

Sathvik Nair

sathvik@umd.edu • sathvikn.github.io

Education

- 2022 – Present **University of Maryland, College Park**
PhD in Linguistics with certificate in Neuroscience & Cognitive Science
Dissertation title (expected defense Spring 2027): *The Proper Role of Probabilistic Inference in Modeling Psycholinguistic Processes*
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
- 2026 Keeping Science Going Grant, Maryland Language Science Center \$720
- 2024 Summer Fellowship, University of Maryland College of Arts and Humanities \$1300
- 2023 Summer Fellowship, University of Maryland College of Arts and Humanities \$1200
- 2022 Honorable Mention, NSF GRFP
- 2020 Robert J. Glushko Prize for Distinguished Undergraduate Research, UC Berkeley Cognitive Science

Publications

(* = equal contribution)

- 2026 **Sathvik Nair*** & Byung-Doh Oh* (2026). Clozing the Gap: Exploring Why Language Model Surprisal Outperforms Cloze Surprisal. *Accepted to ACL 2026*
- 2026 Atrey Desai & **Sathvik Nair** (2026). Filling in the Mechanisms: How do LMs Learn Filler-Gap Dependencies under Developmental Constraints? *Accepted to ACL Findings 2026*
- 2026 **Sathvik Nair** & Colin Phillips (2026). Across the Levels of Analysis: Explaining Predictive Processing in Humans Requires More Than Machine-Estimated Probabilities. *Accepted at Behavioral and Brain Sciences*, Commentary on target article by Futrell & Mahowald.

- 2026 Chiebuka Ohams*, **Sathvik Nair***, Shohini Bhattachali, & Philip Resnik (2026). A predictive coding model for online sentence processing. *Journal of Memory and Language*, 146, 104705.
- 2024 Katherine Howitt, **Sathvik Nair**, Allison Dods, & Robert Melvin Hopkins (2024) Generalizations across filler-gap dependencies in neural language models. In *Proceedings of the 28th Conference on Computational Natural Language Learning*, pages 269–279, Miami, FL, USA. Association for Computational Linguistics.
- 2024 Eun-Kyoung Rosa Lee, **Sathvik Nair**, Naomi Feldman (2024). A Psycholinguistic Evaluation of Language Models’ Sensitivity to Argument Roles. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 3262–3274, Miami, Florida, USA. Association for Computational Linguistics.
- 2023 **Sathvik Nair** & Philip Resnik (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 Stephan Meylan, **Sathvik Nair**, & Griffiths, T.L. (2021) Evaluating Models of Robust Word Recognition with Serial Reproduction. *Cognition* 210
- 2020 **Sathvik Nair**, Srinivasan, M., & Stephan Meylan (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.

Peer-Reviewed Conference Presentations

- 2026 **Sathvik Nair** & Byung-Doh Oh. Should Language Models Replace the Cloze Task Forever? 39th Annual Conference in Human Sentence Processing, Massachusetts Institute of Technology, Cambridge, MA. (*Poster*)
- 2026 **Sathvik Nair**, Philip Resnik & Colin Phillips. *Sampling-Based Assumptions Simplify Modeling of Speeded Cloze Response Times*. 39th Annual Conference in Human Sentence Processing, Massachusetts Institute of Technology, Cambridge, MA. (*Poster*)
- 2026 Samer Nour Eddine, **Sathvik Nair**, Alba Jorquera Jiménez de Aberásturi, Jonathan Simon & Philip Resnik Oh. *Should Language Models Replace the Cloze Task Forever?* 39th Annual Conference in Human Sentence Processing, Massachusetts Institute of Technology, Cambridge, MA. (*Oral*)
- 2025 Eun-Kyoung Rosa Lee, **Sathvik Nair (presenter)**, & Naomi Feldman. *(Un)likely words in context: A divergence between humans and large language models*. 38th Annual Conference on Human Sentence Processing, University of Maryland, College Park, MD. (*Talk*)

- 2025 **Sathvik Nair**, Cassandra Jacobs, Philip Resnik, & Colin Phillips. *Evidence for the contribution of semantic factors on speeded cloze production times*. 38th Annual Conference on Human Sentence Processing, University of Maryland, College Park, MD. (Poster)
- 2024 **Sathvik Nair**, Katherine Howitt, Allison Dods & Robert Melvin Hopkins. *LMs are not good proxies for human language learners*. 49th Annual Meeting of the Boston University Conference on Language Development (BUCLD). Boston, MA. (Talk)
- 2024 **Sathvik Nair (presenter)** & Philip Resnik. *Words, Subwords, and Morphemes: What Really Matters in the Surprisal-Reading Time Relationship?* 7th Annual Meeting of the Society for Computation in Linguistics. University of California, Irvine, Irvine, CA. (Talk)
- 2024 **Sathvik Nair**, Colin Phillips & Philip Resnik. *Words, Subwords, and Morphemes: Surprisal Theory and Units of Prediction*. 37th Annual Conference on Human Sentence Processing, University of Michigan, Ann Arbor, MI. (Poster)
- 2024 Katherine Howitt, **Sathvik Nair**, Allison Dods & Robert Melvin Hopkins. *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. Johns Hopkins University, Baltimore, MD. (Poster)
- 2024 **Sathvik Nair**, Konstantine Kahadze & Philip Resnik. *The Impacts of Subword Tokenization on Psycholinguistic Modeling*. 11th Mid-Atlantic Student Colloquium on Speech, Language and Learning. Johns Hopkins University, Baltimore, MD. (Poster)
- 2023 **Sathvik Nair**, Shohini Bhattachali, Philip Resnik, & Colin Phillips. *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, University of Pittsburgh, Pittsburgh, PA. (Poster)

Invited Talks

- September 2025 Panelist, *LLMs in Research and Practice* with Drs. Rachel Rudinger, José Ortiz, & Jing Liu, moderated by Katherine Howitt, Maryland Language Science Center
- April 2025 *Context-driven expectations in human sentence processing: A sample of data from language models and cloze timing*. Computation & Psycholinguistics Lab (PI, Dr. Tal Linzen) New York University.
- February 2025 *Statistics or Morphology: Subword Units and Surprisal in Sentence Processing*, Lab meeting, Cognitive Science Department (PI: Dr. John Hale) Johns Hopkins University.
- February 2024 *Subword Tokenization in Psycholinguistic Modeling*. Virtual lab meeting, Saarland University Language Science & Technology Department (PIs: Drs. Michael Hahn, Matthew Crocker, Vera Demberg).
- November 2023 *LLMs for Language Research*, joint presentation with Utku Turk. UCLA National Heritage Language Center Panel on AI.

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

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

Skills

Programming:

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

Technical Tools:

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

Research Methods:

Corpus analyses, computational modeling, behavioral experimentation

Languages:

English (native), Hindi-Urdu, French (conversational), Spanish (basic)

Teaching experience

University of Maryland, College Park

Fall 2025 LING 449B: Seminar in Psycholinguistics: Prediction, Computation, and Language Processing
Instructor of Record

Fall 2024 ARHU 299: Machine Learning in Language and Art
Teaching Assistant & Guest Lecturer: Levels of Analysis and Large Language Models
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.

Mentorship

2026	Rhea Choudhury, UMD Computer Science & Math
2025-2026	Atrey Desai, UMD Computer Science, Linguistics & Korean Studies
2024-2025	Konstantine Kahadze, UMD Linguistics & Computer Science undergraduate, currently Developer Experience Engineer at Open AI
2023-2024	Robert Melvin Hopkins, UMD Linguistics & Computer Science undergraduate, incoming Computational Linguistics MS student at University of Washington

Reviewing

2026	Society for Computation in Linguistics
2024-Present	Workshop on Cognitive Modeling and Computational Linguistics
2023-Present	Association for Computational Linguistics Annual Rolling Review, <i>Recognized as Great Reviewer in NLP Interpretability and Cognitive Modeling tracks</i>
2024	Journal of Memory and Language, Annual Conference of the Cognitive Science Society, 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
2024	Presenter, <i>LLMs in Language Science Research</i> with Rupak Sarkar, Maryland Language Science Center
2023	Discussion Leader, <i>Machine Learning and Language Technology</i> , Gemstone Honors Program, University of Maryland
2019	Presenter, <i>Introduction to Natural Language Processing and Machine Learning with Python</i> , Spectra Hackathon, Make School, San Francisco, CA (July 2019)

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

Professional Membership

Society for Human Sentence Processing, Association for Computational Linguistics