Cory Shain

Curriculum Vitae

Contact

Cory A. Shain
43 Vassar Street
Bldg 46-3037
Cambridge, MA 02139
617-253-5748
cory.shain@gmail.com
https://coryshain.github.io/

Bio

I am an Assistant Professor in the Department of Linguistics at Stanford. I study language processing in the mind and brain using experimental and computational approaches. I received my BA (2009), MA (2009), and PhD (2021) in linguistics from Ohio State, where I studied under William Schuler and Micha Elsner, and I was a postdoc in Brain & Cognitive Sciences with Evelina Fedorenko at MIT.

Positions

2024 – present. Assistant Professor. Linguistics, Stanford University
2021 – 2024. Postdoctoral researcher, Brain & Cognitive Sciences, MIT

Education

2016 – 2021. PhD in Linguistics, The Ohio State University. Dissertation title: *Language, time, and the mind: Understanding human language processing using continuous-time deconvolutional regression*.
2009. MA in Linguistics, The Ohio State University
2009. BA in Linguistics, minor in International Studies, The Ohio State University

Preprints

Reilly, J.; **Shain, C.**; et al (under review). What we mean when we say semantic: A consensus statement on the nomenclature of semantic memory. *PsyArXiv*. https://osf.io/preprints/psyarxiv/xrnb2.

Bruffaerts R.; Pongos A.; **Shain, C.**; Lipkin, B.; Siegelman, M.; Wens, V.; Sjøgård, M.; Pantazis, D.; Blank, I.; Goldman, S.; De Tiège, X.; Fedorenko E. (under review). Functional identification of language-responsive channels in individual participants in MEG investigations. *bioRxiv*. https://www.biorxiv.org/content/10.1101/2023.03.23.533424v1.

Journal Articles

Shain, C.; Kean, H.; Lipkin, B.; Affourti, J.; Siegelman, M.; Mollica, F.; Fedorenko, E. (in press). Distributed sensitivity to syntax and semantics throughout the human language network. *Journal of Cognitive Neuroscience*.

Shain, C.; and Schuler, W. (2024). A deep learning approach to analyzing continuous-time cognitive processes. *Open Mind*, 8: 235-264.

Shain, C. (2024). Word frequency and predictability effects dissociate in naturalistic reading. *Open Mind*, 8: 177-201.

Shain, C.; Meister, C.; Pimentel, T.; Cotterell, R.; Levy, R. (2024). Large-scale evidence for logarithmic effects of word predictability on reading time. *Proceedings of the National Academy of Sciences*, 121(10): e2307876121.

Shain, C.; Paunov, A.; Chen, X.; Lipkin, B. ; Fedorenko, F. (2023). No evidence of theory of mind reasoning in the human language network. *Cerebral Cortex*.

Shain, C.; Blank, I.; Fedorenko, E.; Gibson, E.; Schuler, W (2022). Robust effects of working memory demand during naturalistic language comprehension in language-selective cortex. *Journal of Neuroscience*, 42 (39): 7412-7430.

Fedorenko, E.; and **Shain, C.** (2021). Similarity of computations across domains does not imply shared implementation: The case of language comprehension. *Current Directions in Psychological Science*, 30(6): 526-534.

Shain, C.; and Schuler, W. (2021). Continuous-time deconvolutional regression for psycholinguistic modeling. *Cognition*, 215: 104735.

Wehbe, L.; Blank, I.; **Shain, C.**; Futrell, R.; Levy, R.; von der Malsburg, T.; Smith, N.; Gibson, E.; and Fedorenko, E. (2021). Incremental language comprehension difficulty predicts activity in the language network but not the multiple demand network. *Cerebral Cortex*, 31(9): 4006-4023.

Shain, C.; Blank, I.; van Schijndel, M.; Fedorenko, E.; and Schuler, W. (2020). fMRI reveals language-specific predictive coding during naturalistic sentence comprehension. *Neuropsychologia*, 138: 107307.

Elsner, M.; Sims, A. D.; Erdmann, A.; Hernandex, A.; Jaffe, E.; Jin, L.; Johnson, M. B.; Karim, S.; King, D. L.; Lamberti Nunes, L.; Oh, B.; Rasmussen, N.; **Shain, C.**; Antetomaso, S.; Dickinson, K. V.; Diewald, N.; McKenzie, M.; and Stevens-Guille, S (2019). Modeling morphological learning, typology, and change: What can the neural sequence-to-sequence framework contribute? *Journal of Language Modeling*, 7(1): 53-98.

Shain, C., and Tonhauser, J. (2010). The synchrony and diachrony of differential object marking in Paraguayan Guaraní. *Language Variation and Change*, 22(03), 321–346.

Conference Papers

Tuckute G., Sathe A., Wang M., Yoder H., **Shain C.**, Fedorenko E. (2022). SentSpace: Large-scale benchmarking and evaluation of text using cognitively motivated lexical, syntactic, and semantic features. *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: System Demonstrations*.

Shain, C. (2021). CDRNN: Discovering complex dynamics in human language processing. In *Proceedings of the Joint Conference of the 59th Annual Meeting of the Association for*

Computational Linguistics and the 11th International Joint Conference on Natural Language Processing.

Shain, C.; and Elsner, M. (2020). Acquiring language from speech by learning to remember and predict. In *Proceedings of the 24th Conference on Computational Natural Language Learning*: 195-214. **Best Paper Award.**

Jaffe, E.; **Shain C.**; and Schuler, W. (2020). Coreference information guides human expectations during natural reading. In *Proceedings of the 28th International Conference on Computational Linguistics*: 4587-4599.

Shain, C. (2019). A large-scale study of the effects of word frequency and predictability in naturalistic reading. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*: 4086-4094.

Shain, C.; and Elsner, M. (2019). Measuring the perceptual availability of phonological features during language acquisition using unsupervised binary stochastic autoencoders. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*: 69-85.

Shain, C.; and Schuler, W. (2018). Deconvolutional time series regression: A technique for modeling temporally diffuse effects. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*: 2679-2689.

Shain, C.; van Schijndel, M.; and Schuler, W. (2018). Deep syntactic annotations for broad-coverage psycholinguistic modeling. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*.

Jaffe, E.; **Shain, C.**; and Schuler, W. (2018). Coreference and Focus in Reading Times. In *Proceedings of the 8th Workshop on Cognitive Modeling and Computational Linguistics (CMCL 2018)*: 1-9.

Elsner, M.; **Shain, C.** (2017). Speech segmentation with a neural encoder model of working memory. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*: 1070-1080.

Mahler, T.; Cheung, W.; Elsner, M.; King, D.; de Marneffe, M.; **Shain, C.**; Stevens-Guille, S.; and White, M. (2017). Breaking NLP: Using morphosyntax, semantics, pragmatics and world knowledge to fool sentiment analysis systems. In *Proceedings of the First Workshop on Building Linguistically Generalizable NLP Systems*: 33-39.

Shain, C.; Bryce, W.; Jin, L.; Krakovna, V.; Doshi-Velez, F.; Miller, T.; Schuler, W.; and Schwartz, L. (2016). Memory-Bounded Left-Corner Unsupervised Grammar Induction on Child-Directed Input. In *Proceedings of the 26th International Conference on Computational Linguistics*: 964-975.

Shain. C.; van Schijndel, M.; Futrell, R.; Gibson, E.; and Schuler, W. (2016). Memory access during incremental sentence processing causes reading time latency. In *Proceedings of the Workshop on Computational Linguistics for Linguistic Complexity (CL4LC)*: 49-58.

Peer-reviewed conference abstracts

Shain, C.; Meister, C.; Pimentel, T.; Cotterell, R.; Levy, R. (2023). Large-scale evidence for logarithmic effects of word predictability on reading time. In *36th Annual Conference on Human Sentence Processing*. **(plenary talk)**

Shain, C.; Blank, I.; Fedorenko, E.; Gibson, E.; Schuler, W (2023). Robust effects of working memory demand during naturalistic language comprehension in language-selective cortex. In *36th Annual Conference on Human Sentence Processing*.

Shain, C.; Paunov, A.; Chen, X.; Lipkin, B. ; Fedorenko, F. (2023). No evidence of theory of mind reasoning in the human language network. In *36th Annual Conference on Human Sentence Processing*.

Shain C. (2023). Lexical retrieval and probabilistic inference dissociate in naturalistic reading. In *36th Annual Conference on Human Sentence Processing*.

Shain, C.; Blank, I.; Fedorenko, E.; Gibson, E.; Schuler, W (2021). Robust effects of working memory demand during naturalistic language comprehension in language-selective cortex. In *13th Annual Meeting of the Society for the Neurobiology of Language*.

Shain, C.; and Elsner, M. (2021). Modeling effects of incremental memory and prediction pressures on phoneme learning from speech. In *34th Annual CUNY Conference on Human Sentence Processing*.

Shain, C.; and Schuler, W. (2021). Analyzing complex human sentence processing dynamics with CDRNNs. In *34th Annual CUNY Conference on Human Sentence Processing*.

Tuckute, G.; **Shain, C.**; Blank, I.; Wang, M.; and Fedorenko, E. (2020). Linguistic and conceptual processing are dissociated during sentence comprehension. In *12th Annual Meeting of the Society for the Neurobiology of Language*.

Mollica, F.; **Shain, C.**; Affourtit, J.; Kean, H.; Siegelman, M.; and Fedorenko, E. (2020). Another look at the constituent structure of sentences in the human brain. In *12th Annual Meeting of the Society for the Neurobiology of Language*.

Shain, C.; Blank, I.; van Schijndel, M.; Fedorenko, E.; and Schuler, W. (2020). Prediction in the language network is sensitive to syntactic structure. In *33rd Annual CUNY Conference on Human Sentence Processing*. **(plenary talk)**

Shain, C.; and Schuler, W. (2020). Controlling for human response latency with continuous-time deconvolutional regression. In *33rd Annual CUNY Conference on Human Sentence Processing*.

Jaffe, E.; **Shain, C.**; and Schuler, W. (2020). Coreference information guides human expectations during naturalistic reading. In *33rd Annual CUNY Conference on Human Sentence Processing*.

Shain, C. (2019). A large-scale deconvolutional study of predictability and frequency effects in naturalistic reading. In *32nd Annual CUNY Conference on Human Sentence Processing*.

Shain, C.; and Elsner, M. (2019). Computer modeling suggests patterns of perceptual availability of phonological structure during infant language acquisition. In *32nd Annual CUNY Conference on Human Sentence Processing*.

Shain, C.; Blank, I.; van Schijndel, M.; Schuler, W.; and Fedorenko, E. (2019). fMRI reveals language-specific predictive coding during naturalistic sentence comprehension. In *32nd Annual CUNY Conference on Human Sentence Processing*.

Jaffe, E.; **Shain, C.**; and Schuler, W. (2018). Coreference and focus in broad-coverage stimuli. In *31st Annual CUNY Conference on Human Sentence Processing*.

Shain, C.; and Schuler, W. (2018). Modeling psycholinguistic effect timecourses with deconvolutional time series regression. In *31st Annual CUNY Conference on Human Sentence Processing*.

Shain, C.; Futrell, R.; van Schijndel, M.; Gibson, E.; Schuler, W.; and Fedorenko, E. (2018). Evidence of semantic processing difficulty in naturalistic reading. In *31st Annual CUNY Conference on Human Sentence Processing*.

Elsner, M.; and **Shain, C.** (2017). Word segmentation with neural net working memory. In *Midwest Speech and Language Days 2017*.

Shain, C.; van Schijndel, M.; Futrell, R.; Gibson, E.; and Schuler, W. (2017). Retrieving structures from memory causes difficulty during incremental processing. In *30th Annual CUNY Conference on Human Sentence Processing*. **(plenary talk)**

Shain, C.; Bryce, W.; Jin, L.; Krakovna, V.; Doshi-Velez, F.; Miller, T.; Schuler, W.; and Schwartz, L. (2017). Modeling syntax acquisition via cognitively-constrained unsupervised grammar induction. In *30th Annual CUNY Conference on Human Sentence Processing*.

Shain, C.; van Schijndel, M.; Gibson, E.; and Schuler, W. (2016). Exploring memory and processing through a gold-standard annotation of Dundee. In *29th Annual CUNY Conference on Human Sentence Processing*.

Theses

Shain, C. (2021). *Language, time, and the mind: Understanding human language processing using continuous-time deconvolutional regression*. PhD Thesis: The Ohio State University.

Shain, C. (2009). *The distribution of differential object marking in Paraguayan Guaraní*. MA Thesis: The Ohio State University. (Available at https://etd.ohiolink.edu/!etd.send_file?accession=osu1243450139).

Unpublished works

A Sketch of the Phonology and Morphology of Iyasa. (Available at http://go.osu.edu/iyasaphon).

Iyasa Webonary. 1700+ entries. French and English interfaces. (Available at http://go.osu.edu/ykowebonary).

Invited talks

Uncovering the Algorithmic Foundations of Language Learning and Processing. Department of Cognitive, Linguistics and Psychological Sciences, Brown University. Sep 26, 2023.

Uncovering the Algorithmic Foundations of Language Learning and Processing. Language Imaging Laboratory, Medical College of Wisconsin. Mar 23, 2023.

Uncovering the Algorithmic Foundations of Language Learning and Processing. Department of Linguistics and Halıcıoğlu Data Science Institute, University of California San Diego. Mar 21, 2023.

Uncovering the Algorithmic Foundations of Language Learning and Processing. Department of Cognitive Science, University of California San Diego. Mar 20, 2023.

Uncovering the Algorithmic Foundations of Language Learning and Processing. Department of Cognitive Science, Johns Hopkins University. Mar 6, 2023.

Uncovering the Algorithmic Foundations of Language Learning and Processing. Department of Language Studies, University of Toronto Scarborough. Feb 28, 2023.

Uncovering the Algorithmic Foundations of Language Learning and Processing. Department of Linguistics, University of Florida. Feb 13, 2023.

Uncovering the Algorithmic Foundations of Language Learning and Processing. Neuroscience Institute and Department of Psychology, Carnegie Mellon University. Feb 6, 2023.

Uncovering the Algorithmic Foundations of Language Learning and Processing. Department of Linguistics, Stanford University. Feb 3, 2023.

Uncovering the Algorithmic Foundations of Language Learning and Processing. Department of Linguistics, University of Illinois Urbana-Champaign. Jan 23, 2023.

Incremental story comprehension in the human brain. Language and Cognition talk series, Harvard University. Sep 20, 2022.

Incremental story comprehension in the human brain (Remote). Department of Linguistics, Cornell University. Sep 16, 2022.

Language, Time, and the Mind. Department of Linguistics and Cognitive Science. University of Delaware. Feb 14, 2021.

Prediction and memory in human language comprehension: Evidence from naturalistic fMRI (Online). Centre for Language Studies, Radboud University. Dec 7, 2020.

Studying how language comprehension unfolds over time (Remote). Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology. Sep 16, 2020.

Localizing incremental linguistic prediction in the mind (Remote). Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology. May 7, 2019.

Discovering psycholinguistic effect timecourses with deconvolutional time series regression. Department of Cognitive Science, Johns Hopkins University. Nov 7, 2018.

Unsupervised machine learning as acquisition modeling. Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology. Aug 16, 2017.

Media appearances

Episode 16 of the Language Neuroscience Podcast: Robust effects of working memory demand during naturalistic language comprehension in language-selective cortex. Oct 4, 2021.

Honors and awards

2017. NSF Graduate Research Fellowship Program Honorable Mention

Grants and fellowships

2016. University Fellowship, The Ohio State University

Teaching

2024, Fall. LINGUIST 145/245A, PSYCH 140. Introduction to Psycholinguistics. Stanford University.
2018, Spring. LING 2000, Introduction to Linguistics. Ohio State University
2017, Autumn. LING 2000, Introduction to Linguistics. Ohio State University
2014-2016, All terms. EXP 1100, University Exploration Survey (orientation and study skills course). Ohio State University

Service

Journal reviewing:

Attention, Perception, and Psychophysics
Brain and Language
Brain Structure and Function
Cerebral Cortex
Cognition
Cognitive Science
Computational Linguistics
Computer Speech and Language
European Journal of Neuroscience
Imaging Neuroscience
iScience
Journal of Cognitive Neuroscience
Journal of Memory and Language
Journal of Neuroscience
Language and Speech
Language, Cognition and Neuroscience
Language Resources and Evaluation
Nature Communications
Nature Human Behavior
Nature Scientific Reports
Neurobiology of Language
NeuroImage
Neuron
Open Mind
Perspectives on Psychological Science
Phonology
Proceedings of the National Academy of Sciences
Psychonomic Bulletin and Review
Speech Communication
Transactions of the Association for Computational Linguistics

Conference reviewing:

AAAI Conference on Artificial Intelligence (AAAI)
ACL rolling review

Asia-Pacific Chapter of the Association for Computational Linguistics (AACL)
Association for Computational Linguistics (ACL), **outstanding reviewer (2020)**
Cognitive Modeling and Computational Linguistics (CMCL)
Conference on Computational Natural Language Learning (CoNLL)
Empirical Methods in Natural Language Processing (EMNLP)
European Chapter of the Association for Computational Linguistics (EACL)
International Conference on Computational Linguistics (COLING), **outstanding reviewer (2018)**
North American Association for Computational Linguistics (NAACL)
Meeting of the Cognitive Science Society (CogSci)
Meeting of the Society for Computation in Linguistics (SCiL)
Widening Natural Language Processing (WiNLP)