Son, S., Hong, C., Shin, H., & Lee, K. (2016, October). *Decision tree analysis utilizing sankey diagram*. Poster presented at the Academic Association of Global Cultural Contents (AAGCC).

Choi, S., Kim, J., Kim, S., Mun, S., Kim, K., & Lee, K. (2016, August). *Analysis of linguistic difference in writing primed with regulatory focus*. Poster presented at the 6th International Conference on Language, Literature, Culture and Education (ICLLCE) 2016.

Mun, S., Choi, G., Han, H., Kim, J., & Lee, K. (2016, June). *VoteStreamVis: Visual analysis of congressional votes in the annals of the Joseon Dynasty*. Paper presented at the Interdisciplinary Research Methodology for 'Big Data' in Historical Studies; The beginning of 'Digital History'.

Mun, S., Kim, G., Karim, R. M., & Lee, K. (2015, October). *Internet review opinion mining and data visualization*. Poster presented at the Symposium on Visualization in Data Science (VDS at IEEE VIS 2015).

Lee, S., Choi, H., Mun, S., Han, H., Ha, H., & Lee, K. (2014, December). *Visualization approach to a Korean genealogy data*. Paper presented at the 1st International Symposium on Digital Humanities, Digital Humanities and Culture: Trends and Prospects.

GRANTS, AWARDS, & SCHOLARSHIPS

<u>**Major**</u>

Research Grant (KRW 30,000,000; app. USD 27,000) | National Research Foundation of Korea (PI: Prof. Kyungwon Lee); February 2018 ~ February 2019

Research Scholar Grant (KRW 15,000,000; app. USD 13,500) | Ajou University; September 2017 ~ January 2018

Research Scholar Grant (KRW 15,000,000; app. USD 13,500) | Ajou University; February 2017 ~ August 2017

Research Scholar Grant (KRW 15,000,000; app. USD 13,500) | Ajou University; September 2016 ~ January 2017

Research Grant (KRW 60,000,000; app. USD 54,000) | Korea Meteorological Administration (PI: Prof. Kyungwon Lee); March 2015 ~ December 2015

Research Assistance Scholarship (KRW 24,500,000; app. USD 22,050) | Department of History, Ajou University; November 2014 ~ November 2018

Research Grant (KRW 1,000,000,000; app. USD 900,000) | National Research Foundation of Korea (PI: Prof. Sangkuk Lee); November 2014 ~ November 2024

Research Grant (KRW 10,800,000; app. USD 9,720) | National Research Foundation of Korea, Brain Korea 21, September 2014 ~ February 2016

Others

Research Assistance Scholarship (EUR 5,000; app. USD 6,000) | Department of Linguistics, Université Paris 8; February 2019 ~ August 2019

Research Assistance Scholarship (EUR 2,500; app. USD 3,000) | Department of Linguistics, Université Sorbonne Nouvelle Paris 3; November 2018 ~ February 2019

Best Paper Award | Journal of Design Research; 2018 KSDS Fall International Conference; November 2018

Research Assistance Scholarship (KRW 1,000,000; app. USD 900) | Ajou University; August 2018

Research Grant (KRW 2,000,000; app. USD 1,800) | National Research Foundation of Korea, Brain Korea 21; July 2018 ~August 2018

Research Assistance Scholarship (KRW 3,000,000; app. USD 2,700) | Department of History, Ajou University; June 2018 ~ September 2018

Research Award from the Minister of Education | The Education Ministry of Korea; National Research Foundation of Korea; March 2018

Ph.D. Research Scholarship (KRW 7,000,000; app. USD 6,300) | National Research Foundation of Korea, Brain Korea 21, February 2018 ~ August 2018

Research Assistance Scholarship (KRW 1,500,000; app. USD 1,350) | Department of History, Seongbuk-gu government public relations division; November 2017 ~ March 2018

Full Scholarship for Registration (KRW 5,000,000; app. USD 4,500) | Ajou University; September 2017

Full Scholarship for Registration (KRW 5,000,000; app. USD 4,500) | Ajou University; February 2017

Kantar Information is Beautiful Awards | Kantar; November 2016

Ajou Undergraduate Research Award (KRW 1,000,000; app. USD 900) | Ajou University; December 2016

Full Scholarship for Registration (KRW 5,000,000; app. USD 4,500) | Ajou University; September 2016

Research Assistance Scholarship (KRW 1,800,000; app. USD 1,620) | Leaders in INdustry-university Cooperation (LINC); March 2016 ~ August 2016

Ph.D. Research Scholarship (KRW 7,000,000; app. USD 6,300) | National Research Foundation of Korea, Brain Korea 21, February 2016 ~ August 2016

80% Scholarship for Registration (KRW 4,000,000; app. USD 3,600) | National Research Foundation of Korea, Brain Korea 21; February 2016

Half Scholarship for Registration (KRW 2,500,000; app. USD 2,250) | National Research Foundation of Korea, Brain Korea 21; September 2015

Research Assistance Scholarship (KRW 4,000,000; app. USD 3,600) | Korea Meteorological Administration; March 2015 ~ December 2015

Half Scholarship for Registration (KRW 2,500,000; app. USD 2,250) | National Research Foundation of Korea, Brain Korea 21; February 2015

Full Scholarship for Registration (KRW 5,000,000; app. USD 4,500) | Ajou University; September 2014

Special Admission Scholarship (KRW 5,000,000; app. USD 4,500) | Ajou University; February 2014

Half Scholarship for Registration (KRW 2,000,000; app. USD 1,800) | Pyeongtaek University; September 2013

Full Scholarship for Registration (KRW 4,000,000; app. USD 3,600) | Pyeongtaek University; February 2013

Full Scholarship for Registration (KRW 4,000,000; app. USD 3,600) | Pyeongtaek University; September 2012

TEACHING EXPERIENCE

Exploring History through Statistics | Korea University, October 2018 ~ November 2018
- Aim: To create a dataset on Korean history and analyze the data automatically
- Target: Graduate
- Language(s) / package(s) used: R / *MASS, ggplot2, scatterplot3d, lmtest, RCurl, stringr*
- Contents
  - Web crawling / POS tagging / regression analysis / cluster analysis

Recipes: Data Handling and Analysis with R | Ajou University, July 2018 ~ August 2018
- Aim: To introduce how to handle/analyze data through common data mining techniques
- Target: Undergraduate
- Language(s) / package(s) used: R / *KoNLP, stringr, RCurl, d3Network, wordcloud*
- Contents
  - N-gram extraction / word cloud / network visualization

Recipes: Text Mining with R | Ajou University, May 2018 ~ June 2018
- Aim: To introduce how to analyze language data in an R environment
- Target: Undergraduate
- Language(s) / package(s) used: R / *wordcloud, KoNLP, stringr, RCurl, rsconnect, shiny*
- Contents
  - POS tagging / dependency parsing / web crawling / word cloud

Text Mining for Historical Data | Ajou University, March 2018 ~ June 2018, March 2016 ~ June 2016, September 2015 ~ December 2015
- Aim: To introduce data mining techniques for historical data
- Target: Undergraduate
- Language(s) / package(s) used: R / *wordcloud, KoNLP, stringr, RCurl, rsconnect, shiny*
- Contents
  - Word cloud / POS tagging / dependency parsing / network visualization

Data Analysis through Visualization | Lecipes, January 2016 ~ March 2016

- Aim: To train people for acquiring commonly used data visualization techniques, in combination with statistical analysis, in order to apply this knowledge to their workplace
- Target: Industry workers
- Language(s) / package(s) used: R / *MASS, ggplot2, lmtest, RCurl, stringr, corrplot, hebin*
- Contents
    - Visualization: Hitmap, Hebin, MDS
    - Statistics: Correlation, Regression, Cluster

Visual Information Graphics | Ajou University, March 2015 ~ June 2015

- Aim: To develop web-based data visualization systems
- Target: Undergraduate
- Language(s) / package(s) used: Java, JavaScript, HTML/CSS, SQL / *D3.js, Jquery.js, KKMA, Eunjeon*
- Contents
    - Sunburst / Radial tree visualization
    - Sankey diagram / Parallel Coordinates

Visual Analysis of Historical Data | Ajou University, March 2015 ~ June 2015

- Aim: To introduce how to apply data mining through visualization to historical data
- Target: Undergraduate
- Language(s) / package(s) used: R / *MASS, ggplot2, KoNLP, stringr*
- Contents
    - Cluster analysis / MDS visualization / POS tagging

PERSONAL EXPERIENCE

Research Assistant (Palacký University Olomouc; February 2021 ~ June 2021)

> Developed and designed a web-based Korean test system for L2 learner

Research Assistant (Université Paris 8; March 2020 ~ November 2020)

> Developed COHA pre-trained models with RoBERTa & a data visualization system to evaluate the models

Research Assistant (Université Sorbonne Nouvelle Paris 3; November 2018 ~ March 2019)

> Developed a web-based data archiving system for efficient handling and safe management of child language development records

Research Assistant (Ajou University & Université Paris Nanterre; December 2017 ~ August 2018)

> Developed an automatic recognition system to extract multiword expressions from raw text data

Research Assistant (Ajou University Hospital; March 2016 ~ August 2016)

> Designed a data visualization system to investigate diagnosis records of dementia patients to better identify patterns involving the symptoms of dementia by patients for the enhanced treatment

Research Assistant (Ajou University, Leaders in INdustry-university Cooperation (LINC); March 2016 ~ August 2016)

> Data mining/analysis, system development, research paper writing as a co-author

Research Assistant (Sungshin Women's University; December 2015)

> Developed a data analysis program for the relationship between formative elements of arts and degree of art quality by using Classification and Regression Trees

Research Assistant (Korea Road Traffic Authority & Korea Meteorological Administration; March 2015 ~ December 2015)

> Developed a data mining process of accident reports and an algorithm for searching for better places for traffic cameras

Research Assistant (Ajou University, Department of History; November 2014 ~ November 2018)

> Developed data visualization systems and text mining algorithms to investigate power changing patterns of strengthening or weakening as the power of the family was inherited to later generations in the Joseon dynasty

Research Assistant (Ajou University, Brain Korea 21; September 2014 ~ August 2018)

> Data mining/analysis, system development, research paper writing as a co-author

Statistical investigation agent (Gyeongin Regional Statistics Office; July 2013 ~ August 2013)

> Managed survey data of the Population and Housing Census