

Ling Sun

Researcher in Linguistics and Cognitive Science · Computational Linguist

📍 Brooks Hall, 310 Herty Drive, Athens, GA 30602, USA

✉ Ling.Sun@uga.edu [git github.com/LingSyrina](https://github.com/LingSyrina) [G sites.google.com/ling-sun](https://sites.google.com/ling-sun)

EDUCATION

Ph.D. Linguistics

2020 – June 2026

Ph.D. Cognitive Science

Indiana University Bloomington, USA

Thesis: Grammars of Scale: How Adjective Interpretation Emerges from Structure and Scale

Advisor: Thomas Grano, Robert Goldstone

M.S. Computer Science (Admitted, coursework completed)

2023 – 2025

Indiana University Bloomington, USA

Specialization: Deep learning, systems programming, quantum computing, speech processing, and NLP.

M.A. Linguistics

2017 – 2018

University College London, UK

Thesis: Measure Phrase Distribution in Measure Constructions

Advisor: Yasutada Sudo

B.A. Applied Linguistics

2013 – 2017

Sichuan University, China

Thesis: Zero Copula in Archaic Chinese: Pause Position in Declarative Sentences

Advisor: Guo Li

PROFESSIONAL APPOINTMENTS

Assistant Professor

Aug 2026-Present

Linguistics and Institute for Artificial Intelligence, University of Georgia

Research Assistant

2025-May 2026

Speech Lab, Indiana University

Train and engineer ASR models for atypical speech; design and apply speech data augmentation; conduct model behavior analysis.

Research assistant

May - Aug 2025

Percepts and Concepts Lab, Indiana University

Designed and executed web-based data collection paradigms. Collected, maintained, and analyzed human behavior data in perception and concept learning.

Teaching Assistant

2025

Department of Psychological and Brain Sciences, Indiana University

Graded assignments/exams for K300: Statistical Techniques

Associate Instructor

2022-2024

Linguistics department, Indiana University

Designed and taught undergraduate linguistics courses (L103, L203), including syllabus creation, lectures, assignments, and exams.

Teaching Assistant

2021-2022

Linguistics department, Linguistics department

Led discussions, gave guest lectures, and graded assignments in L103: Intro to Linguistics

Language Instructor

2019-2020

New Oriental Corp., China

Designed and taught language courses for L2 English speakers, including course plan and assignments.

Semantic Annotator

2017

Qizhi Robotic Tech., China

Topic classification and semantic labeling for early dialogue systems; contributed training data and QA for ML pipelines.

HONORS AND AWARDS

Householder Best Paper Award <i>Department of Linguistics, Indiana University</i>	2025
Daniel Dinnsen Excellence in Teaching Award <i>Department of Linguistics, Indiana University</i>	2024
Gene Portteus Branigin COAS Memorial Fund <i>College of Arts & Sciences, Indiana University</i>	2024–2025
Third Place, NeurIPS 2025 Embodied Agent Interface Challenge <i>NeurIPS Competition Track; team award (OpenReview paper available)</i>	2025

RESEARCH FUNDING AND GRANTS

Indiana University Seed Grant <i>PI: S. Shi</i> This project investigates how atypical and L2 speech, linked to cognitive or motor impairments (e.g., apraxia, Down syndrome, Parkinson's) and cross-linguistic variability, impacts the interpretability and fairness of ASR systems. It models error patterns to support equitable and cognitively informed approaches to speech technology.	Candidate
---	-----------

PUBLICATIONS

A. Articles in Peer-Reviewed Conference Proceedings

Published

- [1] L. Sun, "Measure phrase modification in name adjective constructions in mandarin," *Proceedings of the Linguistic Society of America*, vol. 8, no. 1, pp. 5478–5478, 2023.
- [2] L. Sun and M. Fort, "A case study of tone realization in shaoxing opera sprechstimme," in *International Symposium on Frontiers of Research in Speech and Music*, Springer, 2024, pp. 139–149.
- [3] L. Sun, S. Kim, and R. Goldstone, "Does precision affect categorization? magnitude categorization and measurement scales," in *Proceedings of the Annual Meeting of the Cognitive Science Society*, vol. 47, 2025.
- [4] L. Sun, C. Zhu, and S. Shi, "Proficiency-aware adaptation and data augmentation for robust l2 asr," in *ICASSP 2026-2026 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2026, pp. 18 212–18 216.
- [5] L. Sun, P. Sullivan, M. Martin, and Y. Zhou, "Quantum nlp models on natural language inference," in *International Conference on Quantum Artificial Intelligence and Natural Language Processing*, Springer, 2025, pp. 114–133.

Accepted/In-press

- [1] L. Sun and R. Goldstone, "From distributional structure to meaning: Learning, syntax, and the emergence of categorical and graded interpretations," in *Proceedings of the Annual Meeting of the Cognitive Science Society*, Accepted, 2026.

Under Revision

- [1] C. Zhu, L. Sun, and S. Shi, "Speech-LLM for faithful disfluency-aware L2 transcription," in *IEEE Spoken Language Technology Workshop (SLT)*, Under review, 2026.
- [2] E. Naske, L. Sun, B. H. Glidden, and S. Shi, "A systematic study of dysarthric speech recognition across etiologies and severity levels," in *IEEE Spoken Language Technology Workshop (SLT)*, Under review, 2026.

B. Articles in Peer-Reviewed Workshop Proceedings

- [1] L. Sun, S. Kim, X. Dong, and S. Kübler, "Debiasing static embeddings for hate speech detection," in *Proceedings of the The 9th Workshop on Online Abuse and Harms (WOAH)*, 2025, pp. 67–76.

C. Peer-Reviewed Journal Articles

- [1] L. Sun, "A dynamic typology of adjectives: Measurement theory and syntactic interaction," *Logics*, vol. 3, no. 4, p. 17, 2025.

PRESENTATIONS

- [1] L. Sun, C. Zhu, and S. Shi, "Adapting speech foundation models for l2 speakers under proficiency imbalance," in *TICC Workshop 2025*, 2025.
- [2] L. Sun, "Dynamic typology of adjectives: Measurement theory and syntactic interaction," in *European Summer School in Logic, Language and Information (ESSLLI) Proceedings*, Earlier conference version, vol. 8, 2024, p. 66.

- [3] L. Sun, "The role of polarity in measure phrase constructions," in *North American Conference on Chinese Linguistics (NACCL-32)*, 2020.
- [4] L. Sun, "Measure phrase constructions in mandarin, german, dutch and english," in *3rd Crete Summer School of Linguistics*, 2019.
- [5] Q. Xiao, B. Su, L. Sun, Z. Zhu, and T. Le, "Third Place, NeurIPS EAI Challenge. CtrlAct: Grounding LLMs to Bridge the Gap between Embodied Instruction and Action," Presented at the NeurIPS 2025 Competition Track: Embodied Agent Interface (EAI) Challenge, San Diego, USA, 2025.

METHODOLOGICAL SKILLS

- **Language skills:** Mandarin (Native); English (Professional); German (Intermediate); Swahili (Intermediate); French (Beginner)
- **Programming & Systems:** Python; C; JavaScript; R; MATLAB; SQL; Linux; Git; HPC; Bash/Shell
- **Modelling & Machine Learning:** Genetic Algorithms; Quantum Computing; RNN /LSTM/GRU modeling
- **NLP & Speech:** Transformers; Embedding Space Evaluation & Finetuning, ASR; Speech Enhancement
- **Evaluation & Statistics:** Bayesian & Frequentist Stats; Data Visualization

RESEARCH HIGHLIGHTS

Proficiency-Aware Adaptation and Data Augmentation for Robust L2 ASR

Jun–Oct 2025

Submitted to ICASSP 2026

This project develops a proficiency-sensitive framework for automatic speech recognition (ASR), addressing how linguistic and cognitive variability among second-language (L2) speakers shapes communicative understanding. Using a large CEFR-graded English corpus, it demonstrates that conventional ASR systems systematically privilege higher-proficiency speakers—reflecting structural bias in how meaning and fluency are modeled. The proposed approach integrates multitask learning (ASR + proficiency prediction) and targeted augmentation to balance representation across proficiency levels. Beyond improving recognition equity, the study frames ASR misrecognition as a form of pragmatic misalignment, showing that communicative breakdowns arise not only from acoustic variation but from divergent expectations about fluency and discourse structure.

A Dynamic Typology of Adjectives: Measurement Theory and Syntactic Interaction

2025

Round 2 revision, Logics (Special Issue on Logic and Language)

Develops a formal and empirical framework linking syntactic structure, semantic interpretation, and conceptual representation in adjective meaning. Drawing on measurement theory, fuzzy logic, and philosophy of language, the paper argues that adjectives do not have fixed lexical types but acquire their interpretive scale dynamically through syntactic and pragmatic context. Using Mandarin and English data, it formalizes how syntax constrains gradability and evaluativity, which are core aspects of meaning that depend on shared interpretive norms. The approach bridges formal semantics, linguistic pragmatics, and cognitive modeling, offering a general account of how compositional structure interacts with conceptual and logical systems in natural language understanding.

Quantum Natural Language Processing for Inference

2025

Presented at Q-NLP AI 2025

Develops a formal-computational model of linguistic inference inspired by quantum logic and compositional semantics, applying it to the task of natural language inference. The framework encodes syntactic and semantic composition directly into quantum circuits, allowing meaning relations such as entailment and contradiction to be represented as operations over quantum states. By integrating category theory, formal logic, and semantic theory, the study models inference as a process of structural alignment and information exchange—a computational analogue to shared understanding and common ground in linguistic communication. The project illustrates how novel logical architectures can capture human-like reasoning about meaning and context under extreme resource constraints.

Does Precision Affect Categorization? Magnitude Categorization and Measurement Scales *2024–2025*

Published, Cog Sci 2025

Investigates how measurement systems shape conceptual structure and linguistic meaning, exploring whether precision in numerical systems influences how people form and categorize magnitude concepts. Using behavioral experiments with novel adjectives and morph continua, the study shows that exposure to high-precision systems leads to finer conceptual distinctions and expanded representational scales. By linking semantic gradability, cognitive scaling, and measurement theory, this work contributes to understanding how linguistic categories emerge from shared cognitive frameworks—connecting philosophy of language, cognitive semantics, and pragmatic reasoning about comparison and context.

Debiasing Static Embeddings for Hate Speech Detection

2024–2025

Published, WOA@ACL 2025

This work investigates how linguistic meaning representations in word embeddings encode and perpetuate social and pragmatic bias, shaping how models interpret hate and non-hate language. By treating bias not only as a statistical artifact but as a semantic and pragmatic phenomenon, the study examines how embeddings reflect implicit presuppositions and stereotype associations that distort meaning attribution. Using multi-dimensional bias metrics, it shows how differences in embedding geometry correspond to asymmetries in interpretation—for instance, how models conflate evaluative or group-referential meaning with offensiveness.

The Power of Experiment Settings in Unlocking Antonym Evaluativity Contrast

2024

HSP Conference manuscript

This project investigates how experimental framing shapes evaluativity judgments in antonym pairs (e.g., tall/short, heavy/light) and how methodological differences reveal distinct pragmatic inferences. Two behavioral experiments with Mandarin speakers compared inference-validation and sentence-completion paradigms to test how participants interpret evaluative polarity in degree and equative constructions. The findings show that evaluativity is not a fixed semantic property but an emergent pragmatic effect, modulated by task design and contextual framing. This work highlights the importance of experimental methodology as part of the communicative context, offering a dynamic perspective on how meaning and interpretation interact across linguistic polarity, context, and speaker expectation.

A Case Study of Tone Realization in Shaoxing Opera Sprechstimme

2023–2024

Published, FRSM 2024

This study examines how tonal meaning is negotiated between linguistic structure and expressive performance in Shaoxing opera, a Chinese genre characterized by *sprechstimme*, which is a singing voice acting that blurs speech and melody. Using acoustic and phonological analysis, it investigates how performers modulate lexical tones to balance semantic intelligibility, aesthetic form, and auditory perception. The project situates tone realization within the broader question of how meaning persists under expressive distortion. By analyzing how prosodic meaning correlates to deviation, the study contributes to theories of embodied pragmatics of music performance.

Use of Gradable Adjectives: Pragmatic Account and Bayesian Modelling

2022

Independent research project

This project investigated Bayesian models of how contextual standards for gradable adjectives (e.g., tall, full) emerge from pragmatic reasoning under uncertainty. Building on Qing & Franke (2014), the study compared discrete, continuous, and multi-reference Bayesian models to explore how speakers select thresholds and endpoints in adjective interpretation. The project treats vagueness and context-dependence as manifestations of dynamic inference, linking formal semantics and pragmatic efficiency with cognitive models of scalar reasoning.

Phenomenology and Neuromorphic Computing: Toward a Dynamic Theory of Meaning and Intentionality

Ongoing

Manuscript in preparation

This project bridges phenomenology, philosophy of language, and cognitive modeling to develop a neuromorphic account of meaning and intentionality. It challenges classical representational theories of language, which treat meaning as internal correspondence, by modeling it as an emergent property of agent–environment interaction. Drawing on Husserlian intentionality, embodied cognition, and dynamical systems theory, the study proposes that linguistic understanding arises from feedback-based coupling rather than symbolic mapping. Using examples from speech perception and referential communication models, the project demonstrates how communicative meaning can stabilize through temporal coordination and shared dynamics, without pre-specified internal representations. This view integrates phenomenological insight with neuromorphic computation, suggesting that common ground and mutual understanding emerge from interactive processes of adaptation and alignment—reframing intentionality as co-determined stability rather than mental correspondence.

A Neural Model Approach to Individual Differences in Syntactic Adaptation

Ongoing

This study investigates how individuals differ in adapting to syntactic ambiguity, offering a computational framework that connects psycholinguistic adaptation, formal language modeling, and cognitive variability. Using human reading-time data and neural network models trained on author-specific corpora, the project demonstrates that exposure-based syntactic learning is not uniform across speakers but reflects distinct

individual adaptation profiles. By modeling surprisal and processing difficulty through individualized neural architectures, the study reframes syntactic adaptation as a problem of pragmatic accommodation and shared expectation—how interlocutors align grammatical predictions over time. The results link variation in model behavior to conceptual and experiential differences among speakers, showing that language understanding involves dynamic negotiation of structure and meaning. The broader implication situates syntactic adaptation within theories of common ground formation, cognitive flexibility, and the logic of linguistic coordination.