Zachary McCormick

+1 (847) 890-9492 **Personal:** zmmccormick3@gmail.com linkedin.com/zachary-mccormick
School: zmccorm1@jh.edu

Education

Johns Hopkins University Jul 2025 – May 2030 (anticipated)
*Ph.D. in Cognitive Science, advised by **Dr. Margaret Renwick** (primary) & **Dr. John Hale** (secondary)* *Baltimore, MD*

- **Courses I am taking include:** Deep Learning for Cognitive Neuroscience, Information Coding in Neural Activity, Neurolinguistics, Psycholinguistics
- **Courses for which I am serving as a teaching assistant:** Acoustic Phonetics (Fall 2025)

Brandeis University Aug 2023 – May 2025
*M.S. in Computational Linguistics (GPA: 3.98 / 4.00), capstone advised by **Dr. James Pustejovsky*** *Waltham, MA*

- **Courses taken include:** Principles of Neuroscience Research, Applied Bayesian Modeling (R), Computational Semantics (Haskell), Advanced Machine Learning for Natural Language Processing (Python), Introduction to Research for Computational Linguistics, Computational Neuroscience (MATLAB), Data Structures and Algorithms (Java), Visual Cognition, Theory of Computation, Human Neuropsychology, Automatic Speech Recognition, Mathematical Logic
- **Courses for which I served as a course assistant:** Morphology (Spring 2025), Semantics I (Spring 2025), Syntax I (Fall 2024), Historical Linguistics and Language Change (Spring 2024), Phonology II: Optimality Theory (Fall 2023)

Purdue University Jun 2021 – May 2023
*B.A. in Linguistics, Spanish, and Italian Studies (GPA: 4.00 / 4.00), capstone advised by **Dr. Elena Coda*** *West Lafayette, IN*

- **Courses taken include:** Anatomy and Physiology of the Speech Mechanism; Corpus Linguistics (Python); Cognitive Psychology; Phonology; Syntax; The Indo-European Language Family; Spanish Phonetics, Phonology, and Dialectology; Spanish Morphosyntax
- **Courses for which I tutored:** Analytic Geometry and Calculus I, Beginning Italian

Experience

Johns Hopkins University Aug 2025 (i.e., presently)
Teaching Assistant *Baltimore, MD*

Brandeis University Jun 2024 – May 2025
Research Assistant *Waltham, MA*

- Assisted in publishing the Bilingual Russian Child Speech (BiRCh) corpus, ensuring adherence to NSF outcomes and maintaining documentation for research outputs.
- Managed the organization and publication of audio files, ELAN transcripts, and FoLiA files.
- Created a mapping from transcription conventions and disfluency annotations to CHAT format.

Babel Street Feb 2024 – Jun 2024
Natural Language Annotator *Remote*

- Annotated spans of natural language in journal articles for legal event extraction.
- Met with colleagues weekly to ensure sufficient inter-annotator agreement and mutual understanding of the ontology employed.

Brandeis University Aug 2023 – May 2025
Course Assistant *Waltham, MA*

- Held weekly office hours to respond to students' questions and concerns regarding the course material and assignments.
- Graded weekly homework assignments and provide general comments to the professor concerning common mistakes in students' responses.

Purdue University Sep 2022 – May 2023
School of Languages and Culture (SLC) Intern *West Lafayette, IN*

- Planned, organized, and coordinated events for current and prospective SLC students interested in learning more about the language course offerings, study abroad opportunities, and other incentives and activities run by the SLC.
- Made flyers, posters, and sign-up sheets for SLC-run events and advertised them on social media (i.e., Twitter, Instagram, and Facebook in particular).

Purdue University Aug 2022 – May 2023

Resident Assistant *West Lafayette, IN*

- Planned and coordinated various events for my residents at least once a week and partnered with campus resources and cultural centers (e.g., the LGBTQ Center, University Residences, the Native American Educational And Cultural Center) in the process.
- Was part of the Diversity, Equity, and Inclusion (DEI) Committee and devised solutions with fellow RAs to make the Purdue campus a safer and more welcoming space for every resident.
- Led community meetings and attended weekly staff meetings to gauge the success of recent RA-led initiatives and events.

Purdue University Mar 2022 – May 2022

Research Assistant *West Lafayette, IN*

- Worked under the supervision of a Linguistics PhD student to mark intonation contours (e.g., HH, HL, LH, LL) at the ends of Spanish utterances produced by native Peninsular Spanish speakers using Praat software.
- Held weekly meetings to discuss the progress of the research task and evaluating inter-annotator agreement in intonation contour markings.

Mathnasium of South Elgin Jul 2019 – Jul 2021

Lead Instructor *South Elgin, IL*

- Tutored students (anywhere from kindergarten- to college-age) in a vast array of mathematics topics, ranging from basic arithmetic to calculus.
- Designed personalized lesson plans to address individual student learning styles and challenges, improving engagement and comprehension.
- Kept a daily log of each student's progress and their completed assignments during each session.

Projects

"Hljóð and clear": Estimation, noise robustness, and visualization of Icelandic phonological posteriors | *Python*

- Estimated phonological posteriors from Icelandic child speech, using data from the Samrómur Children Corpus and phonetic alignments from WebMAUS using a bidirectional recurrent neural network (BRNN) model.
- Conducted systematic evaluations of model robustness across varying types and levels of additive noise, analyzing feature-level degradation patterns.
- Built *Hljóðasýn* using Streamlit, an interactive Python-based tool for visualizing and analyzing frame-level phonological posterior probabilities.
- Executed end-to-end pipeline including data preprocessing, forced alignment, model training, evaluation, and visualization for Icelandic speech processing research.

Retrieval-augmented generation (RAG) for science abstracts | *Python*

- Designed and implemented a Retrieval-Augmented Generation (RAG) pipeline for question answering over scientific corpora, integrating abstracts from PubMed and arXiv.
- Conducted comparative evaluation of two LLM generators—FLAN-T5-Base (encoder-decoder) and GPT-Neo-1.3B (decoder-only)—highlighting architectural impacts on output quality.
- Demonstrated that GPT-Neo-1.3B produced more comprehensive and contextually relevant answers in scientific QA tasks, informing model selection in domain-specific RAG applications.
- Built a full-stack NLP system combining dense retrieval, context chunking, and prompt-based generation to support efficient synthesis of scientific knowledge.

Computational modeling of predictive coding in the auditory cortex (A1) – Literature Review

- Conducted an in-depth literature review on predictive coding in the auditory cortex, synthesizing current theories and computational models in relation to neural hierarchies.
- Identified key gaps in existing models, including limited integration of subcortical auditory structures and biologically implausible assumptions in cortical feedback mechanisms.
- Evaluated the biological fidelity of leading predictive coding frameworks, highlighting discrepancies between theoretical models and empirical neural data.
- Proposed future research directions emphasizing the need for multi-level models that incorporate both cortical and subcortical processing pathways.

"Flow motion": Navigating the currents of optical flow in computational and biological systems – Literature Review

- Conducted a critical literature review on optical flow estimation in computational and biological systems, emphasizing its role in visual motion perception.
- Evaluated the limitations of optical flow models in handling occlusions, non-rigid motion, and other complex scene dynamics.
- Assessed the extent to which computational models of optical flow align with biological mechanisms of motion perception, identifying gaps in biological plausibility.
- Highlighted the potential of probabilistic models to bridge the divide between engineered vision systems and cognitive neuroscience research on dynamic scene understanding.

Modeling competition via spike-timing-dependent plasticity (STDP) | *MATLAB*

- Simulated the membrane potential of a leaky-integrate and fire (LIF) neuron receiving input spike trains and updated their synaptic strengths using a batch method STDP rule.
- Trained the system and altered the phase offset across trials, plotting and analyzing the differences among them.
- Plotted and analyzed the differences in behavior between uncorrelated and correlated stimuli.

Headline metonymy recognition | *Label Studio*, *Python*, *BERT*

- Devised annotation guidelines and conducted internal annotations in Label Studio to facilitate the systematic tagging of headline metonymies.
- Leveraged a dataset annotated by three external annotators, assessing inter-annotator agreement and conducting adjudication to resolve inter-annotator discrepancies.

Sentiment classification of film reviews | *Python*, *scikit-learn*, *Pandas*

- Leveraged the IMDb Movie Reviews Dataset from the IEEE DataPort to conduct multiclass sentiment classification on 100,000 film reviews using models from the scikit-learn library.
- Improved baseline accuracy by 12.73 (percentage) points in employing bag of words features with a logistic regression model and learning rate of 0.01.

Ferzan Özpetek's gays and gaze: Redefining the cinematic queer gaze – Conference Presentation

- Extended Mulvey's principle of the male gaze to an alternative queer gaze as it is presented in the works of Turkish-Italian film director Ferzan Özpetek.
- Examined the psychoanalytic and social underpinnings of the queer gaze as a gaze of foresight, balance, careful calculation, and reclamation.

Coarticulatory effects of the lateral approximant in Colombian Spanish – Pilot Study

- Designed and executed a pilot phonetic study investigating coarticulatory effects of /l/ on adjacent vowels in bilingual Colombian Spanish-English speakers (n=4).
- Recorded structured utterances using Praat, ensuring consistent speech elicitation conditions across language modes (i.e., Spanish vs. English).
- Manually segmented and labeled acoustic data in Praat to extract second formant (F2) values from target segments.
- Interpreted results in light of language mode theory and bilingual phonetic interference, identifying trends for future large-scale study.

Technical Skills

Programming Languages and Frameworks: Python, Java, MATLAB, R, PyTorch, Pandas, scikit-learn
Natural Languages: English (native), Italian (proficient), Spanish (proficient), Portuguese (intermediate), French (intermediate), Icelandic (basic), German (basic), Arabic (basic), Persian/Farsi (basic), Indonesian (basic)
Transferable Skills: attention to detail, perspicuous communication, self-motivation, intercultural competence, problem analysis from multiple perspectives, passion for lifelong learning, mentorship