Curriculum Vitae of Dr. Stephan C. Meylan, Ph.D.

Massachusetts Institute of Technology
Department of Brain and Cognitive Sciences
77 Massachusetts Avenue
Cambridge, MA 02139-4307

Phone: (727) 278-1398
Email: smeylan@mit.edu
Homepage: stephanmeylan.com

Education

Ph.D. Psychology, University of California, Berkeley, 2018.
Advisor: Dr. Thomas L. Griffiths
Committee: Dr. Thomas L. Griffiths, Dr. Mahesh Srinivasan, Dr. Terry Regier, Dr. Michael C. Frank
Dissertation title: "Representing linguistic knowledge with probabilistic models"

B.A. *magna cum laude* Linguistics, Brown University, 2010.

Study Abroad: Casa De Las Americas in Havana, Cuba (through Brown University), Fall 2008

Academic Positions

2024 - Present: Research Scientist, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology. PI: Dr. Roger Levy.

2019 - 2024: Postdoctoral Fellow, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology. PIs: Dr. Roger Levy (MIT) and Dr. Elika Bergelson (Duke University / Harvard University).

2023 - Present: Research Associate, Harvard University Department of Psychology.

2018 - 2019: Postdoctoral Associate, Department of Psychology and Neuroscience, Duke University. PIs: Dr. Elika Bergelson (Duke University) and Dr. Roger Levy (MIT).

Publications

Journal Articles

Meylan, S.C., Griffiths, T.L. (2024) Word forms reflect trade-offs between speaker effort and robust listener recognition. *Cognitive Science* 48 (e13478).

Meylan, S.C., Foushee, R., Wong, N.H., Bergelson, E. and Levy, R.P. (2023) How adults understand what young children say. *Nature Human Behavior*.

Zettersten, M., Yurovsky, D., Xu, T.L., Uner, S., Tsui, A., Schneider, R.M., Saleh, A., **Meylan, S.C.,** Marchman, V.A., Mankewitz, J., MacDonald, K., Long, B., Lewis, M., Kachergis, G., Handa, K., deMayo, B., Carstensen, A., Braginsky, M., Boyce, V., Bhatt, N., Bergey, C., and Frank, M.C. (2022) Peekbank: An open, large-scale repository for developmental eye-tracking data of children's word recognition. *Behavior Research Methods*.

Meylan, S.C., Bergelson, E. (2022) Learning through processing: Towards an integrated approach to early word learning. *Annual Review of Linguistics* 8.

Meylan, S.C., Nair, S., and Griffiths, T.L. (2021) The challenges of large-scale, web-based language datasets: Word length and predictability revisited. *Cognitive Science* 45 (6).

Meylan, S.C., Nair, S., and Griffiths, T.L. (2021) Evaluating models of robust word recognition with serial reproduction. *Cognition* 210.

*Sanchez, A., ***Meylan, S.C**., Braginsky, M., MacDonald, K.E., Yurovsky, D., and Frank, M.C. (2019). childes-db: a flexible and reproducible interface to the Child Language Data Exchange System. *Behavior Research Methods* 51 1928–1941.

Meylan, S.C., Frank, M. C., Roy, B.C., and Levy, R. (2017). The emergence of an abstract grammatical category in children's early speech. *Psychological Science* 28 (2), pp. 181 - 192.

Kurumada, C., **Meylan, S. C.**, and Frank, M. C. (2013). Zipfian frequency distributions facilitate word segmentation in context. *Cognition* 127, 439–453.

* indicates co-first authorship.

Manuscripts Under Review / In Revision / In Preparation

Meylan, S.C., Levy, R.P., and Bergelson, E. (in revision) Children's expressive and receptive knowledge of the English regular plural.

Meylan, S.C., Levy, R.P., and Beguš, G. (in prep.) A multitask, multimodal neural language learner.

Meylan, S.C., Foushee, R. (in prep.) Discourse-first language acquisition.

Meylan, S.C., Mankewitz, J., Floyd, Sammy, Rabagliati, H., and Srinivasan, M. (in prep.) The prevalence of polysemy in first language learning.

Zhi, S., Levy, R.P., and **Meylan, S.C.** (in prep.) Modeling the contribution of facial visual input to phonetic category learning.

Refereed Conference Proceedings

Meylan, S.C., Foushee, R., Bergelson, E., and Levy, R.P. (2021). Child-directed Listening: How Caregiver Inference Enables Children's Early Verbal Communication. *Proceedings of the 43rd Annual Meeting of the Cognitive Science Society*, Vienna. *(talk & proceedings paper)*

***Meylan, S.C.**, *Mankewitz, J., Floyd, S., Rabagliati, H., and Srinivasan, M. (2021). Quantifying Lexical Ambiguity in Speech To and From English-Learning Children. *Proceedings of the 43rd Annual Meeting of the Cognitive Science Society*, Vienna. *(talk & proceedings paper)*

Zettersten, M., Bergey, C.A., Bhatt, N.S., Boyce, V., Braginsky, M., Carstensen, A., deMayo, B., Kachergis, G., Lewis, M., Long, B. MacDonald, K., Mankewitz, J., **Meylan, S.C.**, Saleh, A.N., Schneider, R.M., Tsui, A.S.M, Uner, S., Xu, T.L., Yurovsky, D., Frank, M.C. Peekbank: Exploring children's word recognition through an open, large-scale repository for developmental eye-tracking data. *Proceedings of the 43rd Annual Meeting of the Cognitive Science Society*, Vienna. *(poster & proceedings paper)*

Nair, S., Srinivasan, M., and **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 / COLING*) VI ,129–141. *(talk & proceedings paper)*

Meylan, S.C., Levy, R.P., and Bergelson, E. (2020). Children's Expressive and Receptive Knowledge of the English Regular Plural. 42nd Annual Meeting of the Cognitive Science Society, 29 July - 1 August 2020. *(poster & proceedings paper)*

Nematzadeh, A., **Meylan, S.C.**, and Griffiths, T.L. (2017). Evaluating Vector-Space Models of Word Representation. Proceedings of the 37th Annual Meeting of the Cognitive Science Society, London. *(talk & proceedings paper)*

Meylan, S. C., and Griffiths, T.L. (2015). A Bayesian framework for learning words from multiword utterances. *Proceedings of the 37th Annual Meeting of the Cognitive Science Society*, Pasadena. *(talk & proceedings paper)*

Meylan, S. C., and Gahl, S. (2014). The divergent lexicon: Lexical overlap decreases with age in a large corpus of conversational speech. *Proceedings of the 36th Annual Meeting of the Cognitive Science Society*, Quebec City. *(talk & proceedings paper)*

Meylan, S. C., Frank, M. C., and Levy, R. (2013). Modeling the development of determiner productivity in children's early speech. *Proceedings of the 35th Annual Meeting of the Cognitive Science Society*, Berlin. *(talk & proceedings paper)*

Meylan, S. C., Kurumada, C., Börschinger, B., Johnson, M., and Frank, M. C. (2012). Modeling online word segmentation performance in structured artificial languages. *Proceedings of the 34th Annual Meeting of the Cognitive Science Society*, Sapporo. *(poster & proceedings paper)*

Kurumada, C., **Meylan, S. C.**, and Frank, M. C. (2011). Zipfian word frequencies support statistical word segmentation. *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society (BUCLD)*, Boston. *(talk & proceedings paper)*

* indicates co-first authorship.

Refereed Conference Presentations and Posters

Zhi, S., Levy, R.L., and **Meylan, S.C.** (2024). Multimodal Input Aids a Bayesian Model of Phonetic Learning. Member abstract presented at the 46th Annual Meeting of the Cognitive Science Society. *(poster)*

Mankewitz, J., **Meylan, S.**, Sammy, F., Rabagliati, H., and Srinivasan, M. (2021). English-learning children hear and use multiple meanings for words in early speech. 46th Annual Boston University Conference on Language Development (BUCLD), Boston (*talk*)

Zettersten, M., Saleh, A., Bhatt, N., Yurovsky, D., Xu, T.L., Uner, S., Tsui, A.S.M., Schneider, R., **Meylan, S.**, Marchman, V., Mankewitz, J., MacDonald, K., Long, B., Lewis, M., Kachergis, G., DeMayo, B., Carstensen, A., Braginsky, M., Boyce, V., Bergey, C., and Frank, M. (2021). Increases in speed and accuracy of children's online word recognition measured via a large-scale, open database of developmental eye-tracking data. 46th Annual Boston University Conference on Language Development (BUCLD), Boston (*talk*).

Meylan, S.C., Bergelson, E., and Levy, R.P. (2020). Characterizing Child-Directed Listening with Corpus and Model-based Analyses. 22nd Biennial International Conference of Infant Studies, 6-9 July 2020. *(talk)*

Meylan, S., Braginsky, M., DeMayo, B., Sanchez, A., Schonberg, C., Srinivasan, M., Vlach, H., Lupyan, G., Griffiths, T., and Frank, M. (2019). Wordful: Tracking Early Productive Vocabulary Growth with Smartphones. 44th Annual Conference on Language Development (BUCLD), Boston. *(talk)*

Meylan, S.C., and Griffiths, T.L. (2017). Statistical language models reveal cross-linguistic differences in processing difficulty. Poster presentation at the 2017 CUNY Sentence Processing Conference, Boston, 30–31 March 2017. *(poster)*

Braginsky, M., **Meylan, S.C.**, and Frank, M.C. (2016). Gender differences in lexical input and acquisition. 41st Annual Boston University Conference on Language Development (BUCLD), Boston. *(talk).*

Suchow, J.W., Morgan, T.J.H., Hamrick, J., Pacer, M., **Meylan, S.C.**, and Griffiths, T.L. (2016). Wallace: automating cultural evolution experiments through crowdsourcing. Tutorial presentation at the 38th Annual Conference of the Cognitive Science Society, Philadelphia. *(workshop talk)*

Meylan, S.C., Frank, M.C., Roy, B.C., and Levy, S.M. (2016). Bayesian modeling reveals the early emergence of syntactic abstraction in children's speech. 20th Biennial Congress of the International Congress on Infant Studies, New Orleans. *(talk)*

Suchow, J.W., Morgan, T.J.H., Hamrick, J., Pacer, M., **Meylan, S.C.**, and Griffiths, T.L. (2015). Wallace: A platform for simulating cultural evolution in structured populations online. ACM Conference on Economics and Computation, Portland, Oregon. *(talk)*

Meylan, S. C., and Griffiths, T.L. (2015). Word forms—not just their lengths—are optimized for efficient communication. Poster presentation at the 2015 CUNY Sentence Processing Conference, University of Southern California, 19–22 March 2015. *(poster)*

Meylan, S. C., and Griffiths, T.L. (2014). Exploring Inductive Biases In Naturalistic Language Use. Member abstract presented at the 36th Annual Meeting of the Cognitive Science Society. *(poster)*

Non-Refereed Technical Reports

Robbins, L.L., Hansen, M.E., Kleypas, J.A., and **Meylan, S.C.** (2010). CO2calc: A user-friendly seawater carbon calculator for Windows, Max OS X, and iOS (iPhone). U.S. Geological Survey Open-File Report 2010–1280. *425 citations as of September 2021*

Funding and Grants

Economic and Social Research Council "Charting lexical development through dense coding and analysis of word senses'.' 2021 - 2023. Lead PI: Dr. Hugh Rabagliati. £300k.

National Institutes of Health F32 Postdoctoral Fellowship, "Linguistic Experience and Generalization: Early Links between Sounds, Words, and Grammar'.' 2019 - 2022. PI: Stephan Meylan. $197k.

National Science Foundation Graduate Research Fellowship (GRFP), 2012-2017, $147k.

Invited Talks

University of Edinburgh School of Informatics "Reverse engineering the child language learner" March 12th, 2024

University of Edinburgh Psychology Department "Language Learning, Language Processing, And The Space In Between" October 30, 2023

Harvard Language and Cognition Talk Series "A discourse-first model of early child language learning" October 3, 2023

MIT Cog Lunch "Word forms reflect trade-offs between speaker effort and robust listener recognition" March 14, 2023

Northwestern University Linguistics Department "Closing the loop: Characterizing child and caregiver language learning within dyads" March 11, 2022

Berkeley Language and Cognition Lab (PI: Dr. Mahesh Srinivasan) "Child-directed listening: How adults understand what young children say"' January 27, 2022

Dartmouth Cognitive Brown Bag "Child-directed Listening: How Caregivers' Noisy Channel Inferences Enable Children's Early Verbal Communication"' December 6, 2021

Boston College Language Learning Lab (PI: Dr. Joshua Hartshorne) "Child-directed Listening: How Caregivers' Noisy Channel Inferences Enable Children's Early Verbal Communication" October 26, 2021

Max Planck Institute for Evolutionary Anthropology, Leipzig Lectures on Language Talk Series (PI: Dr. Angela D. Friederici) "The nature and origins of grammatical productivity" July 21, 2021

Stanford Language and Cognition Lab (PI: Dr. Mike Frank) "Child-directed Listening: How Caregivers' Noisy Channel Inferences Enable Children's Early Verbal Communication" May 7, 2021

MIT Cog Lunch "Child-directed Listening: How Caregivers' Noisy Channel Inferences Enable Children's Early Verbal Communication' May 18, 2021.

MIT Cog Lunch "The origins of grammatical productivity" August 4, 2020

UC Berkeley Computation and Language Lab (PI: Dr. Steven Piantadosi) "The Telephone Game: Evaluating and improving language models with a large-scale serial transmission experiment." March 22nd, 2019

UC Berkeley NLP Seminar. "Word forms—not just their lengths—are optimized for efficient communication." February 13th, 2017.

MIT Computational Psycholinguistics Lab (Brain and Cognitive Science). "A rose by any other name would be less communicatively efficient: Investigating the interface of language structure and processing." November 1st, 2017

Workshops and Hackathons

"childes-db-derived: Developing a consistent interface to heterogeneous language datasets" (March 2024)

"Peekbank: A database of child eyetracking datasets for studying word learning" (August 2021, 2022, and 2023)

"Measuring grammatical productivity with Bayesian graphical models" Leipzig Lectures on Language (PI: Dr. Angela D. Friederici) October 20-21, 2021.

"Scientific Reports" June 4, 2020. Computational Psycholinguistics Lab, MIT (with Mika Braginsky)

"How to Compute: An Introduction to Linux-based Scientific Computing" June 1, 2020. Computational Psycholinguistics Lab, MIT.

"childes-db: A flexible and reproducible interface to CHILDES." March 8, 2019. Duke - University of North Carolina, Chapel Hill Talk Series.

"An introduction to plotting in R." October 6, 2017. Quantitative Analysis and Coding Knowledge series, UC Berkeley Department of Psychology.

Teaching Experience

Guest Lecturer, Language Acquisition, September 2021 and November 2022 (Principal Instructor: Dr. Elika Bergelson)

Guest Lecturer, Data on the Mind, April 2017 (Principal Instructor: Dr. Yang Xu)

Graduate Student Instructor, Computational Models of Cognition, Spring 2015 (Principal Instructor: Dr. Tom Griffiths)

Graduate Student Instructor, Computational Models of Cognition, Spring 2014 (Acting Graduate Instructor: Joshua Abbott)

Undergraduate Teaching Assistant, Introduction to Linguistic Theory, Fall 2009 (Principal Instructor: Dr. Pauline Jacobson)

Research Experience

Graduate Student Researcher, DARPA Next Generation Social Science (NGS2) program, Fall 2017 - Spring 2018 (Principal Investigator: Dr. Thomas Griffiths)

Graduate Student Researcher, Data on the Mind: Center for Data-Intensive Psychological Science, Fall 2014-Spring 2015 (Principal Investigator: Dr. Thomas Griffiths)

Research Assistant, Stanford Language and Cognition Lab, Fall 2010 - Summer 2012 (Principal Investigator: Dr. Michael C. Frank)

Research Assistant, USGS Coastal and Marine Geology Program, summers 2006 - 2010 (Principal Investigator: Dr. Lisa Robbins)

Industry Experience

Platform Analyst / Statistician, Crowdflower (later Figure Eight, acquired subsequently by Appen), San Francisco CA, Fall 2011 – Spring 2012.

Data Science Intern, Crowdflower, San Francisco CA, Summer 2011.

Maintained Software Projects

Database designer and backend engineer for *Peekbank*, A database of child eyetracking datasets for studying word learning (`peekbank.stanford.edu`)

Project lead and backend engineer for *Wordful*, A smartphone app for tracking early vocabulary growth (`wordfulapp.com`)

Project lead and backend engineer for *childes-db*: a reproducible interface to the Child Language Data Exchange System (`childes-db.stanford.edu`)

Developer for *CO2calc*: A User-Friendly Seawater Carbon Calculator for Windows, Mac OS X, and iOS (`pubs.usgs.gov/of/2010/1280`). Technical report has received approximately 400 citations by July 2020.

My Github profile is `github.com/smeylan`. Additional code samples available on request (less than 20% of my code is publicly available at present).

Mentorship

Master's Theses Advised

Sophia Zhi, "Unsupervised Phonetic Category Learning from Audio and Visual Input" (MIT, Masters in Engineering): 2022-2023 (Principal Thesis Advisor: Dr. Roger Levy)

Undergraduate Theses Advised

Sathvik Nair (UC Berkeley): Spring 2020 (Principal Thesis Advisor: Dr. Mahesh Srinivasan). Received Robert J. Glushko Prize for Distinguished Undergraduate Research at UC Berkeley

Aucher Serr (UC Berkeley): Spring 2014 (Principal Thesis Advisor: Dr. Mahesh Srinivasan)

Research Assistants

Arjun Pawar (UCLA, through MIT Brain and Cognitive Sciences): Summer 2024

Sophia Zhi (MIT): Spring 2021, Fall 2022. Machine learning intern at Meta; Machine learning engineer at Kensho

Nicole Wong (MIT): Spring - Summer 2021; Summer 2023. Masters of Engineering student at MIT

Esat Boucaud (MIT Summer Research Program): Summer 2022

Isha Mahadeshwar (Duke; North Carolina School of Science and Mathematics): Spring 2022

Anirudh Kothapalli (Berkeley): Spring 2022

Andrew Wang (UC Berkeley): Spring 2020; as of 2020, engineer at Salesforce

Jessica Mankewitz (UC Berkeley): Fall 2019 - Spring 2020; lab manager for Dr. Michael Frank, now PhD student at U. Wisconsin Psychology

Dee Dong (post-bac RA at UC Berkeley): Spring - Fall 2019; senior backend engineer at Snowflake, as of 2023, MIT Sloan MBA

Sathvik Nair (UC Berkeley): Fall 2017 - Summer 2018; engineer at Amazon AWS then PhD student at U. Maryland Linguistics; NSF Graduate Research Fellowship Recipient

Neha Dabke (UC Berkeley): Summer - Fall 2017

Aditya Bhumbla (UC Berkeley): Spring 2017

Teeranan (Ben) Pokaprakarn (UC Berkeley): Spring - Fall 2015; Ph.D. student in bioinformatics at UNC Billings School of Public Health

Wesley Hsieh (UC Berkeley): Spring 2015

Naomi Jing (UC Berkeley): Spring 2015 - Fall 2016

Xiyue (Sissi) Wang (UC Berkeley): Spring 2014 - Fall 2016 (please note that this is *not* the Princeton History department student detained by Iran; I can't help with related inquiries)

Reviewing

Conferences

Annual Conference of the Cognitive Science Society (CogSci), 2014 - 2024

Architectures and Mechanisms of Language Processing (AMLAP) 2024

Human Sentence Processing (HSP) 2024

Many Paths to Language (MPaL) 2023

Conference on Computational Natural Language Learning (CONLL) 2021-2022

Conference on Empirical Methods in Natural Language Processing (EMNLP) 2022

International Conference on Learning Representations (ICLR) 2020

Journals ad hoc reviewer for:

Cognition

Cognitive Science

Journal of Experimental Psychology: General

Open Mind

Journal of Memory and Language

Developmental Psychology

Language Development Research

Language Learning and Development

Behavior Research Methods

Scientific Reports

Evolution and Human Behavior

Funding Organizations

ad hoc reviewer for National Science Foundation (USA) doctoral dissertation improvement grant

Media Coverage

Newsweek, October 2023, "Why you can understand what your baby is trying to say"

MIT News, October 2023, "How adults understand what kids are saying"

MIT News, April 2017, "Articles of faith"

Stanford News, February 2017, "Toddlers' grammar skills not inherent, but learned, new Stanford research says"'