

Vy Ai Vo, PhD

<https://vyaivo.github.io> | vyaivo@gmail.com

POSITIONS

2019 – 2024: AI / ML Research Scientist
Intel Labs, Intel Corporation, Hillsboro, OR
Brain-Inspired Computing Lab

EDUCATION

2013 – 2019: M.S., Ph.D. in Neurosciences (Computational specialization)
University of California, San Diego, La Jolla, CA
Thesis committee: John Serences (advisor), Ed Vul, Tim Gentner, Ed Callaway, Doug Nitz

2007 – 2011: B.A. with High Honors, double major in Biology & Cognitive Science
Swarthmore College, Swarthmore, PA
Advisors: Kathleen K. Siwicky & Frank H. Durgin

PREPRINTS

Pink, M., **Vo, V.A.**, Wu, Q., Mu, Q., Turek, J.S., Hasson, U., Norman, K.A., Michelmann, S., Huth, A., Toneva, M. (2024). Assessing episodic memory in LLMs with Sequence Order Recall Tasks. *arXiv:2410.08133*. Under review; also presented by J. Turek at *Workshop for LatinX in AI, Neural Information Processing Systems (NeurIPS)*.

Vo, V.A.*, Jain, S.*, Beckage, N., Chien, H.Y.S., Obinwa, C., Huth, A.G. (2023). A unifying computational account of temporal context effects in language across the human cortex. *bioRxiv (under review)*

Chen, L., Ahmed, N.K., Dutta, A., Bhattacharjee, A., Yu, S., Mahmud, Q.I., Abebe, W., Phan, H., Sarkar, A., Butler, B., Hasabnis, N., Oren, G., **Vo, V.A.**, Munoz, J.P., Willke, T.L., Mattson, T., Jannesari, A. (2024). The Landscape and Challenges of HPC Research and LLMs. *arXiv:2402.02018*

PUBLICATIONS

Racah, O., Chen, P., Gureckis, T., Poeppel, D., **Vo, V.** (2024). The “Naturalistic Free Recall” Dataset: four stories, hundreds of participants, and high-fidelity transcriptions. *Scientific Data*.

Leto, A., Aguerrebere, C., Bhati, I., Willke, T., Tepper, M., **Vo, V.A.** (2024). Toward optimal search and retrieval for RAG. *arXiv:2411.07396*. *2nd Workshop on Attributing Model Behavior at Scale, Neural Information Processing Systems (NeurIPS)*.

Kadosh, T., Hasabnis, N., Soundararajan, P., **Vo, V.A.**, Capotă, M., Ahmed, N., Pinter, Y., Oren, G. (2024). OMPar: Automatic Parallelization with AI-Driven Source-to-Source Compilation. *arXiv:2409.14771*. *Workshop on ML for Systems, Neural Information Processing Systems (NeurIPS)*.

Kadosh, T., Hasabnis, N., **Vo, V.A.**, Schneider, N., Krien, N., Capotă, M., Wasay, A., Tamir, G., Willke, T., Ahmed, N., Pinter, Y., Mattson, T., Oren, G. (2024). MonoCoder: Domain-Specific Code Language Model for HPC Codes and Tasks. **Outstanding paper award**, *IEEE High Performance Extreme Computing*.

- Schneider, N., Hasabnis, N., **Vo, V.A.**, Kadosh, T., Krien, N., Capotă, M., Tamir, G., Willke, T., Ahmed, N., Pinter, Y., Mattson, T., Oren, G. (2024). MPIrigen: MPI Code Generation through Domain-Specific Language Models. *Workshop on AI for Systems, High Performance Parallel and Distributed Computing (HPDC)*.
- Chen, L., Bhattacharjee, A., Ahmed, N., Hasabnis, N., Oren, G., **Vo, V.A.**, Jannesari, A. (2024). OMPGPT: A Generative Pre-trained Transformer Model for OpenMP. *International European Conference on Parallel and Distributed Computing (Euro-Par)*.
- Ma, G., **Vo, V.A.**, Willke, T.L., Ahmed, N.A. (2023). Memory-augmented Graph Neural Networks: A Brain-Inspired Review. *IEEE Transactions on Artificial Intelligence*.
- Ma, G., **Vo, V.A.**, Willke, T.L., Ahmed, N.A. (2023). Augmenting recurrent graph neural networks with a cache. *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*.
- Tang, J., Du, M., **Vo, V.A.**, Lal, V., Huth, A.G. (2023). Brain encoding models based on multimodal transformers can transfer across language and vision. *Neural Information Processing Systems (NeurIPS)*.
- Jain, S., **Vo, V.A.**, Wehbe, L., Huth, A.G. (2023). Computational language modeling and the promise of in silico experimentation. *Neurobiology of Language*.
- Racah, O., Chen, P., Willke, T., Poeppel, D., **Vo, V.A.** (2022). Memory in humans and deep language models: Linking hypotheses for model augmentation. *Workshop on Memory in Artificial and Real Intelligence (MemARI), Neural Information Processing Systems (NeurIPS)*.
- Ma, G., **Vo, V.A.**, Ahmed, N., Willke, T. (2022). Cache-memory gated graph neural networks. *Workshop on Memory in Artificial and Real Intelligence (MemARI), Neural Information Processing Systems (NeurIPS)*.
- Antonello, R., Turek, J.S., **Vo, V.A.**, Huth, A. Low-Dimensional Structure in the Space of Language Representations is Reflected in Brain Responses. (2021). *Neural Information Processing Systems (NeurIPS)*.
- Kumar, M., Anderson, M., Antony, J.W., Baldassano, C., Brooks, P., Cai, M., Chen, P-H.C., Ellis, C., Henselman-Petrusek, G., Huberdeau, D., Hutchinson, J.B., Li, Y.P., Lu, Q., Manning, J., Mennen, A., Nastase, S., Richard, H., Shapiro, A.C., Schuck, N., Shvartsman, M., Sundaram, N., Suo, D., Turek, J.S., **Vo, V.**, Wallace, G., Wang, Y., Zhang, H., Zhu, X., Capotă, M., Cohen, J., Hasson, U., Li, K., Ramadge, P.J., Turk-Browne, N., Willke, T., Norman, K.A. (2021). BrainIAK: The Brain Imaging Analysis Kit. *Aperture*.
- Vo, V.A.**, Sutterer, D.W., Foster, J.J., Sprague, T.C., Awh, E., Serences, J.T. (2021). Shared representational formats for information maintained in working memory and information retrieved from long-term memory. *Cerebral Cortex*.
- Chien, H-Y.S., Beckage, N.M., **Vo, V.A.**, Turek, J.S., Honey, C., Willke, T.L. (2021). Long short-term memory with slower information decay. *Workshop for LatinX in AI, International Conference on Machine Learning (ICML)*.
- Mahto, S., **Vo, V.A.**, Turek, J.S., Huth, A.G. (2021). Multi-timescale representation learning in LSTM language models. *International Conference on Learning Representations (ICLR)*.
- Jain, S., **Vo, V.**, Mahto, S., LeBel, A., Turek, J., Huth, A. (2020). Interpretable multi-timescale models for predicting fMRI responses to continuous natural speech. *Neural Information Processing Systems (NeurIPS)*.
- Turek, J., Jain, S., **Vo, V.**, Capotă, M., Huth, A., Willke, T. (2020). Approximating stacked and bidirectional recurrent architectures with the delayed recurrent neural network. *International Conference on Machine Learning (ICML)*.

Itthipuripat, S.I.*, **Vo, V.A.***, Sprague, T.C., Serences, J.T. (2019). Value-driven attentional capture enhances distractor representations in early visual cortex. *PLOS Biology*.

Henderson, M.H.*, **Vo, V.A.***, Chunharas, C., Sprague, T.C., Serences, J.T. (2019). Multivariate analysis of BOLD activation patterns recovers graded depth representations in human visual and parietal cortex. *eNeuro*.

Sprague, T.C.*, Adam, K.C.S.*, Foster, J.J.*, Rahmati, M.*, Sutterer, D.W.*, **Vo, V.A.*** (2018). Inverted encoding models assay population-level stimulus representations, not single-unit neural tuning. *eNeuro*.

Sprague, T.C., Itthipuripat, S., **Vo, V.A.**, and Serences, J.T. (2018). Dissociable signatures of visual salience and behavioral relevance across attentional priority maps in human cortex. *Journal of Neurophysiology*.

Vo, V.A., Sprague, T.C., and Serences, J.T. (2017). Spatial tuning shifts increase the discriminability and fidelity of population codes in visual cortex. *Journal of Neuroscience*.

Vo, V.A., Li, R., Kornell, N., Pouget, A., Cantlon, J.F. (2014). Young children bet on their numerical skills: Metacognition in the numerical domain. *Psychological Science*.

* These authors made equal contributions.

SELECTED ACADEMIC TALKS

"Building explicitly multi-timescale artificial neural networks that explain natural language processing in the human brain." (2022) Talk at Computational and Systems Neuroscience (COSYNE) Workshop on *Mechanisms, functions, and methods for diversity of neuronal and network timescales*.

"Memory systems in cognitive neuroscience and machine learning." (2019) Invited talk for University of Oregon cognitive neuroscience group, Eugene, OR.

"Neural representations of spatial position recalled from long-term and short-term memory diverge across the cortical hierarchy." (2017) Talk at Vision Sciences Society meeting, St. Pete Beach, FL.

"Spatial attention modulates voxel receptive fields to boost the fidelity of multi-voxel stimulus representations." (2016) Nanosymposium talk at Society for Neuroscience meeting, San Diego, CA.

PROFESSIONAL ACTIVITIES

Workshops: **Co-organizer:** ICLR "How can findings about the brain improve AI systems?" (2021), NeurIPS "Memory in Artificial and Real Intelligence" (2022); **Panel member:** SIGKDD "LLM Day" (2023); **Proposal reviewer:** NeurIPS (2024); **Program committee (paper reviewer):** NeurIPS "Context and compositionality in biological and artificial neural systems" (2019), NeurIPS "Gaze Meets ML" (2022), ICLR "Representational Alignment" (2024), Brain Informatics (2024)

Paper reviewer: **Computer science:** NeurIPS, ICML, ICLR, AAAI, IMWUT (Interactive, Mobile, Wearable, and Ubiquitous Technologies); **Neuroscience/Psychology:** Cerebral Cortex, eNeuro, Journal of Cognitive Neuroscience, NeuroImage, Proceedings of the National Academy of Sciences, PLOS Computational Biology, Learning and Memory

Intel activities: **Organizer:** Women at Intel Labs (2022-2024), IL Equity & Inclusion Taskforces (2023-2024);

Presenter: Intel Labs Open House (2019, 2021, 2023), Intel Spoken Language Technologies Summit (2021)

Courses: Attended "Computational Neuroscience: Vision" summer course, Cold Spring Harbor Laboratory (2016)

SELECTED PRESENTATIONS

- Chen, P.*, Raccach, O.*, Gureckis, T.M., Poeppel, D., **Vo, V.A.** (2024). Automatic detection of spontaneous false memories in the naturalistic recall of narratives. Poster at NYU Minds, Brains, and Machines.
- Toneva, M., **Vo, V.**, Turek, J., Jain, S., Michelmann, S., Capotă, M., Huth, A., Hasson, U., Norman, K. (2022). Language models that can remember. Poster at Context and Episodic Memory Symposium.
- Vo, V.***, Turek, J., Capotă, M., Willke, T. (2022). Using neuroscience to improve long-range sequential processing in language models. Poster at From Neuroscience to Artificially Intelligent Systems (NAISys) meeting at Cold Spring Harbor Labs.
- Jain, S.*, **Vo, V.***, Beckage, N.M., Chien, H.Y.S., Obinwa, C., Huth, A.G. (2021). A unifying computational account of temporal processing in natural speech across cortex. Poster slide slam at Society for Neurobiology of Language (SNL) meeting.
- Vo, V.**, Beckage, N., Capotă, M., Turek, J. Brain-inspired multi-timescale language models. (2021). Intel Labs Open House demo.
- Vo, V.A.**, Serences, J.T. (2018). The effects of attentional scope on voxel receptive fields and population codes for space. Poster at Vision Sciences Society meeting, St. Pete Beach, FL.
- Itthipuripat, S., **Vo, V.A.**, Sprague, T.C., Serences, J. (2017). Reward and selection history shape neural representations of an attentional priority in human visual and parietal cortex. Poster at Society for Neuroscience meeting, Washington, D.C.
- Itthipuripat, S., Chang, K., **Vo, V.**, Serences, J. (2017). Dissociable effects of stimulus strength, task demands, and training on occipital and parietal EEG signals during perceptual decision-making. Talk at Vision Sciences Society meeting, St. Pete Beach, FL.
- Sprague, T.C., Itthipuripat, S., **Vo, V.A.**, and Serences, J.T. (2016). Graded representations of stimulus salience and attentional priority across visually-responsive cortex. Nanosymposium talk at Society for Neuroscience meeting, San Diego, CA.
- Vo, V.A.**, Herrera, E.I., Serences, J.T. (2016). Orientation selective responses as measured with EEG track both featural and temporal attention enhancements. Poster at Vision Sciences Society meeting, St. Pete Beach, FL.
- Henderson, M., Chunharas, C., **Vo, V.**, Sprague, T., Serences, J. (2016). Reconstructing 3D stimuli using BOLD activation patterns recovers hierarchical depth processing in human visual and parietal cortex. Poster at Vision Sciences Society meeting, St. Pete Beach, FL.
- Vo, V.A.**, Sprague, T.C., Serences, J.T. (2015). Linking attentional modulations of single-voxel population receptive fields and region-level spatial reconstructions. Poster at Society for Neuroscience meeting, Chicago, IL.
- Emerson, R., **Vo, V.A.**, Kurtz, T., Cantlon, J.F. (2014). Mathematics expertise predicts structural and functional variability in the intraparietal sulcus. Nanosymposium talk at Society for Neuroscience meeting, Washington, D.C.
- Emerson, R., **Vo, V.**, Cantlon, J.F. (2013). Longitudinal changes in children's IPS responses are number-specific and mathematics related. Poster at Cognitive Neuroscience Society meeting, San Francisco, CA.

Vo, V., Li, R., Kornell, N., Cantlon, J.F. (2012). Metacognition in children is specific to domain knowledge. Poster at Cognitive Science Society meeting, Sapporo, Japan.

Vo, V., Ning, A., Bhattacharjee, A., Li, Z., Durgin, F. (2011). Pointing accurately at a target doesn't require perceiving its location accurately. *Journal of Vision*, 11(11):944. Poster at Vision Sciences Society meeting, Naples, FL.

MENTORSHIP & TEACHING

- 2024 *Mentor/collaborator* for Shailee Jain (UCSF, UT Austin) to win Glushko Dissertation Prize, Jerry Tang (UT Austin) for NIH F32 award; *Intern mentor* for Alex Leto (CU Boulder), resulting in NeurIPS workshop paper.
- 2022 *Intern mentor* for Omri Raccach (NYU), resulting in NeurIPS workshop paper and NIH F32 award.
- 2021 *Project mentor* for Neuromatch Academy (computational neuroscience). Project extension with Hanxiao Lu and Andréanne Proulx, resulting in Neuromatch 4.0 main track talk "Untangling contributions of distinct features of images to object processing in inferotemporal cortex".
- 2013 – 2019 *Graduate student mentor* for independent study students and volunteers (UCSD) in Psychology and Neuroscience. Isabel Asp; Kaylee Craig; Eduardo Herrera; Avery Rogers; Matthew Jaconetta; Kia Shams; Wenjing Dong; Rie Davis; Naomi Lee.
- Spring 2016 *Student Lecturer*, Fundamentals in Statistics and Computation for Neuroscientists graduate course (UCSD; entirely student-taught). Prepared two video lectures and wrote accompanying quizzes and problem sets on basic hypothesis testing and permutation testing.
- Spring 2015 *Teaching Assistant*, Data Analysis in MATLAB graduate course (UCSD; Prof. John Serences). Reviewed student code & algorithmic solutions to weekly problem sets on advanced topics in data analysis, such as bootstrapping & permutation statistics, time-frequency analysis, pattern classification, and nonlinear curve & surface fitting.
- Spring 2015 *Guest Lecturer*, Sensation & Perception undergraduate course (UCSD Psychology; Prof. John Serences). Lectured on sensory development.
- 2014 - 2015 *Teaching Assistant*, Neurosciences Graduate Program Boot Camp for incoming students (taught by Stefan Leutgeb & Jing Wang). Led computational neuroscience workshops & labs designed to familiarize students with programming & data analysis techniques using Python and MATLAB.
- 2011 – 2013 *Graduate student mentor* for independent study students (University of Rochester). Laura Ackerman (Brain & Cognitive Sciences), undergraduate thesis; Emily Kasman (Brain & Cognitive Sciences); Matthew Mullen (Neurosciences).

AWARDS AND HONORS

- Intel Labs Intelligent Systems Research Division Recognition Award (2024)
Outstanding paper, *IEEE High Performance Extreme Computing* (2024)
Intel Labs Emergent AI Research Division Recognition Award (2021)
Intel Innovator High 5 Award (2020) for filing 5+ patent applications
Intel Labs Emerging Technologies Research Division Recognition Award (2020)
Vision Sciences Society Student Travel Award (2017)
Center for Visual Sciences Symposium Student Travel Award (2016)
NSF Graduate Research Fellowship (awarded 2013)
Howard Hughes Medical Institute Summer Research Fellow (2010)