

Hahn Koo

Department of Linguistics and Language Development
San Jose State University
One Washington Square
San Jose, CA, 95192-0093, U.S.A.

Phone: +1-408-924-7093
Fax: +1-408-924-4703
Email: hahn.koo@sjsu.edu
Web: <http://www.sjsu.edu/faculty/hahn.koo>

Research Interests

computational linguistics, corpus linguistics, phonetics, psycholinguistics

Education

University of Illinois at Urbana-Champaign, Urbana, IL

Ph.D. in Linguistics, December 2007

Committee: Richard Sproat, Jennifer Cole, Gary Dell, Cynthia Fisher

Thesis: Change in the adult phonological processing system by learning non-adjacent phonotactic constraints from brief experience: An experimental and computational study

Seoul National University, Seoul, Korea

B.A. in English Language and Literature, August 2002

Employment

San Jose State University, San Jose, CA

Associate Professor, Department of Linguistics and Language Development, August 2014 – Present

Assistant Professor, Department of Linguistics and Language Development, August 2008 – August 2014

Funding

Koo, H. (PI) “Annotation of a speech database collected from Chinese, Japanese, and Korean learners of English”, Naver Corporation, August 2015 – January 2017, **\$355,713**.

Koo, H. (PI) “Development of an American English multiple pronunciation lexicon”, Electronics and Telecommunications Research Institute, July 2011 – September 2011, **\$50,000**.

Koo, H. (PI) “Development of multilingual grapheme-to-phoneme conversion and part-of-speech tagging modules”, Electronics and Telecommunications Research Institute, July 2010 – January 2011, **\$100,000**.

Koo, H. (PI) “Performance enhancement of the multilingual (Canadian French, Mexican Spanish) grapheme-to-phoneme conversion modules”, Electronics and Telecommunications Research Institute, April 2009 – September 2009, **\$50,000**.

Koo, H. (PI) “Development of multilingual (Canadian French, Mexican Spanish) grapheme-to-phoneme conversion modules”, Electronics and Telecommunications Research Institute, September 2008 – January 2009, **\$50,000**.

Publications

Journal articles

- Koo, H., Kataoka, R., Swan, J. T., & Tzeng, C. Y. (2023). Effects of lexical frequency on the post-exposure magnitude of recalibration in lexically guided perceptual learning: An explorative analysis. *JASA Express Letters*, 3(8), 085201.
- Frazier, S., & Koo, H. (2019). The use of progressive aspect in discourse: an analysis of a corpus of interview data. *Corpus Pragmatics*, 3(2), 145-171.
- Frazier, S., & Koo, H. (2019). Discourse-pragmatic functions of the present perfect in American English TV and radio interviews. *Text & Talk*, 39(1), 77-98.
- Kataoka, R., & Koo, H. (2017). Comparing malleability of phonetic category between [i] and [u]. *The Journal of the Acoustical Society of America*, 142(1), EL 42-48.
- Koo, H. (2015). An unsupervised method for identifying loanwords in Korean. *Language Resources and Evaluation*, 49(2), 355-373.
- Koo, H., & Oh, Y.-i. (2013). Beyond tier-based bigrams: an artificial grammar learning study. *Language Sciences*, 38, 53-58.
- Koo, H., & Callahan, L. (2012). Tier-adjacency is not a necessary condition for learning phonotactic dependencies. *Language and Cognitive Processes*, 27(10), 1425-1432.
- Koo, H., & Oh, Y.-i. (2007). Onset-to-onset probability and gradient acceptability in Korean. *Language Research*, 43(2), 287-307.

Conference proceedings

- Koo, H. (2011). A weighted finite-state transducer implementation of phoneme rewrite rules for English-to-Korean pronunciation conversion. *Procedia – Social and Behavioral Sciences*, 27, 202-208.
- Koo, H. (2011). A statistical model of Korean loanword phonology. In Ho-min Sohn, et al. (Eds.): *Japanese/Korean Linguistics*, 19 (pp. 15-26), Stanford, CA: CSLI Publications.
- Koo, H. (2010). Perplexity of bi-phone phonotactic models in Korean loanword phonology. *Proceedings of the 2010 Seoul International Conference on Linguistics*.
- Koo, H., & Cole, J. (2009). Gradient perceptual facilitation from phonotactic knowledge. *Proceedings from the Annual Meeting of the Chicago Linguistic Society*, 43(1), 167-181.
- Koo, H., & Cheng, Y. M. (2006). Incremental learning of MAP context-dependent edit operations for spoken phone number recognition in an embedded platform. In *Proceedings of ISCA International Conference on Spoken Language Processing, INTERSPEECH 2006* (pp. 2310-2313).
- Koo, H., & Cole, J. (2006). On learnability and naturalness as constraints on phonological grammar. In Antonis Botinis (Ed.): *Proceedings of ISCA Tutorial and Research Workshop on Experimental Linguistics* (pp. 165-168). Athens: University of Athens.

Presentations

Refereed

- “A Praat script for unsupervised phoneme segmentation based on spectro-temporal representation”, Poster presented at the 189th Meeting of the Acoustical Society of America, Honolulu, U.S.A., December 2025.

“Overfitting in Deep Learning Models for Disfluency Detection in Bilingual Children”, (with Caryn Iwakiri, Pei-Tzu Tsai, and Janet Bang) Poster presented at the 2025 ASHA (American Speech-Language-Hearing Association) Convention, Washington D.C., U.S.A., November 2025 (**designated as an ASHA Convention Centennial Session**)

“Understanding AI model development and application in disfluency detection”, (with Caryn Iwakiri, Pei-Tzu Tsai, and Janet Bang) Poster presented at the 2025 ASHA (American Speech-Language-Hearing Association) Convention, Washington D.C., U.S.A., November 2025

“Exploring automatic fluency detection using bilingual children’s speech”, (with Caryn Iwakiri, Pei-Tzu Tsai, and Janet Bang) Poster presented at the 2024 ASHA (American Speech-Language-Hearing Association) Convention, Seattle, U.S.A., December 2024.

“A digital pattern playback system implemented in Python”, Poster presented at the 182nd Meeting of the Acoustical Society of America, Denver, U.S.A., May 2022.

“Change in pitch and voice quality of L2 speakers of English as a function of similarity between their L1 and English”, Poster presented at the 181st Meeting of the Acoustical Society of America, Seattle, U.S.A., November, 2021.

“Dynamic vowel qualities among multiethnic Santa Clara County talkers”, (with Julia Swan and Reiko Kataoka) Poster presented at the 95th Annual Meeting of the Linguistic Society of America, Virtual Meeting, January 2021.

“Voice quality differences between American English speakers and Korean learners of English”, Poster presented at the 179th Meeting of the Acoustical Society of America, Virtual Meeting, December, 2020.

“Effects of variability and internal structure on malleability of phonetic categories: A case of [i] versus [u]”, (with Reiko Kataoka) Poster presented at the 178th Meeting of the Acoustical Society of America, San Diego, U.S.A., December 2019.

“Is perceptual learning influenced by inherent variability of phonetic category?”, (with Reiko Kataoka) Poster presented at the 171st Meeting of the Acoustical Society of America, Salt Lake City, U.S.A., May 2016.

“A weighted finite-state transducer implementation of phoneme rewrite rules for English-to-Korean pronunciation conversion”, Poster presented at the 12th Conference of the Pacific Association for Computational Linguistics (PACLING2011), Kuala Lumpur, Malaysia, July 2011.

“Perplexity of bi-phone phonotactic models in Korean loanword phonology”, Talk presented at the 2010 Seoul International Conference on Linguistics, Seoul, Korea, June 2010.

“Statistical model of Korean loanword phonology”, Talk presented at the 19th Japanese/Korean Linguistics Conference, Honolulu, U.S.A., November 2009.

“Phonotactic learning beyond tier-adjacency”, Talk presented at the 83rd Annual Meeting of the Linguistic Society of America, San Francisco, U.S.A., January 2009.

“Effect of onset-to-onset probability on non-word speech processing”, (with Young-il Oh) Talk presented at the 82nd Annual Meeting of the Linguistic Society of America, Chicago, U.S.A., January 2008.

“Gradient perceptual facilitation from phonotactic knowledge”, (with Jennifer Cole) Talk presented at the 43rd Annual Meeting of the Chicago Linguistic Society, Chicago, U.S.A., May 2007.

“Complexity and perceptual factors in phonotactic learning: Evidence from artificial grammar learning”, (with Jennifer Cole) Talk presented at the 81st Annual Meeting of the Linguistic Society of America, Anaheim, U.S.A., January 2007.

“Incremental learning of MAP context-dependent edit operations for spoken phone number recognition in an embedded platform”, (with Yan Ming Cheng) Poster presented at INTERSPEECH 2006, Pittsburgh, U.S.A., September 2006.

“On learnability and naturalness as constraints on phonological grammar”, (with Jennifer Cole) Talk presented at the ISCA Tutorial and Research Workshop on Experimental Linguistics, Athens, Greece, August 2006.

Invited

“Unsupervised phoneme segmentation using a transformer autoencoder”, BayPhon2025, UC Santa Cruz, May 2025.

“Comparing malleability of phonetic category between [i] and [u]”, (with Reiko Kataoka) The Berkeley Phonetics and Phonology Forum, April 2018.

“Using sub-lexical statistics to find loanwords in unlabeled monolingual corpora: A case study in Korean”, Stanford Phonetics and Phonology Workshop, November 2015.

“Detecting and recovering transliterated foreign words”, eBay Language Technology Forum, March 2014.

“An unsupervised method for identifying foreign loanwords in Korean”, Seoul National University, July 2012.

“On contiguity and similarity as constraints on phonotactic learning”, UCLA Linguistics Colloquium, October 2011.

“Noisy channel model for error correction in speech and language technology”, San Jose State University, February 2008.

“Using an error corrective post-processor to improve speech recognition accuracy”, AT&T Labs-Research, May 2007.

“Confusability and perceptual facilitation from knowledge of phonotactic constraints”, San Jose State University, March 2007.

Teaching

Courses taught

San Jose State University (Fall 2008 – Present)

LING 20 Nature of Language

LING 26 Quantitative Reasoning in Linguistic Diversity

LING 26W Quantitative Reasoning in Linguistic Diversity Workshop

LING 101 Introduction to Linguistics

LING 111 Introduction to Linguistic Phonetics

LING 113 Introduction to Phonology

LING 115 Corpus Linguistics

LING 123 Sound and Communication

LING 124 Introduction to Speech Technology

LING 161 Psycholinguistics

LING 162 Introduction to Morphology

LING 165 Introduction to Natural Language Processing

LING 201 Phonology – Theory and Applications

Seoul National University (Summer 2012)

Seminar on Sublexical Co-occurrence Statistics – Implications and Applications

Theses supervised

As Committee Chair

Simon, Megan (2019). *A Phonetic Distance Approach to Intelligibility between Mam Regional Dialects* [Master's thesis, San Jose State University].

Tracy Rohlin (2016). *Popularity Prediction of Reddit Texts* [Master's thesis, San Jose State University].

Elena Grossfeld (2013). *Evaluating the Contribution of Hashtags to Sentiment Analysis of Microblogs* [Master's thesis, San Jose State University].

As Committee Member

Iwakiri, Caryn W. C. (2025). *Generalizability of Existing AI Models for Automatic Disfluency Detection in Children's Speech* [Master's thesis, San Jose State University].

Angier, Danielle Miles (2023). *Qualifying Quantifiers: A Usage-Based, Diachronic Analysis of Quantifier Constructions* [Master's thesis, San Jose State University].

Koffman, David (2020). *Easy in Any Language: Syntax, Semantics, and Pragmatics of Activity-Oriented Adjectives in Six Languages* [Master's thesis, San Jose State University].

Ryan, Mary Catherine Moran (2017). *Quantifying the Functional Consequences of Spanish [s] Lenition: Plural Marking and Derived Homophony in Western Andalusian and Castilian* [Master's thesis, San Jose State University].

Hernandez, Antonio de Jesus (2013). *Boom and Whoosh: Verbs of Explosion as a Change-of-State Class* [Master's thesis, San Jose State University].

Panfil, Laura Maggia (2013). *Grading Non-Gradable Adjectives: A "Totally Unique" Corpus Study* [Master's thesis, San Jose State University].

Service

Professional Service

Reviewer for *Language, Laboratory Phonology, Linguistics Applied, International Joint Conference on Natural Language Processing, Linguistics and Language Compass, Annual Meeting of the Association for Computational Linguistics, Phonology, Glossa*.

University site coordinator for the North American Computational Linguistics Olympiad, February 2011 – February 2014.

Local organization committee for the 83rd Annual Meeting of the Linguistic Society of America, January 2009.

Service at San Jose State University

Chair, Department Curriculum Committee, Fall 2024 – Present

Member, Department RTP Committee, AYs 2018-22, Fall 2023 – Present

Member, Department Curriculum Committee, Fall 2008 – Present

Member, Department Lecturer Evaluation Committee, Fall 2008 – Present

Director, Center for Human Language Technology, Fall 2008 – Present

Member, International Programs and Students Committee, AY 2021-22, AY 2023-25

Member, Department Tenure-Track Hiring Committee, AYs 2016-18, 2019-22, AY 2024-25

Member, College Curriculum Committee, AYs 2016-17, 2018-20, Fall 2022

Member, College Research Committee, AY 2010-11

Awards

Poster selected as a Centennial Session, ASHA 2025 Convention: Overfitting in deep learning models for disfluency detection in bilingual children

Faculty RSCA Assigned Time Program Award, San Jose State University, AY 2023-24

Dean's Professional Development Grant, San Jose State University, AY 2021-22

RSCA Seed Grant Program Award, San Jose State University, AY 2021-22

Faculty RSCA Assigned Time Program Award, San Jose State University, AYs 2018-23

Junior Faculty Career Development Grant, San Jose State University, AY 2011-12

College Release Time Award, San Jose State University, Spring 2009

University Fellowship, University of Illinois at Urbana-Champaign, AYs 2002-04

References

Available upon request.