Sathvik Nair

sathvik@umd.edu • sathvikn.github.io

Education

2022 – Present **University of Maryland, College Park**
PhD, Linguistics
Advisors: Dr. Philip Resnik, Dr. Colin Phillips

2016 – 2020 **University of California, Berkeley**
B.A., majors in Computer Science & Cognitive Science
Highest Honors in Cognitive Science
Thesis Title: Attention-Based Neural Networks Encode Aspects of Human-Like Word Sense Knowledge
Advisor: Dr. Mahesh Srinivasan, direct supervision by Dr. Stephan Meylan (MIT)

Fellowships and Awards

2024-2027 National Science Foundation Graduate Research Fellowship (GRFP) *$159,000*

2023-2024 Great Reviewer, Association for Computational Linguistics (Linguistic Theories, Cognitive Modeling, and Psycholinguistics & Interpretability and Analysis of Models for NLP tracks)

2022 Honorable Mention, NSF GRFP

2020 Robert J. Glushko Prize for Distinguished Undergraduate Research, UC Berkeley Cognitive Science

Publications

2024 **Nair, S.** *, Howitt, K. G.*, Dods, A., Hopkins, R.M. (2024) Generalizations across filler-gap dependencies in neural language models. *Accepted to CoNLL.* *equal contribution

2024 Lee, E.K., **Nair, S.**, Feldman, N. H. (2024) A Psycholinguistic Evaluation of Language Models' Sensitivity to Argument Roles. *Accepted to Findings of EMNLP.*

2023 **Nair, S.** & Resnik, P. (2023) Words, Subwords, and Morphemes: What Really Matters in the Surprisal-Reading Time Relationship?. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 11251–11260, Singapore. Association for Computational Linguistics.

2021 Meylan, S.C., **Nair, S.**, & Griffiths, T.L. (2021) Evaluating Models of Robust Word Recognition with Serial Reproduction. *Cognition* 210

2020 **Nair, S.**, Srinivasan, M., & Meylan, S.C. (2020). Contextualized Word Embeddings Encode Aspects of Human-Like Word Sense Knowledge. Proceedings of the Workshop on the Cognitive Aspects of the Lexicon (*CogALex*) VI at Interational Conference in Computational Linguistics (COLING 2020), 129 - 141.

Conference Presentations

2024 **Nair, S.**, Howitt, K. G., Dods, A., Hopkins, R.M. LMs are not good proxies for human language learners. 49th Annual Meeting of the Boston University Conference on Language Development (BUCLD). Boston, Massachusetts. *(Talk)*

2024 **Nair, S.** & Resnik, P. Words, Subwords, and Morphemes: What Really Matters in the Surprisal-Reading Time Relationship? 7th Annual Meeting of the Society for Computation in Linguistics. Irvine, California. *(Talk)*

2024 **Nair, S.**, Phillips, C. & Resnik, P. Words, Subwords, and Morphemes: Surprisal Theory and Units of Prediction. 37th Annual Conference on Human Sentence Processing, Ann Arbor, Michigan. *(Poster)*

2024 Howitt, K. G., **Nair, S.**, Dods, A., Hopkins, R.M. Acquiring generalizations across unbounded dependencies: How language models can provide insight into first language acquisition. 11th Mid-Atlantic Student Colloquium on Speech, Language and Learning. Baltimore, Maryland. *(Poster)*

2024 **Nair, S.**, Kahadze, K., Resnik P. The Impacts of Subword Tokenization on Psycholinguistic Modeling. 11th Mid-Atlantic Student Colloquium on Speech, Language and Learning. Baltimore, Maryland. *(Poster)*

2023 **Nair, S.**, Bhattasali, S., Resnik, P.S., and Phillips, C. How far does probability take us when measuring psycholinguistic fit? Evidence from Substitution Illusions and Speeded Cloze Data. 36th Annual Conference on Human Sentence Processing, Pittsburgh, Pennsylvania. *(Poster)*

Invited Talks

February 2024 Saarland University Language Science & Technology Department Virtual Lab Meeting (PIs: Drs. Michael Hahn, Matthew Crocker, Vera Demberg), Saarland, Germany. *Subword Tokenization in Psycholinguistic Modeling.*

February 2024 Maryland Language Science Center, Language Science Lunch Talk Series. *From Code Switching to Coding to Both? "Growing Up" Academically in a Changing Interdisciplinary Field.*

November 2023 — UCLA National Heritage Language Center Panel on AI, joint presentation with Utku Turk, *LLMs for Language Research.*

October 2023 — ONR MURI on Document Comprehension Annual Review Meeting, George Mason University. *Prediction in Language Comprehension: Exploring Differences Between Humans and Machines.*

Teaching experience

University of Maryland, College Park

Fall 2024 — ARHU 299: Machine Learning in Language and Art.
Teaching Assistant
Primary Instructor: Dr. Omar Agha

Fall 2023 — LING200: Introductory Linguistics.
Teaching Assistant
Primary Instructor: Dr. Margaret Antonisse
LING499-D: Undergraduate Seminar in Psycholinguistics
Guest Lecturer: Prediction in Humans and Machines
Primary Instructor: Eun-Kyoung Rosa Lee

University of California Berkeley

Spring 2020 — COGSCI 131: Computational Models of Cognition.
Undergraduate Student Instructor
Primary Instructor: Dr. Steven Piantadosi

Spring 2018-Fall 2019 — DATA 8: Foundations of Data Science
Undergraduate Student Instructor (Spring 2019 and Fall 2019), Tutor (Spring, Summer, and Fall 2018)
Primary Instructors: Drs. Ani Adhikari, Will Fithian, David Wagner, Ramesh Sridharan and Swupnil Sahai, Vinitra Swamy and Fahad Kamran.

Industry Experience

2020-2022 — **Amazon Web Services.** Boston, Massachusetts
Software Development Engineer, EC2 Elastic Block Store

2019 — **Workday.** Pleasanton, California
Software Development Engineering Intern, Object Management System

Reviewing

2024-Present — Workshop on Cognitive Modeling and Computational Linguistics

2023-Present — Association for Computational Linguistics Annual Rolling Review

2024 Annual Conference of the Cognitive Science Society, Journal of Memory and Language, Secondary reviewer for Annual Reviews in Linguistics and Glossa Psycholinguistics

2022 Language Resources & Evaluation

Service & Outreach

2023-Present Site Coordinator, North American Computational Linguistics Olympiad (2023-2024)

2024 Presenter, *LLMs in Language Science Research* with Rupak Sarkar, Maryland Language Science Center

2023 Discussion Leader, *Machine Learning and Language Technology*, Gemstone Honors Program, University of Maryland

2019 Presenter, *Introduction to Natural Language Processing and Machine Learning with Python*, Spectra Hackathon, Make School, San Francisco, CA (July 2019)

2018 Panelist, *Undergraduate Data Science, Pedagogy and Practice*, Division of Data Sciences, UC Berkeley

Mentorship

2024-Present Konstantine Kahadze, UMD Linguistics & Computer Science

2023-2024 Robert Melvin Hopkins, UMD Linguistics & Computer Science

Professional Membership

Cognitive Science Society, Society for Human Sentence Processing, Association for Computational Linguistics

Technical Skills

Programming:
Python, Java, R, SQL, HTML/CSS/JavaScript

Technical Tools:
Data Analysis (Jupyter, Pandas, Numpy, Matplotlib, Seaborn, dplyr, ggplot), Machine Learning/NLP (Huggingface, Pytorch, NLTK, Scikit-Learn), Other (UNIX, SLURM, bash, Git, Docker, AWS).

Research Methods:
Corpus analyses, computational modeling, behavioral experimentation

Languages:
English (native), Hindi-Urdu, French (conversational; spoken & written)